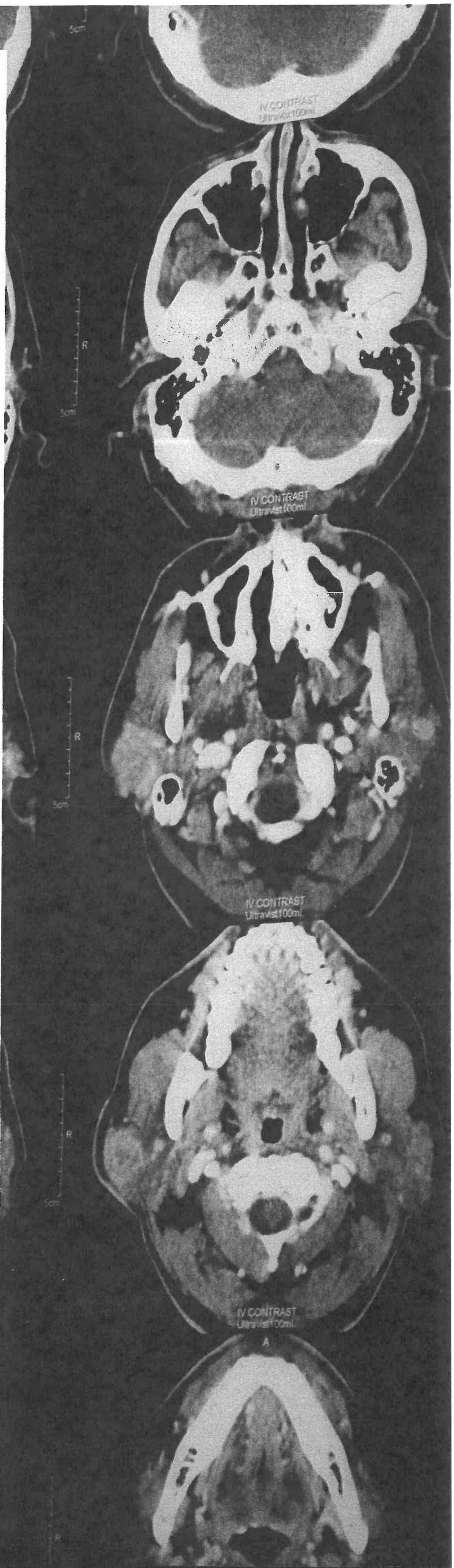
CHAPTER 01

SCIENCE INQUIRY

UNIT 3 AND 4 CONTENT

SCIENCE INQUIRY SKILLS

- identify, research and construct questions for investigation; propose hypotheses; and predict possible outcomes
- design investigations, including the procedure(s) to be followed, the materials required, and the type and amount of primary and/or secondary data to be collected; conduct risk assessments; and consider research ethics, including animal ethics
- conduct investigations, including the collection of data related to homeostasis and the use of models of disease transmission, safely, competently and methodically for the collection of valid and reliable data
- conduct investigations, including the use of virtual or real biotechnological techniques of polymerase chain reaction (PCR), gel electrophoresis for deoxyribose nucleic acid (DNA) sequencing, and techniques for relative and absolute dating, safely, competently and methodically for valid and reliable collection of data
- represent data in meaningful and useful ways, including the use of mean, median, range and probability; organise and analyse data to identify trends, patterns and relationships; discuss the ways in which measurement error, instrumental accuracy, the nature of the procedure and the sample size may influence uncertainty and limitations in data; and select, synthesise and use evidence to make and justify conclusions
- interpret a range of scientific and media texts, and evaluate models, processes, claims and conclusions by considering the quality of available evidence, including interpreting confidence intervals in secondary data; and use reasoning to construct scientific arguments
- select, use and/or construct appropriate representations, including diagrams, models and flow charts, to communicate conceptual understanding, solve problems and make predictions
- communicate to specific audiences, and for specific purposes, using appropriate language, nomenclature, genres and modes, including scientific reports



In science the important thing is to modify and change one's ideas as science advances.

Herbert Spencer

Herbert Spencer, an English philosopher who lived from 1820 to 1903, neatly summarised what science is all about. Science is a process of inquiry aimed at finding answers to problems and discovering new knowledge about the natural world. The knowledge discovered as a result of scientific inquiry becomes a part of science. That is, science means two things: a process of discovery and the knowledge that is discovered. The information presented in this book is science. It is some of our present knowledge about the human species, knowledge that has been obtained by scientific investigation.

Types of investigation

Observations

Observation is an essential part of science. Any investigation, regardless of the procedure used, will involve some form of observation. In investigations based on observation, scientists are looking for patterns. When a pattern becomes evident it may be possible to draw tentative conclusions.

An example of an investigation based on observation is the discovery that peptic ulcers are caused by a bacterial infection. In 1979 Dr Robin Warren, a pathologist at the Royal Perth Hospital, observed the presence of bacteria in samples of stomach tissue that he was examining. Continued observation over the next few years showed that the bacteria were often present in the stomachs of patients suffering from stomach inflammation. Warren's discovery was not taken seriously because, at the time, it was believed that the stomach contents were too acidic for bacteria to survive.

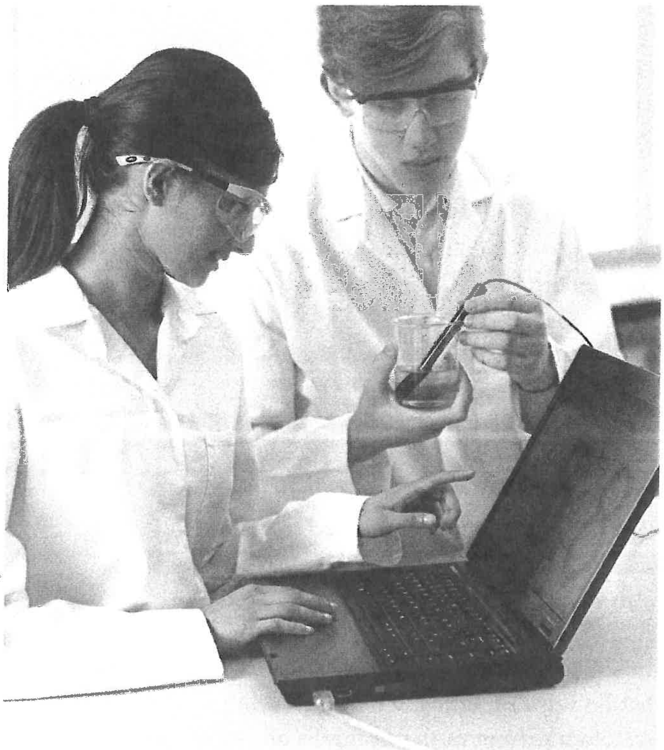
In collaboration with Dr Barry Marshall, a doctor specialising in stomach disease, Warren was able to show that a particular species of bacterium was present in the majority of cases of stomach and duodenal ulcers. They also found that it was rare to have an ulcer without being infected by the bacterium that Warren had discovered. Marshall and Warren went on to culture the bacterium and to show that it did indeed cause ulcers. As a result of their discovery it became easy to treat stomach and duodenal ulcers with antibiotics. The discovery that the bacterium *Helicobacter pylori* was the cause of ulcers began with simple observation and became one of the most significant events in Australian medical history. Robin Warren and Barry Marshall were awarded the Nobel Prize for Physiology or Medicine in 2005.

