# Key Science Skills
- develop aims and questions, formulate hypotheses and make predictions
- plan and undertake investigations
- comply with safety and ethical guidelines
- conduct investigations to collect and record data
- analyse and evaluate data, methods and scientific models
- draw evidence-based conclusions
- communicate and explain scientific ideas.

## Steps in Psychological Research
- Steps in psychological research 16
- Research methods 19
- Types of data 56
- Organising, presenting and interpreting data 58
- Ethics in psychological research and reporting 69
- Use of animals in psychological research 75
- Reporting conventions 78
The development of key science skills is a core component of the study of VCE Psychology and applies across Units 1 to 4 in all areas of study. These skills are applied when planning and conducting your own research, as well as when analysing or evaluating the research of others (VCE Psychology study design 2016, p. 11). All skills are evident in the different research methods commonly used by psychologists.

A research method is a particular way of conducting a research study (‘investigation’) to collect accurate and reliable data on a topic of interest. For example, experiments and self-reports are different types of research methods.

In an experiment, the researcher manipulates and controls a research participant’s experiences in some way to find out whether this causes a particular response. For example, a researcher may conduct an experiment to find out whether learning a list of previously unseen words by repeating the words and their definitions aloud three times improves performance on a test of those words. This could be compared with learning the words and their definitions by writing them down three times.

For some research topics, the most appropriate way of collecting data may be to ask participants about their thoughts, feelings or behaviour. This is when a self-report method such as a questionnaire or interview could be used. For example, participants may be asked questions about their attitudes towards school and the reasons for their attendance and absences. The questions may be presented in a questionnaire for which participants respond to a written set of questions, or by interviewing participants and recording their verbal answers in writing or electronically.

In some cases it is appropriate to use a combination of research methods to collect data. For example, a researcher conducting an experiment on different learning techniques used by students when studying for an exam may also interview the students to find out what motivates them to study for an exam, when they study and how much time they spend.

These are just some of the many research methods available to psychologists. Other methods include cross-sectional studies, case studies and observational studies.

In this chapter we examine the research methods and key science research skills prescribed in VCE Psychology. We start with an overview of the steps involved in conducting scientific research in psychology.

**FIGURE 2.1** Experiments in psychology can be conducted in both laboratory and field settings. (a) In a laboratory setting, social behaviour may be observed in a controlled situation established by the researcher. (b) In a field setting, social behaviour may be observed in a real-world situation such as a bar, but less control of conditions is possible.
**Steps in Psychological Research**

Most of what psychologists know about behaviour and mental processes is based on empirical evidence. *Empirical evidence* is data collected through systematic observations and/or carefully controlled experiments. This type of evidence allows psychologists to draw accurate conclusions which are more likely to be valid and free from personal biases, as compared to our ‘common sense’ conclusions based on everyday observations of behaviour.

Because psychological research involves the collection of empirical evidence using scientific attitudes and practices, it is often called *empirical research*. Generally, a systematic step-by-step procedure is followed. There are variations of the steps and they do not guarantee that accurate and justifiable conclusions will be reached. However, the steps ensure empirical evidence is collected and minimise the chance for bias, errors, faulty conclusions and results that cannot be tested through replication.

**Step 1: Identify the research topic**

The first step in conducting psychological research is to identify the topic, question or issue of research interest. For example, a researcher might be interested in ways of reducing the number of accidents caused by red P-plate drivers. To do this, they may conduct a *literature review*. This involves searching major psychological journals to find and review published reports of research that have already been conducted on this topic. For example, they may consider research that has been conducted on defensive driving programs such as the Smith System.

The Smith System involves five *rules* to train the eyes to identify what is important when driving. As shown in figure 2.3, the rules are:

1. **Aim your vision high** *(to steer accurately and anticipate problems)*.
2. **Keep your eyes moving** *(avoid staring and stay alert)*.
3. **Look at the total driving picture** *(don’t focus your eyes on one area of the road)*.
4. **Leave yourself a ‘way out’**.
5. **Look for a position on the road that ensures other drivers can see you**.

Conducting a literature review enables the researcher to become more familiar with their topic of research interest. It also enables them to refine their ideas and propose a relevant research question that will be the aim and therefore the focus of their research; for example, ‘Does training red P-plate drivers with the Smith System help to reduce the number of accidents they cause?’

