Chapter 2: Experimental Investigations

Most scientific investigations begin with an observation.

Observation	Led to
When elements are arranged in order of increasing atomic mass, their properties recurred periodically.	Mendeleev's periodic law, and ultimately the modern periodic table
The line emission spectrum of hydrogen was observed.	This was a key experimental observation that led Bohr to develop his model of the atom.

At its core, science is an attempt to answer questions. Scientists are naturally curious people. Many discoveries have been made by scientists wondering…

- *What* would happen if?
- *How* does that work?
- *Why* does that happen?
- What would happen if we *changed* variable X?

Scientists spend their lives asking – and attempting to answer – questions.

Good research questions are:

- Testable
- Measurable (the results can be quantified or observed)
- Specific (focus on a clear concept or variable)
- Objective (the questions should not reply on opinions or personal bias)
- Realistic (it must be practical, safe and ethical to investigate)
- Relevant to scientific concepts
- Open-ended (should not have a yes/no answer or be answerable by a single number or word)
- Suitably challenging

Bad research questions are:

- Untestable
- Vague (lack focus or clarity)
- Too broad (or too simple)
- Unrealistic (impractical, unsafe, unethical etc.)

Hypothesis

A hypothesis is a proposed explanation for an observation that can be tested by carrying out an experiment or further observation.

Once you have formulated your hypothesis, you should make a **prediction.** A prediction outlines the possible or expected result of an experiment that was designed and carried out to investigate the hypothesis.

A prediction is the expected outcome of a scientific test or experiment.

Designing and Planning Investigations

A fair test is an investigation in which only one variable is changed (the independent variable) and everything else is kept constant (controlled variables)

A variable is any factor or condition that can be changed during an experiment.

An independent variable is the variable that is deliberately **changed** during the experiment. It is the variable being tested.

The dependent variable is the variable that is **measured** during the experiment (to determine the effect of changing the independent variable).

Controlled variables are the variables that are **kept constant** throughout the experiment to ensure a fair test.

In conclusion, a fair test has one independent variable, one dependent variable and multiple controlled variables.

A control is a version of the experiment that is identical to the original experiment except that it does not include the independent variable.

The function of a control is to act as a comparison. Controls allow scientists to determine whether the changes observed in the original experiment are due to the independent variable or other factors.

Method

An experimental method must be **repeatable, precise** and presented in a logical **step-by-step sequence.** It should be so clear that another person could follow it exactly without needing to ask you any questions.

An important part of designing an experiment is **selecting suitable equipment**, e.g. a beaker is more suitable for stirring a solution with a glass rod, while a conical flask is more suitable for swirling.

Safety

Before beginning an experimental investigation, it is necessary to complete a risk assessment. Risks and procedures taken to minimise these risks must be identified. Often a table is used listing the hazard, risk, safety precaution and the response plan.

- Hazard: The chemical or apparatus
- Risk: What could go wrong?
- Safety precaution: How will you minimise the risk?
- Response plan: If the worst occurs, how will you respond?

Safety symbols

Corrosive: Can cause severe chemical burns to the skin and eyes, may corrode metals. Require careful handling to avoid contact with skin or inhalation of fumes. Safety precautions include wearing protective googles, gloves and clothing.

Toxic: Toxic substances pose a serious risk to health, potentially causing severe harm or death if swallowed, inhaled, or absorbed through the skin. Proper handling includes avoiding contact with skin or inhalation and using protective equipment such as gloves and masks.

<u>Harmful/ Irritant:</u> Can cause respiratory issues, skin irritation or allergic reactions. They should not be inhaled, swallowed or allowed to come into contact with skin or eyes. Protective measures include using gloves, eye protection and working in well-ventilated areas.

<u>Oxidising:</u> Can intensify fires or cause explosions when in contact with flammable materials. Should be kept away from heat and ignition sources. Rinse any contaminated skin and clothing immediately with water. Wear appropriate protective gear during handling.

<u>Flammable:</u> Can easily catch fire under normal conditions. Keep them away from heat, sparks and open flames. Store in tightly closed containers, keep cool and protect from sunlight.

<u>Hazardous to the environment:</u> Can be toxic to aquatic life and have long-lasting harmful effects on ecosystems. Prevent environmental contamination by avoiding release into waterways and collecting any spillage for proper disposal.