Another example of systematic observation is when it is used to gain knowledge of animal behaviour. Jane Goodall, the first person to observe the social organisation of chimpanzees in the wild, documented the interactions of chimps with one another, their social hierarchy, their tool making and many other features of their society. She began her observations in 1960, and her work is being continued today through the Jane Goodall Institute.

Controlled experiments

Controlled experiments, sometimes called fair tests, are designed to investigate relationships between factors (or variables). They involve changing one variable while all the other variables are kept the same. Any differences in the results should be due to the changed variable.

Howard Florey, an Australian working at Oxford in England, used controlled experiments to demonstrate the effectiveness of penicillin in treating bacterial infections. In 1940 a crucial experiment to test the effectiveness of penicillin as an antibiotic was carried out. Eight mice, all the same weight and age, were each injected with 100 million streptococci, a type of bacterium. Previous experiments had shown that an injection of that size would kill all mice injected. After the injection of streptococci, four of the mice were put back in their cages and given no further treatment. The other four mice were given injections of penicillin. The mice in the control group – those that did not receive penicillin – all died within 12 hours. Mice in the experimental group, which were given penicillin, survived for many days – one for more than six weeks.



Science Photo Library

Figure 1.1
All scientific investigation involves the collection of data. These students are collecting data on the pH of a liquid.

The only difference between the mice in the two groups was the injection of penicillin, so the survival of those mice was good evidence that penicillin was effective in combating bacterial infections. Further positive feedback from repeated controlled experiments gave Florey and his co-workers the confidence to try penicillin on humans suffering from bacterial infections. The results were outstanding. Florey and his colleague Ernst Chain were awarded the Nobel Prize in Physiology and Medicine in 1945.

Surveys

A **survey** is a process of systematically collecting, analysing and interpreting information about an aspect of a study. Surveys are usually designed to collect data from a large number of subjects. The information may be collected using a questionnaire or by interview. Using the large amount of information collected, the researcher can then look for patterns in the data.

Dr Karl Kruszelnicki was stimulated by a listener to his radio program to carry out a survey into the origins of belly button lint. The survey was conducted over the Internet, with 4799 people responding, and publicised on the radio station Triple J. The patterns in the responses revealed that people more likely to have belly button lint are male, hairy, with a concave belly button, and that the amount of lint increases with age. This was a light-hearted exercise but it does demonstrate the principles involved in conducting a survey. In 2002 'Dr Karl' received an Ig Nobel Prize for his efforts. Ig Nobel Prizes are awarded for research that makes people laugh and then makes them think.

Trial and error

Dr Karl's survey

Trial and error sounds like a random process but when used in scientific research it is systematic. The process involves one attempt to solve a problem being followed by another. Each trial is recorded and the results allow the investigator to gradually home in on the solution to a problem. Thomas Edison, who developed the electric light globe, had to find a suitable material for the filament in the light globe. Using trial and error, he examined more than 600 different materials before finding one that was satisfactory.

Many new drugs, such as antibiotics, have been discovered using trial and error. Chemical compounds extracted from plants can be tested on cultures of bacteria to see whether they have any effect. Those that show promise can then be subjected to further testing under different conditions. Meticulous records of the results of each trial must be kept. Such research is often prolonged and tedious but it is often the only way to find effective substances.

Case studies

A **case study** is an in-depth investigation of one particular person or situation. Case studies are frequently used in areas such as education and business management. However, they may also be useful in some areas of science. For example, in medicine the progress of a particular disease in one person may be documented. Such a case study can extend or help to confirm what is already known about the disease.

Longitudinal studies

A **longitudinal study** is conducted over a long period of time. It is similar to a case study but is more prolonged. Longitudinal studies may take place over many years, even decades. They can also be done retrospectively, which means that the researcher can examine records of past events to build up a picture of change over time.

The Busselton Health Study is a longitudinal study of the population of Busselton, a coastal town in the south-west of Western Australia. Begun in 1966 and continuing today, it is the world's longest-running study of the health of a population.

Busselton Health Study

A longitudinal study on a smaller scale was carried out by an American doctor, William Bean, who studied the growth of his fingernails for 35 years. He did this by filing a horizontal line on his thumbnail just above the cuticle and recording how long it took the mark to reach the tip of his thumbnail. From his records he was able to calculate the growth rate. In 1980, after 35 years of measurements, Bean was able to conclude that the growth of his nails had slowed from 0.123 mm a day when he was 32 years of age to 0.095 mm a day at the age of 67.

EXTENSION

1 A research method sometimes used by scientists is meta-analysis. Find out what is meant by meta-analysis and give an example of an investigation that used this method of research.
2 Some controlled experiments are said to be 'double-blind' experiments. Find out what is meant by a double-blind experiment and give examples of how such experiments might be used.

Scientific method

Scientists investigate in a methodical and systematic way. Wherever possible the data they collect is quantitative – that is, in the form of numbers. Detailed records are kept of the procedure followed and the results obtained.

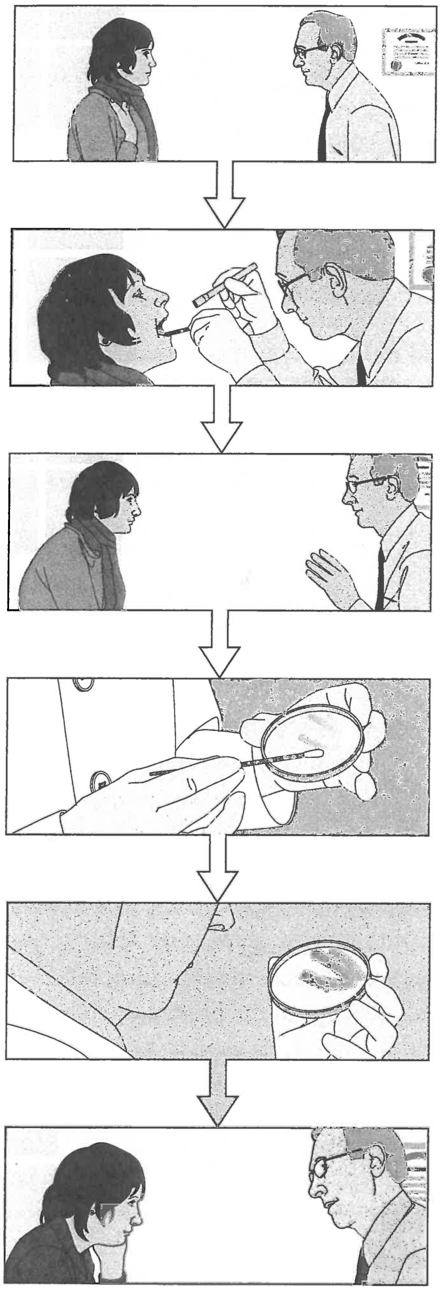
As we have seen, there are many ways of investigating in science. The method used will be the one that best suits the situation. However, many investigations lead to the testing of a hypothesis and will follow a pattern similar to the following.