**Step 2: Formulate the research hypothesis**

When the topic has been identified and refined, the next step is to formulate a hypothesis for the research. This is essentially a thoughtful prediction about the results that will be obtained when the hypothesis is tested.

An example of a research hypothesis for the driver study is ‘Red P-plate drivers who receive defensive driving training will make fewer driving errors than red P-plate drivers who have not received defensive driving training’.

The research hypothesis must be testable and written as a very specific statement. The testability of the hypothesis for the Smith System study is reflected by the way in which it is stated. For example, ‘number of accidents’ is defined more specifically as ‘driving errors’ and driving errors will be measured through a practical driving test in a driving simulator. There are several other important characteristics of a research hypothesis. These are described on pages 21–2.

**Step 3: Design the research**

The third step is to select the research method(s) and design the most appropriate procedures to collect the data required to test the hypothesis (but this is also considered when the hypothesis is being formulated). Generally this involves selecting a specific research method (such as a particular type of experiment) and determining the procedures that will be used.

When designing the research, the researcher must consider the hypothesis, decide which participants will be studied, how many there will be, and how they will be selected and be allocated to different groups that may be used in the study. The participants’ responses provide the data that become the results for the research.

**Figure 2.2** Psychological research may be conducted on ways of reducing the number of accidents caused by red P-plate drivers.
1. Aim your vision high (to steer accurately and anticipate problems).

2. Keep your eyes moving (avoid staring and stay alert).

3. Look at the total driving picture (don’t focus your eyes on one area of the road).

4. Look for and leave yourself a ‘way out’.

5. Look for a position on the road that ensures other drivers can see you.

**FIGURE 2.3** The Smith System involves five rules to identify what is important when driving.

There are also organisational matters to attend to; for example, ensuring there is access to a driving simulator and preparation of materials such as consent forms and instructions for participants.

When designing the research it is vital that relevant ethical standards and guidelines are followed to ensure the rights and wellbeing of all participants are protected. These apply to all stages of the research (including the report) and are described on pages 69–72.

There are advantages and limitations of each type of research method that may be used to test a hypothesis. Some are more suited to particular research hypotheses and data collection than others.

In the driver study, the researcher may decide to conduct an experiment using an independent groups design. Ten male and 10 female red P-plate drivers with a similar amount of driving experience in city and regional areas as participants in the study could be used. Half the participants would be in one group and receive a defensive driving training session using the Smith System. The other half would be another group and not receive the training so that the potential benefits of the training could be compared.

Alternatively, the researcher may decide to give questionnaires to a number of driving instructors who have taught young people, with and without the Smith System, to obtain information based on their experience as to which of the two methods resulted in better driving skills. Another option would be to conduct an experiment and interview the research participants.

**Step 4: Collect the data**

The fourth step involves actually collecting the required data in order to determine the results and conclusions that can be drawn.

Based on their plan and research design, the participants and materials are organised and the study is conducted.

Data may be primary or secondary and quantitative or qualitative. These different types of data are described on pages 56–7.

**Step 5: Analyse the data**

When the data have been collected they are in a ‘raw’ format because they have not been processed or analysed. The next step in psychological research is therefore to summarise, organise and represent the raw data in a logical and meaningful way to help determine whether the hypothesis is supported and to draw conclusions.

This usually involves breaking down a large set of numbers into smaller sets (e.g. raw data summarised as percentages in a table or graph) or even a single number or two (e.g. a mean score or standard deviation). The researcher is then better able to consider the data when determining whether the hypothesis is supported based on the results obtained.
Step 6: Interpret and evaluate the results

Once the data has been analysed, it needs to be interpreted and evaluated. This includes drawing conclusions about what the results mean and identifying, describing and explaining limitations.

One type of conclusion relates directly to the research hypothesis. The focus of this conclusion is on whether or not the results support the hypothesis. The researcher may also consider how widely the findings of the research can be applied. Descriptive and inferential statistics may be used by the researcher to help decide what the results mean and what conclusion(s) can legitimately be made (see pages 58–67).