<u>Explosive:</u> May explode when exposed to heat or ignited in air. A sudden shock can also cause them to detonate, as can abrupt handling or the friction caused by dropping the container.

SDS (safety data sheets) are available online. If you are unsure about the risks associated with using specific chemicals, refer to these sheets (e.g. google 'SDS for ethanol').

<u>Data</u>

Data refers to the information collected during an experiment or investigation.

Anomalous data is a data point or result that deviates significantly from the rest of the dataset (or that deviates from the expected pattern or trend).

An anomaly (or outlier) is a value in a dataset that appears unusually large or small compared to the rest of the data.

If anomalies are identified during an investigation, it is advisable to repeat that part of the experiment.

Quantitative data

Quantitative data is data that can be represented numerically.

All numerical data is quantitative data, e.g. 33 °C, 5 m, 4.2 kg (it involves *quantities).*
Continuous data can have a **numerical value within a range** (including fractions, decimals etc.). This data is collected through **measurement.**

Discrete data also has numerical values, but the values must be **whole numbers**. This data is collected through **counting.**

Raw data refers to data collected directly from experiments; it has not been processed or altered in any way. It is the unrefined data that you first recorded, such as measurements, counts or observations made during the experiment.

Processed data refers to data that has been manipulated in some way to make it more usable for analysis. Processed data is often graphed, as graphs provide a clear visual representation of trends and patterns.

Qualitative data is non-numerical data that describes characteristics, qualities or categories.

Primary data is original data that you have collected yourself.

Secondary data is data that you have not collected yourself, i.e. data that someone else has collected.

Collecting and recording data

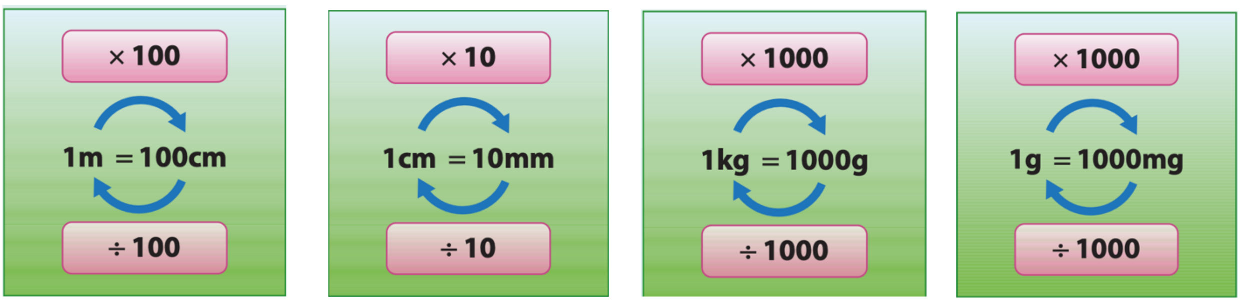
Data is often recorded in tables. It is good practice to draw out the data table before starting the experiment.

The headings in data tables should **include units**, e.g. 'Temperature (°C)' rather than just 'Temperature'. Typically, the **first column is the independent variable, with the dependent variable placed in the second.**

It is good practice to conduct more than one trial or test. If doing this, then the column for the dependent variable is divided into the number of tests carried out.

Units

Processing data often involves converting between different units.



Significant figures

When carrying out calculations involving large numbers, a good rule of thumb is to round answers to **four significant figures.** It is a good idea to round your final answer only (rounding answers too early in a multi-step calculation can lead to incorrect answers).

We start counting significant figures from the first non-zero digit. For example, these numbers all have four significant figures: 1821g, 0.1821 g and 0.01821 g. (Note that the number 0.018210 g has five significant figures; any zeros after the first non-zero number are significant).

Writing the molarity of a solution as 1.00 M means that this value has three significant figures. If we wrote 1 M, then this would indicate that this value has only one significant figure and is less precise.

When taking measurements in a lab and processing data, it's important to **round your result to the same number of significant figures as the measurement with the fewest significant figures.** This will ensure that your result **reflects the reliability of the least accurate measurement.**

Analysing data

Data analysis refers to the **interpretation and examination** of data to extract meaningful insights and draw conclusions.

A **trend** in data analysis refers to a consistent direction that the data is moving (e.g. an increase or decrease in data values).