1 Recognise a problem and define a question.
2 Collect as much information as possible relating to the problem.
3 Propose a hypothesis – a possible explanation for the problem.
4 Test the hypothesis using an experiment.
5 Analyse and interpret the data collected from the experiment.
6 Draw conclusions about whether the hypothesis was supported or disproved.
7 Report on the investigation.

Note that although a hypothesis may be disproved, it cannot be proved. The results of an experiment can only provide *support* for the hypothesis. As Albert Einstein said:

> No amount of experimentation can ever prove me right; a single experiment can prove me wrong.

The scientific method outlined above can be applied to many problem-solving situations. Figure 1.2 shows how a doctor uses the scientific method to arrive at a diagnosis. A mechanic would probably go about solving the problem of a car that would not start in the same way.



Problem
A patient with a sore throat visits a doctor. The doctor recognises the problem and formulates a question:
'What is causing the patient's sore throat?'

Collect information
The doctor asks the patient how long her throat has been sore; takes the patient's temperature; feels the throat for swelling; looks carefully inside the patient's mouth.

Propose a hypothesis
The doctor proposes a hypothesis:
'The sore throat is caused by a bacterial infection.'

Test the hypothesis
The doctor takes a swab from the patient's throat and sends it for testing at a laboratory.

Analyse and interpret data
The laboratory staff examine the culture from the patient's swab and check for disease-causing bacteria. They send their findings to the doctor.

Draw conclusions
As a result of the information from the pathology laboratory, the doctor decides that the patient's sore throat is caused by an infection of *Streptococcus*. The doctor's hypothesis is supported by the data collected from the test.

Scientific method
This website has one section that describes the scientific method and another that shows how it is applied in astronomy.

How science works
This website provides a series of modules designed to answer the question: 'What is science and how does it work?'

Figure 1.2 The scientific method

Some scientists are not happy with descriptions of the 'scientific method' because of the many different ways that can be used to gather data. One of the main criticisms of using this 'method' as a model of scientific investigation is that it does not give enough weight to the importance of observation as a means of obtaining knowledge. In some investigations, observation alone can lead to new understanding without the need for any experimentation.

Despite the criticisms, the scientific method is a useful model because it can be applied to many investigations. When you do investigations yourself, you will find the model a useful planning tool.

Planning an investigation

Literature review

One of the early steps in the scientific method described above is to 'collect as much information relating to the problem as possible'. One way of collecting the information is to carry out a literature review. A **literature review** is a survey of the material that has been written about the subject under consideration. Literature reviews used to involve long hours of library research but since the Internet was established they have become much easier.

The purposes of a literature review are as follows.

- To help the researcher define the problem: defining the problem carefully helps in the design of an investigation that will contribute to a solution of the problem.
- To find out what is already known about the problem: this prevents duplication of effort and allows the researcher to build on knowledge that is already available.
- To assess research methods: methods used by others may be used or adapted for the researcher's own investigation.
- To allow researchers to relate their findings to what is already known: this is particularly useful for the final report on an investigation. It is also helpful in considering areas for further research.

Safety

It is important that an investigation presents no danger to the participants or to the investigators. Examine the design of your investigation to make sure that there are no safety risks associated with any of the steps in the procedure. Consider the need for safety devices such as fire extinguishers, fire blankets, fume hoods and eye washes. Also assess the need for personal protection such as safety glasses, gloves, face masks and protective clothing. For some investigations it may be necessary to assess the participants. For example, do they have any allergies to the substances being used? Do they have health problems that could be affected by the activities involved?

Safety considerations should include not just the physical safety of the people involved, but also factors such as whether the participants will feel insecure, threatened or embarrassed.

Ethics

Ethics are a set of moral principles or values. They are standards that are observed by most people in our society. **Ethical behaviour** is behaviour that follows those principles or values. In scientific research, especially research involving human participants, there are many ethical issues.

Some of the principles that an investigation involving humans must satisfy if it is to be ethically sound are:

- *voluntary participation* – the subjects should not be pressured into taking part in the investigation
- *informed consent* – the subjects should be fully informed about the objectives of the research, the procedures to be followed, any possible risks and the potential benefits of the research; consent should only be sought after all information has been given
- *risk of harm* – as mentioned in the section on safety, there should be no risk of physical or psychological harm
- *confidentiality* – the identities of participants will not be revealed except to people directly involved in the study.

Just as humans must be treated in an ethical way, the same applies to animals. The requirements for investigations involving animals are set out in the *Australian code for the care and use of animals for scientific purposes 8th edition (2013)*. The code sets out detailed requirements, but in general terms any use of animals in research or teaching should be:

- valid
- humane
- justifiable
- considerate.

Controlling variables

A **variable** is any factor that may change during an experiment.

The **independent variable** is the factor that is being investigated – it is deliberately changed to determine its effect. This variable is deliberately made different in the control and the experimental groups in an experiment. By comparing the results from the control and experimental groups, the effect of the independent variable can be determined.

The independent variable may also be called the experimental variable or the manipulated variable.

The **dependent variable** is the factor that changes in response to the changes made to the independent variable. It is sometimes called the responding variable.

Controlled variables are the factors that are kept the same for both the control and the experimental groups in an experiment.

In any experiment it is important that, with the exception of the independent variable, all variables are kept the same for the control and the experimental group of subjects. If one or more is not kept the same, it is impossible to tell which variable is causing any difference between the two groups of subjects.

Sometimes it is difficult or impossible to keep all variables the same. **Uncontrolled variables** are variables that are not kept the same for the control and the experimental groups in an experiment. They may have been overlooked by the experimenter or they may have been impossible to control. If there are uncontrolled variables in an investigation, this must be taken into account when interpreting the results.

Repetition and replication

Scientific experiments always involve repetition or replication. **Repetition** means doing the same experiment many times. **Replication** means having a number of identical experiments running together or performing the experiment on a large number of subjects at the same time. Both repetition and replication help to demonstrate that results are consistent. If results are different each time an experiment is performed, they are of little value.