A researcher usually studies a relatively small number of participants who are selected from the bigger group of interest; for example, 10 male and 10 female red P-plate drivers aged between 18–21 years, rather than all red P-plate drivers. Of particular interest to the researcher is whether the results obtained from a relatively small number of cases (the red P-plate drivers in the study) can be extended to apply to the bigger group of red P-plate drivers from which the smaller group was selected, to the entire group of red P-plate drivers, to green P-plate drivers or even all drivers.

With reference to the driving study, if the results for the experiment indicated that the red P-plate drivers who were trained with the Smith System made significantly fewer errors than the drivers who did not receive the training, then the researcher would interpret the results as providing support for the hypothesis. They would conclude that using the Smith System to teach defensive driving techniques to drivers reduces the likelihood of red P-plate drivers having an accident. On the basis of this conclusion, the researcher may also tentatively (cautiously) conclude that the finding may apply to the bigger group of red P-plate drivers targeted for research or even all red P-plate drivers.

**FIGURE 2.4** Steps commonly followed when planning and conducting psychological research. There are variations of these steps, but all are based on scientific attitudes and practices.
When drawing conclusions, the researcher bases their judgment strictly on what the results show. They would also seek to identify, comment on and take account of any limitations of their study, particularly research procedures or unexpected events that may have influenced their results in an unwanted way.

**Step 7: Report the research and findings**

The final step of psychological research involves preparation of a report for others who may be interested in the research, its results and findings. Typically, psychological researchers prepare a detailed written report which they seek to get published in a professional journal. Each of these journals has reviewers who critically evaluate the research report and it is only on the basis of peer reviews that the research may be accepted for publication. In addition, a poster report may also be prepared for display and discussion at conferences or meetings with other researchers.

The report prepared for publication follows a strict format and describes in detail all aspects of the research, including relevant background information, how the research was conducted, the results and findings, any limitations which may have impacted on the results, and a list of references used in preparing for the study and writing the report. A poster report has a format more suited to display and is less detailed.

Reporting the research and its findings is a very important part of the research process. It is the way other researchers find out about research which has been conducted and the way scientific progress is achieved. It also enables the general public to benefit from the findings of research.

Importantly, the reporting process places the specific study and its research procedures under the critical eye of other psychologists and researchers; for example, to check the accuracy of the findings and to consider alternative conclusions that may be valid. It also enables replication by other researchers — to repeat the study in order to test the accuracy of the results or to test the relevance of the study or results to other groups or situations.

**RESEARCH METHODS**

Research methods are the means or ‘tools’ for observing, measuring, manipulating or controlling what takes place in psychological investigations. Each method has its specific purposes, procedures, advantages and limitations.

The choice of research method depends on which is most appropriate for the specific topic of research interest and hypothesis being tested. This is not unlike the choice made by a motor mechanic when selecting tools to repair a car engine. Their selection will depend on the specific engine problem in need of repair and the work that needs to be done to fix it. Each tool will have a specific use and way of being used. Similarly, each research method has a particular logic underlying its use and how it is used.

Sample selection and formulation of a research hypothesis are common to all psychological investigations requiring participants and are undertaken early in the research process. We consider each of these important procedures before examining the research methods prescribed in VCE Psychology and the specific features that distinguish them from each other.

**FIGURE 2.6** A motor mechanic selects the best tools to solve a mechanical problem, just as a researcher chooses the best research method(s) to solve a research problem.

**FIGURE 2.5** Two of the Australian journals in which psychologists may publish their research reports.
Sample and population

Psychologists mostly conduct research studies with people and, in some cases, animals. The participants being studied are called the sample.

A sample is a subsection or smaller group of research participants selected from a larger group of research interest. For example, suppose that a researcher is interested in conducting an experiment to find out whether children who attended a childcare centre during their preschool years have better language skills than children who did not attend a childcare centre. It would be impractical to test every child who attended a childcare centre and every preschool child who did not. Researchers therefore select a sample with whom they conduct their research. If the sample is selected in a scientific way, the results obtained for the sample can then be applied to the population from which it was drawn or even other groups or situations (assuming the results are valid and reliable).