Patterns describe repeating or recurring features in data. Patterns highlight predictable behaviours or observations that occur repeatedly in a dataset.

Causation refers to a direct cause and effect relationship, where one variable directly causes a change in another variable.

Correlation refers to a relationship between two or more variables where one does not necessarily cause the other.

Correlation does not necessarily imply causation. Even if there is a correlation, it may be simply a coincidence.

Presenting data (Graphs)

Graphs are a very useful way for scientists to present data. They allow for the **quick interpretation of large amounts of data** (datasets) and are **very effective for identifying trends and patterns**. Graphs also provide a clear visual representation of the data collected. When multiple trials have been conducted, the **average** value of each measurement should be used to plot the graph.

There are many different types of graphs, and you must be able to decide which type is most suitable for your data. This choice depends on the **type of data** being graphed:

- If the independent and dependent variable are continuous data, then a **line graph** is the most suitable choice.
- If the independent variable is qualitative data, e.g. the name of a chemical, then a **column graph** is a more appropriate choice.

When drawing graphs, plot the independent variable on the x-axis and the dependent variable on the y-axis.

Drawing graphs

- Always use a pencil and graph paper.
- Plot the independent variable on the x-axis (horizontal)
- Plot the dependent variable on the y-axis (vertical)
- Decide if you are going to draw your graph in a portrait or landscape orientation. (Landscape is often a good option to ensure your graph fits comfortably on the page).
- For a line graph, use consistent marks and numbers to indicate the scaling along each axis. Before plotting a graph, count carefully and decide on the most sensible scale for your graph. Ideally scales should go up in 5s, 10s, etc., rather than awkward numbers like 3s.
- Label each axis, including the name of the variable and its unit of measurement, if applicable. These labels should match the labels in your data table.
- Plot data points as small, neat dots (or little x's if you wish).
- Connect all data points into a line or curve (used when data points are directly related) or draw a line of best fit (to summarise an overall trend).
- Include a title.
- If there are multiple lines on the same graph (e.g. comparing the volume produced using chunks of calcium carbonate vs. crushed), then there should also be a key (legend) to indicate what each line represents.

Reliability & Validity

Reliability refers to consistency and repeatability in measurement over multiple tests.

Validity refers to the suitability of an experiment or other investigation to answer accurately the question being asked, i.e. how well it measures what it is intended to measure.

Accuracy vs. Precision

Accuracy refers to how close the experimental measurements are to the true value.

Precision refers to how close repeated measurements are to each other.

Integrity & Ethics

Integrity refers to the honesty, accuracy, transparency and accountability of the scientific process.

Integrity means that scientists are honest in the way that they report results: not omitting or changing data – even if it is does not suit their hypothesis.

Integrity ensures that:

- Data is not fabricated (made up), falsified (changed), misrepresented or reported selectively.
- Results are reported transparently, with any errors and limitations included.
- Proper credit is given by acknowledging and properly citing sources.
- Experiments are as objective as possible, avoiding bias.

Ethics refers to the moral principles that guide the design, execution and reporting of experiments.

Ethical experiments are experiments that:
- are conducted safely,
- treat human or animal participants with respect and dignity and ensure that humans have given informed consent,
- follow GDPR (General Data Protection Regulation) guidelines,
- minimise environmental impact and harm to ecosystems,
- have no conflict of interest or bias.

Errors

Systematic errors cause measurements to differ consistently from the true value by the same amount each time a measurement is taken.

Random errors cause each measurement to deviate from the true value by a random amount and in a random direction.

You can reduce the impact of random errors on the overall results by conducting an experiment multiple times and averaging out the results. The average of repeat readings should be more reliable than a single reading.

Drawing conclusions

Your conclusion should summarise what you have learnt from the investigation. It should clearly explain the relationship between the variables, link the findings back to the research question, hypothesis and prediction, and explain the findings using your scientific knowledge.

Evaluating

Reflect on your
- Has your method produced results that allowed you to reach a clear conclusion?
- Are you certain that only the independent variable affected the dependent variable?
- Could other apparatus have been more suitable?
- If you repeated the experiment multiple times, did you get consistent results?
- Was your data accurate and/or precise?
- Did your data have a large range?

You could also suggest and explain possible changes to your experiment.