Repetition and replication can also help to overcome the effects of uncontrolled variables. For example, if 10 subjects are used in an experiment and one of them is unusual in some way, it will have a big effect on the overall result. If 100 subjects were used, one unusual person in 100 would not have much effect on the average result.

In designing any experiment, plan for as much repetition or replication as time and resources will allow.

Validity and reliability of results

An experiment is **valid** when it tests what it is supposed to test. Some scientists were testing the hypothesis that 'consumption of junk food affects people's memory'. They fed one group of young rats on fatty food for 12 weeks, and fed another group of older rats a low-fat diet. The rats' memories were then tested using an activity that involved pressing a lever. The rats fed on junk food were more forgetful, so it was concluded that the hypothesis was supported. This experiment did not test what it was supposed to test, for two reasons.

1 Testing one species, rats, will only demonstrate the effect on the memory of rats, not any other species.
2 Rats' memories may be affected by age. The two groups of rats should have been of the same age.

Experiments can also be invalid if there are uncontrolled variables – that is, if there are factors that could affect the result of an experiment that are not kept the same for the experimental and the control set-ups. When experimenting with humans, it is often very difficult to design a valid experiment because it is hard to control all the variables.

Reliability is the extent to which an experiment gives the same result each time it is performed. The measuring instruments used in the experiment must also be reliable; that is, they must give the same measurement each time they are used. For example, you may have a set of bathroom scales that give three different weights when you step on them three separate times. Those scales are unreliable, and if used in an experiment would make the results unreliable. The bathroom scales may give the same reading every time but it may be consistently higher or lower than the actual weight. In that case the scales are inaccurate.

Repetition and replication are used to make sure that results are reliable, but they do not improve the accuracy of the experiment.

Analysing results

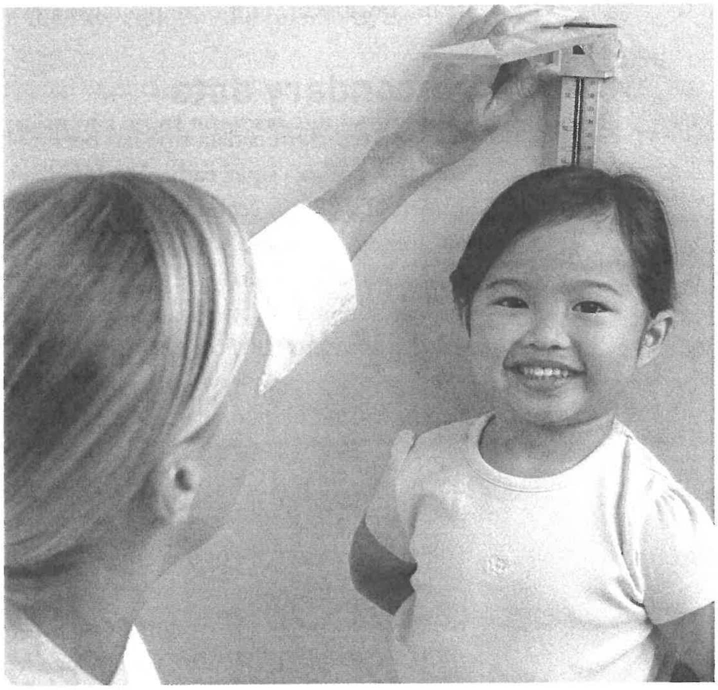
Quantifying results

Data from an investigation can be one of two types:

- **quantitative data** – expressed in numbers and usually involving measurement; for example, 'the students are 174 and 176 cm in height'
- **qualitative data** – observations that do not involve numbers or measurement; for example, 'student A is taller than student B'.

Wherever possible, you should design an investigation so that the results are quantifiable (Figure 1.3). Numerical results can be ranked, averaged and manipulated in other ways. They can also be summarised using graphs.

Sometimes it is possible to quantify qualitative data. For example, if asking people's opinions on something, they can be asked 'Do you disagree strongly, disagree, agree or agree strongly?' or asked to answer using a numerical value such as 1 for 'disagree strongly' to 4 for 'agree strongly'.



Science Photo Library/Adam Gault

Figure 1.3 Where possible, design an experiment so that the results are expressed as measurements. Measuring height in millimetres is much more meaningful than observations such as 'tall' or 'short'.

Errors and limitations in data

It is important that data are checked carefully for errors. In science an error is not necessarily a mistake. It is any deviation from the result that should have been obtained. One of the reasons why scientists provide comprehensive reports on their investigations is so that others can check their data for errors.

Measurements made with any measuring instrument are approximate. For example, if you measure a person's height at several different times, the measurements are unlikely to be the same every time. This may be because there is natural variation in the subject, variation in the measurement process, or both. This uncertainty in measurement is called **measurement error**. In this case the word 'error' does not mean the same as 'mistake'. Your measurements are not wrong; the measurement error is the difference between the measurements you made and the true value of what you were measuring. Repetition can help to reduce measurement error but it cannot overcome error caused by the limitations or deficiencies of the measuring instrument.

It is also important to understand the limitations of data obtained from an investigation. You must not draw conclusions that go beyond the data. Sometimes it is difficult to look objectively at data, and even experienced scientists can draw conclusions that are not necessarily supported by their data.

One example of reading too much into the data obtained in an investigation arose from a report by Norwegian scientists on the incidence of breast cancer in 25 624 women. Published in the *New England Journal of Medicine* (1997, volume 336, number 18, page 1269), a prestigious medical journal, the results of the scientists' survey showed that the incidence of breast cancer in women who exercise regularly was reduced by 37%. The media reported on the investigation with headlines stating that 'exercise prevents breast cancer'.

Other scientists pointed out that women who exercise regularly are also likely to be non-smokers, drink less alcohol, have healthier diets, and have higher levels of education and higher incomes than women who do not exercise. Which of these variables was actually contributing to the reduction in breast cancer? Was it really exercise, or could it be having a healthier diet, being a non-smoker, having a better education and so on? Could it be a combination of some of these factors? Each factor and combination of factors would have to be investigated before arriving at a firm conclusion. This example illustrates some of the pitfalls in analysing data.

Secondary data

Secondary data is data that has been collected by someone other than the people who are using the data. For example, earlier in this chapter we quoted the rate of growth of Dr William Bean's fingernails. This is secondary data – fingernail growth was measured by Dr Bean, not by the authors of this book.