In scientific research, the population does not necessarily refer to all people (or animals) in the world, in a country, or even in a particular city or area. The term population refers to the entire group of research interest from which a sample is drawn and to which the researcher will seek to apply (generalise) the results of their investigation. A population of interest may be all preschool children, all blonde-haired females, all VCE Psychology students, all female VCE Psychology students, all Catholic school educated boys, or all male chimpanzees born in captivity.

A population in a study does not always refer to living things. A population could also be measurable things such as all community health centres in the Goulburn Valley region, all admissions at the Royal Melbourne Hospital, all VCE exam results in English in 2016, all days of school missed by year 9 students, all brands of sports shoes, all calls to the Kids Helpline telephone number, or any other specific source of data.
Research hypothesis

In psychology as in other sciences, different research methods are used to test one or more hypotheses relevant to a topic, question or issue a researcher aims to study. A **research hypothesis** is a testable prediction of the relationship between two or more variables (events or characteristics). It states the existence of a relationship between the variables of interest and the expected relationship between them. For example, it may be a prediction about the relationship between:

- attending a revision lecture (one event) and the score achieved for a psychology test (another event)
- biological sex (one characteristic) and finger dexterity (another characteristic)
- reading (one event) and brainwave activity (one characteristic).

The research hypothesis formulated for a specific investigation is essentially an educated or thoughtful guess about what the results will be. It is usually based on knowledge of other research findings or theories on the topic being studied. This is why it is often referred to as an ‘educated’ guess.

The research hypothesis is formulated before the study is conducted and provides a focus for the research. A research hypothesis usually has the following characteristics:

- refers to events or characteristics that can be observed and measured and is therefore testable
- states the existence of a relationship between two or more variables
- states the expected relationship between the variables (e.g. how one variable will influence the other)
- based on observations, a theory, model or research findings
- prepared as a carefully worded written statement (rather than a question)
- expressed clearly and precisely (rather than vaguely and generally)
- written as a single sentence.

In some cases, the research hypothesis may also refer to the population from which the sample was drawn and therefore the larger group about which the researcher intends to draw conclusions.

Research hypotheses for the examples could be:

- **VCE Units 1 and 2 Psychology students who attend a revision lecture before a test** will achieve a better score on the test than students who do not.
- **Females** have better finger dexterity than males.
- **Beta brainwaves** are predominant when primary school children are reading.

It is not always possible to be entirely certain about the accuracy of a prediction within a hypothesis. This is mainly because the researcher does not necessarily know or can control the influence of the many different variables that can affect the thoughts, feelings or behaviour being studied. Nonetheless, many researchers would consider it pointless to conduct a study when the outcome is certain.

Research hypothesis versus theory and model

A research hypothesis is different from a theory and model. A research hypothesis is a specific prediction that guides the collection, analysis, interpretation and evaluation of data that has been collected to test it. In contrast, a **theory** is a general explanation of a set of observations about behaviour and/or mental processes which seem to be related. A **model** tends to focus more on representing how some behaviour and/or mental process(es) could, should or does occur. It is often presented in the form of a diagram with boxes and arrows to organise and show relationships between different concepts. Figure 2.9 shows an example of a model of human memory.

Theories and models vary in scope, complexity and detail. Some are essentially a hypothesis that has been restated. Others explain many interrelated
research findings and ideas. Along with explaining existing results, a useful theory or model generates new hypotheses and guides further research. Many theories or models of child development, personality, learning, remembering, forgetting and so on, are the products of psychological research and have generated valuable new research. In addition, some theories have generated new models and some models have generated new theories.

Whatever their scope — from tiny to vast — theories and models serve a gap-filling function. They explain how findings and ideas fit together and what they mean, thereby making psychology a discipline that does more than report isolated facts. Psychologists prefer testable theories and models because they can be confirmed or refuted (rejected) by further scientific research. Therefore, theories and models tend not to be judged in terms of their accuracy but rather in terms of their usefulness. This means that a theory or model tends not to be considered as right or wrong. Instead, it is simply regarded as more or less useful.

**FIGURE 2.8** Theories and models are revised and expanded to reflect relevant research findings. New or revised theories and models lead to new observations or questions that stimulate new research.

**FIGURE 2.9** A contemporary model of human memory, representing it as a multi-storage system through which information flows.
Experimental research

One of the most scientifically demanding and controlled research methods in psychology is the experiment. An experiment is used to test a cause–effect relationship between variables under controlled conditions; for example, to test whether talking on a hand-held mobile phone while driving (one variable) causes or influences a change in driver reaction time (another variable), or whether receiving a reward (one variable) has an effect on exam performance (another variable).

Essentially, an experiment enables the researcher to investigate whether there is a causal (cause–effect) relationship between two or more variables; for example, if talking on a hand-held mobile phone while driving causes drivers to react more slowly (and therefore increases the likelihood of an accident).

There are different types of experimental designs that vary in terms of their specific procedures and complexity. All experiments, however, have a number of common features. We consider the essential features of an experiment and why this particular research method can be used to investigate causes of behaviour and links between behaviour and mental processes.

**Independent and dependent variables**

In a research study, a variable is something that can change (‘vary’) in amount or type and is measurable. For example, the time that it takes for a newborn infant to distinguish between different shapes is a variable that changes in amount (time) and type (square or circle). How long it takes to distinguish between the different shapes can also be measured. Although personal characteristics of an individual, such as biological sex, blood type and racial or ethnic background are all inborn and do not actually change, in psychological research they are still considered to be variables because they can be of different types and are measurable. For example, ‘male’ and ‘female’ are two types of biological sex and ‘O’, ‘A’, ‘B’ and ‘AB’ are four different blood types.

If the experiment involves testing whether a particular anger management technique reduced the incidence of road rage in people who had previously been convicted of road rage, the two variables being tested would be (1) the anger management technique and (2) the incidence of road rage. These are two different types of variables called independent variables and dependent variables.

**Independent variable**

Every experiment has at least one independent variable and one dependent variable. In a simple experiment, one of these variables is manipulated or changed by the researcher to observe whether it affects another variable and what those effects are. The variable that is manipulated in order to measure its effects on the dependent variable is called the independent variable (IV). It is sometimes referred to as the ‘treatment’ variable to which participants may be exposed (or not exposed).

In terms of cause and effect, the IV is said to be the cause of any changes that may result in the dependent variable. For example, in the road rage experiment, the IV would be the anger management technique. The researcher would have control over and can therefore ‘manipulate’ which participants would learn the anger management technique and which participants would not, in order to test the effect(s) of the technique on the incidence or extent of road rage-related behaviour; that is, the dependent variable.

![Diagram](FIGURE 2.10 The IV and DV in experimental research designed to investigate a technique that may reduce the incidence of road rage)
Dependent variable

The variable that is used to observe and measure the effects of the IV is called the dependent variable (DV). The dependent variable is often the responses made by the participants and it usually has a numerical value. It is called the dependent variable because whether or not it changes and the way in which it changes ‘depends’ on the effects of the independent variable.

In terms of a cause–effect relationship, the DV is the effect(s) caused by manipulation of the IV; that is, exposure or non-exposure to the IV. In the road rage example, the DV is the measured change in the amount of road rage behaviour displayed by participants as a result of using or not using the anger management technique — the IV.

Operationalising independent and dependent variables

Operationalising the IV and DV involves defining them in terms of the specific procedures or actions (‘operations’) used to measure them. This is an important step because many of the behaviours and mental processes psychologists investigate can have different meanings and can therefore be defined and measured in more than one way.

For example, suppose that a researcher wants to find out whether noise has an effect on problem-solving ability. What is meant by ‘noise’ (the IV) and ‘problem-solving ability’ (the DV), and how will these variables be defined and measured? Will the ‘noise’ be music? If so, will it be classical music, rock music or some other type of music? Will the noise be people talking, whales communicating, an engine revving, the sound of a plane flying overhead or a combination of different types of noises? Will the noise be loud, medium or soft? Will the noise be heard continuously or irregularly?
Similarly, consider ‘problem-solving ability’. What type of problem? Will it be a personal problem, a problem involving someone else or an intellectual problem? Will the problem be simple or complex? Will the problem be presented orally, in writing or audiovisually? Will the problem have one solution or several solutions? Furthermore, precisely how will ‘ability’ to solve the problem be measured? Getting the problem right or wrong? Solving it quickly or slowly? Both accuracy and speed?