Secondary data may sometimes include a *confidence interval*. A confidence interval is used to indicate the reliability of data. It is the range of values above and below a result in which the actual value is likely to fall. For example, opinion polls published in the media may say that 53% plus or minus 1.5% of people will vote for Party X. The confidence interval is 50.5 to 54.5%. A confidence *interval* should be quoted along with a confidence *level*. The *confidence level* most commonly used in research is 95%. This means that if the research were repeated a number of times, the range of values obtained would contain the true value 95% of the time. In the survey of voters, if the survey were repeated many times, the proportion of people who intend to vote for Party X would be between 50.5 and 54.5%, 95% of the time.

Another example may help clarify this concept. Suppose you wanted to find the average height of Year 12 students in Western Australia. You could measure the height of every Year 12 student in the state and then calculate the average height. This would give you an accurate result (the true value) but it would be impractical. A more practical method would be to measure a sample of Year 12 students. If you took a sample of 20 Year 12 students and calculated their average height it would give you an estimated result, but it would not tell you how certain you could be that your result was correct. Using a mathematical formula, a confidence interval could be calculated that would indicate the reliability of your estimate. The calculated confidence interval may show, for example, that using the same sampling method, the average height of Year 12 students will be between 167 and 179 cm 95% of the time.

Processing data

If you have designed an experiment to give quantitative data, you will end up with a mass of figures that you must interpret. In a controlled experiment you will have to compare the control and the experimental results. There are some simple calculations that you can do to make the numbers more meaningful.

Average

In science a description of a set of numbers almost always includes a measure of its centre, or its **average**. Averages are a very common and simple way of handling sets of numerical data. The average that is most often calculated is the **arithmetic mean**, often just called the **mean**. To calculate the mean of a group of measurements, you add up all the measurements in the group and divide by the total number of measurements.

Sometimes in a set of measurements there will be values that are well beyond the range of the rest of the measurements. These are called **outliers**. The mean is affected by outliers, because a very high or very low outlier value will make the mean higher or lower than it would be without the outlier included. Outliers may result from mistakes in measurement, the failure of equipment or other errors. If the outliers clearly result from an error, they may be excluded when the mean is calculated.

Median

The **median** is the middle of a set of numbers. It divides the lower set of numbers from the upper set. For example, the heights of the members of a cricket team were measured and (in centimetres) they were: 164, 176, 177, 177, 178, 181, 182, 182, 183, 185, 191.

The median height of the team was 181 cm. 181 is the middle value – there are five team members with heights lower than 181 cm and five with heights higher. If there is an even number of measurements then the median is taken as the mean of the two values in the middle of the set of numbers.

Using the median of a set of numbers reduces the influence of outliers. Outliers due to measurement error could have a significant effect on the mean of a set of numbers, but would have much less effect on the median.

Range

A measure of the centre of a group of numbers can be misleading. The mean, or the median, gives us no idea about whether all the values are clustered around the centre or whether there is a very wide spread from highest to lowest value. Any description of a set of numbers should therefore include both a measure of centre and a measure of spread.

The simplest way to indicate the spread is to quote the **range** – that is, the highest and lowest measurements in the group. For example, we could say that the heights of students in a Year 12 class ranged from 151 to 183 cm with a mean of 171 cm.

Scientists use a number of other measures of spread, such as quartiles or standard deviation, but range should be sufficient for your investigations.

Ratios and rates

A **ratio** is a numerical statement of how one variable relates to another. That is, it is a comparison of two numbers. Ratios are written as two numbers separated by a colon. For example, on older TV screens the ratio of width to height was 4:3. If the width is 40 cm, the height is 30 cm; if the width is 60 cm, the height is 45 cm and so on. Modern, widescreen TVs have a ratio of 16:9.

A **rate** is a special kind of ratio that shows how long it takes to do something. For example, a good athlete can run 10 000 m (10 km) in around 30 minutes. This is a rate of 1 km per 3 minutes or 20 km per hour. Rate is much more meaningful than a simple count of how often something occurs. If you were investigating the effect of exercise on breathing, counting a person's breaths would be meaningless unless you knew how many breaths there were in a given time. That is, you need to know the rate in breaths per minute.

Percentages

Per cent means 'per hundred'. Percentages are used to express how large one variable is in relation to another. For example, if a breakfast cereal is labelled as containing 1.5% fat, it means that 100 g of the cereal contains 1.5 g of fat.

In Western Australia in 2011, males aged 15 to 19 years made up 6.7% of the population; females of the same age made up 6.4%. This means that for every 100 people in the population, 6.7 (or 67 per thousand) are 15 to 19 year-old males and 6.4 (64 per thousand) are 15 to 19 year-old females. Or we could say that for every 64 girls aged 15 to 19 in Western Australia, there are 67 boys aged 15 to 19.

Percentage change

Calculating a percentage increase or decrease is often a good way to understand changes in a variable over time. For example, if a person weighing 100 kg lost 10 kg after dieting for 6 months, we could say that the person had lost 10% of their body weight as a result of the diet. If another person weighing 120 kg lost 13 kg after 6 months on the same diet, the percentage decrease would be 10.8%. Percentage change is helpful in making such comparisons.

To calculate percentage change:

1 *subtract* the old value (120 kg) from the new value (107 kg)
2 *divide* by the old value (120 kg)
3 *multiply* the result by 100 and add a per cent sign (%) to it.

This can be written as a formula:

$$\frac{\text{new value} - \text{old value}}{\text{old value}} \times 100 = \text{percentage change}$$

If the percentage change is positive, it indicates an increase; if the change is negative, it indicates a decrease.

Frequency

Frequency is the number of times an event occurs. For example, some students conducted a survey to find out how many drinks containing caffeine were consumed by the members of their class in a two-day period. The table of data they collected is called a **frequency distribution** or **frequency table**. A frequency table summarises the data by showing how often the variable in question occurs (Table 1.1). Frequencies can also be presented graphically as a histogram (Figure 1.4).

Table 1.1 Frequency table showing number of caffeine drinks consumed over a two-day period

Number of drinks consumed	Number of students
0	3
1	0
2	3
3	7
4	6
5	3
6	1
7	2

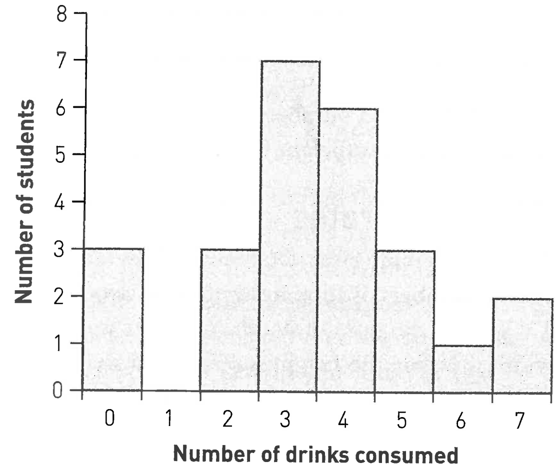


Figure 1.4 Histogram showing number of caffeine drinks consumed by students in a Year 12 class over a two-day period

Probability

Probability is the chance, or likelihood, that a particular event will occur. Many outcomes cannot be predicted with certainty. In such cases the best we can do is to say how likely it is that the event will occur. For example, the weather bureau cannot know with certainty when it will rain. If they forecast that there is a 20% chance of rain, they are stating a probability.