When the IV and DV have been operationalised, they are often stated this way in the hypothesis (but this is not essential). In the study on noise and problem-solving ability, the operationalised variables could be stated in the research hypothesis as follows:

‘Units 1 and 2 VCE students who listen to loud rock music when solving previously unseen written problems will solve fewer problems than students who do not listen to loud rock music’.

Note how the IV and DV have been operationalised:
- IV: continuously listening to loud rock music throughout a one hour session
- DV: the number of problems that are solved.

Note too that the research hypothesis above is also stating the population from which the sample is drawn and therefore the larger group about which the researcher intends to apply the results.

Operationalising the IV(s) and DV(s) ensures that these variables are also precisely defined. The resulting definitions are sometimes referred to as operational definitions.

There are three important benefits of variables being defined precisely through operationalisation:

1. It helps ensure the independent and dependent variables are testable and therefore that the research hypothesis is testable.
2. All researchers involved in conducting the experiment know exactly what is being observed and measured, which helps avoid experimenter biases and differences that can affect the results.
3. When the variables are defined in a very precise way, another researcher interested in the results, or perhaps even doubting them, will be able to repeat the experiment in order to test (‘check’) the results obtained for accuracy or for relevance to other groups or situations.

When a study is replicated using a similar sample and similar results are obtained, there is greater confidence in the validity and reliability of the results.

![FIGURE 2.13 Does facial piercing make a person more or less attractive? The answer depends on how you operationalise ‘facial piercing’ and ‘attractive’.

---

**TABLE 2.1 Ways in which IVs and DVs can be operationalised**

<table>
<thead>
<tr>
<th>Research question of interest</th>
<th>IV example</th>
<th>DV example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do students learn more effectively in early morning or late afternoon classroom lessons?</td>
<td>• time of lesson</td>
<td>• score on a test of recall (amount of information remembered)</td>
</tr>
<tr>
<td>If a teacher ignores a student’s attention-seeking behaviour in class, will this strategy reduce the student’s attention-seeking behaviour?</td>
<td>• teacher not paying attention to attention-seeking behaviours</td>
<td>• frequency of attention-seeking behaviours</td>
</tr>
<tr>
<td>Does playing violent video games cause aggressive behaviour?</td>
<td>• a video game classified by the Commonwealth Government censors as violent</td>
<td>• number of presses of a button that administers a shock to another student</td>
</tr>
<tr>
<td>Does allowing a child to sleep in the same bed as their parents result in the child being overly attached to the parents?</td>
<td>• child sharing bed with both parents over a specified period of time</td>
<td>• frequency of separation anxiety behaviours when either or both parents leave the child alone with a stranger</td>
</tr>
<tr>
<td>What types of jokes are funny to people of different cultural backgrounds?</td>
<td>• different types of jokes</td>
<td>• number of audible laughs detected by an audiometer and number of smiles detected by an electromyograph (measures facial muscle contractions)</td>
</tr>
</tbody>
</table>
LEARNING ACTIVITY 2.4

Review questions
1. Explain the meaning of the phrase ‘operationalising the IV and DV’ with reference to an example.
2. What are three benefits of operationalising the variables for a research hypothesis?
3. Suggest how three of the following variables of research interest could be operationalised.
   a. memory
   b. happiness
   c. attraction
   d. intelligence
4. Suggest operationalised IVs and DVs for three of the following research topics:
   a. Does regular exercise improve psychological wellbeing?
   b. Do people drive less safely when feeling stressed?
   c. Do people talk more after they have eaten than when they are hungry?
   d. Does perception of time change when in a relaxed state?

Experimental and control groups
In a relatively simple experiment, the participants are allocated to one of two groups. One group of participants, called the experimental group, is exposed to the independent variable (i.e. the ‘treatment’). This group is said to be in the experimental condition. A second group of participants, called the control group, is not exposed to the IV. This group is said to be in the control condition.

For example, consider an experiment to investigate whether alcohol consumption affects driving ability. In this experiment, the IV which the researcher will ‘manipulate’ is the amount of alcohol consumed by research participants and the DV will be the number of driving errors made. The experimental group would be tested on their driving skills in a driving simulator after having consumed alcoholic drinks (the experimental condition) and the control group would be tested on their driving ability in the driving simulator after having consumed non-alcoholic drinks (the control condition).