Often the probability of an event occurring can be expressed in more precise mathematical terms. Probability is the number of ways an event can occur divided by the number of possible outcomes. For example, what is the probability that when you throw a die you will get a three? There is only one way you can get a three but there are 6 possible outcomes. The probability of throwing a three is therefore $\frac{1}{6}$.

What is the probability of throwing an even number? There are three ways you can get an even number from 6 possible outcomes, so the probability of throwing an even number is $\frac{3}{6}$ or $\frac{1}{2}$. The probability can also be expressed as a percentage (50%) or as a decimal (0.5).

Probability is useful in many areas of science investigation. For example, in genetics (covered in Unit 2 of Human Biology) it is possible, in a simple cross, to state the probability that the offspring will inherit a certain characteristic.

Presentation of data

Tables

A convenient way to present numerical data is in the form of a table. The table of results for the students' survey of the number of caffeine drinks consumed in a Year 12 class over a two-day period could look something like Table 1.1.

Notice that the table follows these rules:

- It has a title. The title should state the variables represented by the data; in this case, the number of drinks consumed and the number of students are the two variables.
- The data are presented in columns. Usually the independent variable (in this case the number of drinks consumed) is in the left column and the dependent variable (number of students) is in the right column. This rule is not always applied – it is more important that the table be easily understood.
- Each column has a heading and, where appropriate, the heading must state the units in which the data have been measured.

Other examples of tables may be seen in Activity 9.1 on page 125.

Graphs

Graphs are a pictorial way of presenting numerical data. A graph shows how changes in one variable affect another variable. From a graph it is easy to see any trends in the data. It is also possible to predict what the values would have been between the points plotted (interpolation), or the trend beyond the data shown in the graph (extrapolation).

Figure 8.2 on page 103 is a typical *line graph*. When drawing any graph the following rules must be observed.

- The graph should have a title that states the two variables shown on the graph. For example, the title for Figure 8.2 describes what the graph is about.
- The independent variable ('Time' in Figure 8.2) is plotted on the horizontal axis and the dependent variable ('Blood glucose concentration') is plotted on the vertical axis.
- Each axis is labelled with one of the variables and the units in which it is measured.
- Equal intervals of units are used on each axis.

The most commonly used graphs are line graphs (e.g. Figure 3.9 on page 42), histograms, column graphs (e.g. Figure 17.18 on page 263) and bar graphs. Descriptions of how these graphs are used may be found in *Human Perspectives Units* 1 & 2 *ATAR*.

Tables and graphs
At the Australian Bureau of Statistics website you can find more information on graphs, and download the document 'What Graph or Display to Use When'.

Models

In science a **model** is a simplified representation of an idea or process. Figure 1.2 is a model of the scientific method. Once a model has been developed it can be applied to a number of situations. The model for scientific method can be applied to most scientific investigations. It can also be applied to other situations, as in Figure 1.2 where it is being applied to the steps a doctor may take in trying to diagnose an illness.

Figure 2.2 on page 23 is a model showing in simple diagram form how hormones may affect the functioning of a cell. The steady state control system described on page 76 is a model that can be applied to the regulation of body temperature, blood glucose and many other situations.

A model may be a diagram, a flow chart or a physical model such as a model of the atoms in a protein. Scientific models often have to be modified as new data are collected.

Flow charts

A flow chart is a diagram that shows the steps involved in a process. The steps are usually shown in boxes and the sequence of steps is indicated by arrows. Flow charts are very useful in summarising and visualising the steps in a complex process. Numerous flow charts are used in this book – for example, the flow chart illustrating the stimulus-response-feedback model on page 77 and the flow chart on page 90 showing how body temperature is regulated.

Reference to the work of others

An extensive examination of the literature at the start of an investigation allows the researcher to fully grasp the available information relating to the problem under consideration. This review also allows the results to be seen in the context of what is already known. Research done by others can also be used to support or confirm what has been discovered in the investigation. Demonstrating how your findings relate to what is already known will give credibility to your research and will add to the body of knowledge on the subject under review.

Reporting

When an investigation has been completed, the findings need to be made known to others. This is usually done by a written report. Reports are a very important part of communication in science. Scientists inform others of their research by publishing a report in a scientific journal. There are thousands of scientific journals, some of which deal with a very narrow field of science. Examples are *Nature*, *Science*, *Journal of Musculoskeletal Research* and *Journal of Genetics*.

The editors of scientific journals use a process called **peer review** to make sure that the report is worthy of publication. A submitted report is sent to one or more scientists who are experts in the field and who may or may not recommend publication. This process is important as it helps to keep scientific literature free of incorrect, bogus or misleading information.

A scientific **report** includes a description of an investigation, the results that were obtained and any conclusions that can be drawn from the results. The description of how the investigation was done must be sufficiently detailed to allow other scientists to repeat the experiment. It is common practice for scientists to repeat experiments that others have performed. If the results obtained are not the same as those for the original experiment, any conclusions that may have been drawn are questionable.

Reports follow a fairly standard format, similar to that described below.

Scientific report format

Reports may be written using the following headings.

- **Title** of the report and **name** of the author or authors
- **Introduction**, stating the nature of the problem investigated and the hypothesis tested
- **Materials and equipment**, listing the apparatus used, particularly any specialised items of equipment
- **Procedure**, describing the exact method used to carry out the investigation
- **Results**, often presented as tables, graphs, diagrams or photographs
- **Discussion**, including comments about the results and the way they relate to the hypothesis that was tested
- **Conclusion**, summarising the most important parts of the discussion and stating the success or otherwise of the investigation
- **Further research**, as scientific investigations often raise more questions than they answer – many reports suggest areas that need further investigation

- **References**, which list any reports, books, journal articles, websites or other sources of information referred to in the report
- **Acknowledgements**, of people who helped with the investigation or of organisations that provided funds for the research.

The discussion

The most important and longest part of a report is usually the discussion. The discussion is about the results and the method used to obtain the results. The discussion needs to be very thorough and to address all aspects of the research.

A checklist of questions that could be answered in the discussion section is as follows.

- Were there any defects in the design of the investigation or in the procedure?
- Were any results different from those expected?
- How do the results fit into the broader context of what is already known about the topic?
- Are there any practical applications for the results?
- Do the findings relate to any earlier work in the same area?
- Did the results support the hypothesis, or did they indicate that the hypothesis was incorrect?
- Were there any limitations in the research?
- Could the investigation have been improved in any way?
- Were there any variables that could not be controlled?
- Was there any bias in the results?
- Is there any information available from other reliable sources that would support the results?
- Is there a need for further research to clarify any of the results?