The control group provides a ‘baseline’ or standard against which the researcher can compare the performance of the experimental group in order to determine the effect of the IV on the DV. If the driving performance of the experimental group is significantly worse than the driving performance of the control group, the researcher may conclude that the IV (consumption of alcohol) affected the DV (driving errors). A flow chart summary of this experiment is shown in figure 2.14.

The experimental group and the control group need to be as similar as possible in the spread of personal characteristics of participants that can cause a change in the DV. For example, one group should not have significantly more participants who are more experienced (‘better’) drivers so that this does not become a possible reason for differences in the number of driving errors between the groups that may be recorded. It is also important that both groups are treated the same, except for the time when the experimental group is exposed to the IV. Both of these conditions are necessary so that if a change occurs in the experimental group and does not occur in the control group, the researcher can be more confident in concluding that it was the IV that probably caused the change.

Sometimes the experimental condition and control condition are collectively called experimental conditions, which literally means ‘all the conditions of the experiment’. When this expression is used, the condition in which the IV is present may be referred to as the ‘treatment condition’ because the IV is the ‘treatment’ to which the participants are exposed.

![Flow chart of the experiment testing the effect of alcohol consumption on driving ability](image)

**Hypothesis:** Alcohol consumption impairs the driving ability of university students.

**Participants:** 100 third-year university students who responded to an advertisement for research participants. Equal numbers of male and female participants. Ages range from 20–40 years.

**Experimental group:** 50 students (25 male, 25 female) who consumed alcoholic drinks  
**IV:** alcohol consumption

**Control group:** 50 students (25 male, 25 female) who consumed non-alcoholic drinks

Use of driving simulator to test precision of driving skills

**Results:** Participants who consumed alcohol made many more driving errors than participants who did not consume alcohol.

**Conclusions:** Alcohol consumption causes more driving errors, resulting in impaired driving ability in third-year university students. Alcohol consumption may also cause impaired driving ability in all drivers.

**Figure 2.14** A flow chart of the experiment testing the effect of alcohol consumption on driving ability.
Extraneous variables

In an experiment to test whether sleep deprivation causes headaches, the IV is the amount of sleep obtained and the DV is how often a headache is reported. As shown in Table 2.2, the results of this research suggest that the frequency of headaches is likely to increase if people experience six or fewer hours of sleep.

**TABLE 2.2 Frequency of headaches reported and amount of sleep**

<table>
<thead>
<tr>
<th>Hours of sleep</th>
<th>Frequency of headaches reported</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
</tr>
<tr>
<td>≥8</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>38</td>
</tr>
<tr>
<td>≤6</td>
<td>15</td>
</tr>
</tbody>
</table>

However, what would happen if participants who had eight or more hours of sleep also took ‘sleeping pills’ which reduced the likelihood of headaches occurring? Or participants who had six or less hours of sleep were also experiencing considerable stress in their lives or were prone to getting headaches? Or participants had different perceptions of what a headache is and what was reported by one participant as a headache was not reported as a headache by another?

There are variables other than the IV that might influence the DV and therefore the results of an experiment. Researchers try to predict what these might be when planning their experiment. Then, they design the experiment to control or minimise the influence of other, ‘extra’ variables; that is, extraneous variables.

In an experiment, an extraneous variable is any variable other than the IV that can cause a change in the DV and therefore affect the results in an unwanted way. When one or more extraneous variables are present in an experiment, they can make it difficult to conclude with confidence that any change in the DV was caused solely by the presence of the IV and not because of some other variable. This is why they are ‘unwanted’.

In the sleep study described previously, extraneous variables that may have caused or contributed to headaches developing or not developing can include the amount of stress in the person’s life, illness (such as a virus), eye strain, excessive alcohol consumption or use of particular medication. Thus, in the group who had six or less hours of sleep, the greater likelihood of experiencing a headache may not have been a result of insufficient sleep if one or more extraneous variables were present. In order to conclude that the frequency of headaches will increase as a result of a reduction in the amount of sleep obtained, variables other than the IV that can impact on the DV must be controlled or eliminated.