This is not an exhaustive list of questions; when writing a report, you will be able to think of other points that need to be discussed.

Report writing
This Monash University website gives detailed advice on report writing.

Report writing FAQ
This University of New South Wales website gives advice on report writing and links to other useful sites.

A case study of a scientific investigation

French scientist Louis Pasteur (1822–1895) conducted hundreds of investigations. His achievements include showing that micro-organisms cause disease, developing vaccines for rabies and some animal diseases, showing that micro-organisms are responsible for fermentation, and showing how the development of micro-organisms could be prevented by boiling and then cooling a liquid. This last process became known as pasteurisation.

Pasteur's investigations followed the scientific method. We can use aspects of his work as examples of many of the points discussed here. We will focus on Pasteur's demonstration that spontaneous generation does not occur. Spontaneous generation is the idea that living organisms can develop spontaneously from non-living matter.

Italian physician Francesco Redi had shown in 1668 that maggots develop from eggs laid by flies. Until that time it was believed that maggots formed naturally from rotting meat. Another Italian, Lazzaro Spallanzani, demonstrated 100 years later that micro-organisms come from the air and that boiling can kill them.

Despite the work of Redi and Spallanzani, the belief persisted that micro-organisms could spontaneously develop in decaying organic matter. The French Academy of Sciences arranged a contest for scientists to disprove the idea of spontaneous generation. Pasteur took up the challenge in 1859. This is a good example of how scientific

Getty Images/Science Photo Library

Figure 1.5 Louis Pasteur

knowledge builds over time. First Redi, then Spallanzani and later Pasteur and many others were involved in debunking the idea of spontaneous generation.

Pasteur's reports
This website has one of Pasteur's research reports on the growth of micro-organisms, published in 1860.

Pasteur had a problem to be solved and his hypothesis was 'that micro-organisms occur in sterile culture medium only when exposed to contaminated air from the outside'. To test the hypothesis he began a series of meticulous experiments. He opened flasks of sterile broth in the streets of Paris and found that after a time there was abundant growth of micro-organisms in the broth. He opened flasks high in the Alps and the broth nearly always remained sterile. Variables other than the location of exposure were kept the same: the flasks were the same size and shape, with the same volume of the same type of broth. All flasks were opened for the same period of time and kept at the same temperature, and so on. From his results Pasteur was able to conclude that the flasks exposed in Paris became infected because of the large numbers of micro-organisms in the air. The flasks exposed in the Alps remained free of micro-organisms because there are fewer micro-organisms in the air at high altitude.

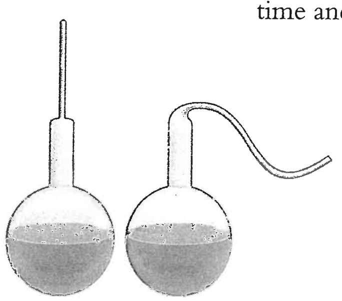


Figure 1.6 The types of flasks used by Pasteur in his experiment to demonstrate that spontaneous generation did not occur

Another experiment that Pasteur performed involved placing broth in flasks and heating them to kill micro-organisms. Some of the flasks then had their necks heated and drawn out into a long S-shaped curve (Figure 1.6). The necks of control flasks were heated but left straight. All flasks were left in the same location with their necks open to the air. After several weeks the broth in the flasks with straight necks had gone cloudy due to the activity of micro-organisms. Broth in the curved-neck flasks remained clear; the micro-organisms and dust in the air settled in the bend of the S-shaped tube and did not reach the broth in the flasks. In the straight-tubed flasks, micro-organisms and dust were able to reach the broth, where the micro-organisms multiplied and made the broth go cloudy. This experiment confirmed Pasteur's earlier conclusion that the air contains micro-organisms. Pasteur summarised his findings in a report titled, 'On the organised bodies that exist in the air. Examination of the doctrine of spontaneous generation'.

Getty Images/Science Photo Library

Figure 1.7 John Tyndall

Support for Pasteur's conclusions came from English physicist John Tyndall (Figure 1.7). He showed that sterile broth exposed to air but kept in a dust-free chamber remained sterile indefinitely. Tyndall and Pasteur were aware of each other's work – an example of one scientist producing evidence that supported the findings of another.

Despite the convincing evidence, the dispute over spontaneous generation continued. Many were not convinced and Pasteur often had to defend his research. At a lecture in 1864 he said:

> there is now no circumstance known in which it can be affirmed that microscopic beings came into the world without germs, without parents similar to themselves. Those who affirm it have been duped by illusions, by ill-conducted experiments, spoilt by errors that they either did not perceive or did not know how to avoid.

This situation persists in science today. The findings of scientists are subject to intense scrutiny by others and are often the subject of criticism – sometimes warranted, sometimes not. One reason for writing reports and presenting papers at conferences is so that other experts can examine the results. Many scientists have to vigorously defend their research in the face of criticism from peers and others.

Ideas about spontaneous generation were not finally laid to rest until 1876, when Pasteur and his assistant, Charles Chamberland, discovered that some bacteria produce spores that are resistant to high temperatures. These resistant spores accounted for the development of micro-organisms in cultures that had apparently been sterile for long periods. Some scientists had claimed that such development was the result of spontaneous generation. Spontaneous generation had finally been refuted 16 years after Pasteur's first convincing experiments and more than 200 years after Redi's research on spontaneous generation.

Pasteur made a great many discoveries about micro-organisms, but there was still a lot to find out. The scientific process of building knowledge continues in microbiology today.

Science inquiry

ACTIVITY 1.1 Validating Pasteur's experiment

You can repeat Pasteur's experiment in which he used flasks with S-shaped necks, to see whether you get the same results.

YOU WILL NEED

For each pair or group:

- beef cubes
- filter funnel and filter paper
- four 100 mL conical flasks
- four one-hole stoppers
- straight glass tubing and S-shaped glass tubing
- source of heat (hot plate or Bunsen burner)

WHAT TO DO

1. Make a broth using the beef cubes.
2. Filter the broth so that it is clear.
3. Place equal volumes of broth in each of the four flasks.
4. In two flasks place a stopper with straight glass tubing and in the other two flasks place a stopper with S-shaped tubing (Figure 1.8).
5. Gently boil the broth in each flask for 15 minutes.
6. Leave the flasks in a warm place and check every couple of days for evidence of the growth of micro-organisms, such as cloudiness, a scum or mould on the surface of the liquid, or bubbles. You may need to leave your flasks for several weeks before any changes are apparent.
7. At the conclusion of the investigation do not open any of the flasks. They must first be autoclaved at 120°C for 20 minutes under 100 kPa pressure to destroy any micro-organisms.

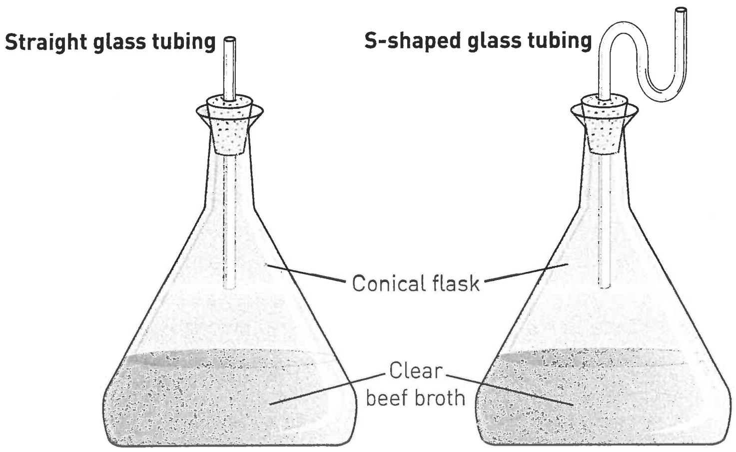


Figure 1.8 Flasks for validating Pasteur's experiment

STUDYING YOUR RESULTS

1. Describe your results, giving a description of the broth in each flask.
2. Combine your results with those of other groups in the class. What is the advantage of combining results?
3. Were your results similar to Pasteur's? Were the class results similar to Pasteur's?
4. If your results were not similar to Pasteur's, can you suggest any explanation?
5. If you were to repeat the experiment, how could you improve it?

ACTIVITY 1.2 Researching for Mightypharm

Researchers working for the pharmaceutical company Mightypharm were extracting chemicals from a new species of toadstool discovered in the rainforests of Brazil. Several of the chemicals were compounds that had never been found before. The researchers decided that the new compounds might have the potential to be used as antibiotics in the treatment of human bacterial infections.

Imagine that you are one of the Mightypharm researchers and your task is to test the new compounds with the goal of eventually producing an antibiotic that can be used to treat bacterial infections in human patients.

- *Stage 1:* Propose a hypothesis linking the two variables (chemical compounds and effect on bacteria). Describe how you would test the hypothesis to find out whether any of the compounds are effective in killing bacteria. Make your description detailed enough for someone else to follow and carry out the same tests that you propose. Describe how you would present your results and what sort of result would indicate that a compound had potential for use as an antibiotic.
- *Stage 2:* Suppose that one of the compounds tested in stage 1 showed promise as an antibiotic. Describe how you would test that compound on animals to find out whether it worked and whether there were any side effects from use of the compound. Make your description detailed enough for someone else to follow your procedure exactly, and remember that there are ethical considerations relating to the use of animals in research.
- *Stage 3:* The promising compound has successfully passed stages 1 and 2. Describe how you would carry out human trials on the compound. Also, describe how you would deal with any ethical issues that may arise.

In writing your descriptions of stages 1, 2 and 3 you may wish, or your teacher may ask you, to present your material as a paper to be published in a scientific journal. Refer to page 14 for the format of a scientific report.

FURTHER INVESTIGATION

You may wish to investigate how a prescription drug currently in use was discovered, developed and marketed.

Review questions

1 Explain the difference between:
  a observations and surveys
  b longitudinal studies and case studies.
2 What is a controlled experiment?
3 **a** What is a hypothesis?
  b Can a hypothesis be proved? Explain.
4 What is a literature review and what are some of the reasons for carrying out such a review?
5 List four principles that must be satisfied if an investigation is to be ethical.
6 **a** Explain the difference between the dependent and the independent variable in an experiment.
  b Explain the difference between controlled and uncontrolled variables.
7 What is the difference between the validity and the reliability of an investigation?
8 Explain the difference between qualitative and quantitative data.
9 Describe how you would calculate the mean of a set of measurements.
10 What are outliers? Should outliers be excluded when drawing conclusions from a set of data?
11 What is a peer review? Why are peer reviews used?
12 Describe some of the points that should be included in the discussion section of a scientific report.
13 What is an 'error' when discussing a scientific investigation?

Apply your knowledge

1 Re-read the account of Florey's experiment in which he injected mice with penicillin (pages 3–4). What variables did Florey control in his experiment?
2 What did Albert Einstein mean when he said: 'No amount of experimentation can ever prove me right; a single experiment can prove me wrong'?
3 What type of investigation would be best for finding a solution to the following problems? Explain the reasons for your choice in each case.
  a Can people taste the difference between two different brands of milk chocolate?
  b What proportion of students in your school are left-handed?
  c What is the ratio of males to females in your Human Biology class?
  d How has a particular person's growth rate changed from birth to age 15?
4 In addition to physical activity that is part of their job or daily routine, many people deliberately exercise by going to a gym or by walking or jogging. Describe how you would conduct a survey to find out the average amount of time the teachers at your school spend on deliberate exercise.
5 The table below shows the systolic blood pressure of students in a Year 12 Human Biology class.

Table 1.2 Systolic blood pressures of Year 12 students

Systolic blood pressure (mmHg)					
109	123	141	115	131	126
144	138	106	115	49	109
125	132	128	114	116	120
195	143	132	116	113	

- **a** Are there any obvious outliers in the data in Table 1.2? If so, which are the outliers and why should they be regarded as outliers?
- **b** Calculate the mean systolic blood pressure for the class, excluding any outliers.
- **c** What is the range of blood pressures in the class?
- **d** What percentage of students had a blood pressure of 130 mmHg or higher?
- **e** The average systolic blood pressure for adults is 120 mmHg. What proportion of students have blood pressures above this average?

6 In 2003 a team of Australian anthropologists discovered skeletal remains on the Indonesian island of Flores. One skeleton was of a small human with a small brain, and dating showed it to be 18 000 years old. The team claimed it was a new species of human and named it *Homo floresiensis*. Experts are divided on whether the discovery is a new type of human or whether there is some other explanation for the small stature and small brain. This is a good example of scientific debate about the meaning of data. Use the Internet to find out some of the hypotheses put forward to explain why the skeleton is really our own species, *Homo sapiens*.

7 If you randomly draw a card from a standard pack of 52 playing cards,

- **a** what is the probability that you will draw a spade?
- **b** what is the probability that you will draw a king?

8 Researchers investigating the benefits of exercise in preventing heart disease studied the health outcomes for women after participating in an exercise program. They calculated the risk of heart disease at 0.18 with a confidence interval of 0.04 to 0.80 at the 95% confidence level. Explain what the data mean.