Potential extraneous variables are often identified prior to the research. Sometimes, the researcher does not become aware of relevant extraneous variables until after the experiment has commenced; for example, during the experiment or when evaluating the experiment after it has been conducted. In some cases, the researcher remains unaware of relevant extraneous variables until another researcher points them out after reading the report on the experiment.

There are potentially many extraneous variables that may affect the DV of an experiment and it can be difficult for the researcher to predict and control all of them. Consequently, researchers tend to focus on controlling those variables that are likely to have a significant effect on the DV. For example, in an experiment to determine the softest noise a person can hear, it would be very important to control background noise. However, in an experiment to test the effect of caffeine on performance of some physical task, background noise may not be so critical.
**Confounding variables**

Every experiment used in psychological research is designed to answer the same basic question: *Does the IV cause the predicted change in the DV?* The researcher recognises that there are other variables that may affect participants' responses (the DV), such as all those variables collectively referred to as extraneous variables.

Extraneous variables are inevitable and do not pose a problem if controlled in an appropriate way. By strictly controlling unwanted effects of relevant extraneous variables on the DV, the effects of the IV on the DV can be isolated. If there is a measurable change in the DV, then the researcher can confidently conclude that the IV caused the change in the DV. If a variable that can affect the DV is not controlled, then its effect on the DV may not be able to be clearly distinguished from that of the IV. When this happens, the uncontrolled variable is referred to as a confounding variable.

A **confounding variable** is a variable other than the IV that has had an unwanted effect on the DV, making it impossible to determine which of the variables has produced the change in the DV. Basically, a confounding variable is a second unwanted IV that has influenced the DV together with the IV. It is called a confounding variable because its effects are entangled and therefore potentially ‘confused’ with those of the IV, thereby preventing the researcher from concluding that the IV caused the predicted (hypothesised) change in the DV.

The presence of one or more confounding variables does not necessarily mean that the IV did not cause the changes in the DV. However, the presence of a confounding variable suggests that there may be one or more alternative explanations for the results obtained in the experiment. The more alternative explanations there are for the results, the less confident the researcher will be that the IV alone was responsible for the results.

A confounding variable is different from an extraneous variable. A confounding variable produces a measurable change in the DV. This change is consistent with what was predicted in the hypothesis, whereas an extraneous variable may or may not affect the DV. What both types of variables have in common is that they create problems for the researcher in isolating the real effects of the IV.

An experiment with uncontrolled variables compromises the validity and interpretation of the results. The more alternative explanations there might be for an observed result, the less confidence a researcher will have in their research hypothesis, which states that the IV will be the cause of a particular result. Because humans are complex and there are often multiple causes of how they may think, feel or behave in any given situation, good experimental design involves anticipating potential extraneous and confounding variables and developing strategies to minimise their influence or ensuring that extraneous variables do not become confounding variables.

![Diagram showing the relationship between independent variable, dependent variable, extraneous variable, and confounding variable.](image)

**FIGURE 2.16** When an extraneous variable influences the DV in an unwanted way, it may be difficult to isolate its effects from those of the IV.

**LEARNING ACTIVITY 2.6**

**Review questions**

1. What are extraneous and confounding variables?
2. In what way are extraneous and confounding variables alike yet different?
3. When is it best to identify extraneous variables? Explain your answer.
4. Explain why the presence of extraneous and/or confounding variables is problematic for the researcher.
5. Give an example of an extraneous variable that may be relevant in one experiment and therefore require control but irrelevant in another experiment and not requiring control. Explain your answer.

6. For each of the following research topics, identify the IV, the DV and three potential extraneous variables that could affect the DV.
   (a) The effect of shyness on the ability to make new friends at school
   (b) Whether meditation can improve performance on a VCE English exam
   (c) Whether males are more willing than females to taste different foods
   (d) Whether students who have breakfast concentrate better in class
   (e) Whether having a pet in an aged-care nursing home improves happiness for elderly people who live there
**Identifying potential extraneous and confounding variables**

Researchers have described many variables that can be extraneous or confounding variables in an experiment. VCE Psychology prescribes the study of individual participant differences, use of non-standardised instructions and procedures, order effects, experimenter effect and placebo effect. We examine each of these in turn and then consider how researchers can minimise their potential influences.

**Individual participant differences**

Individual participant differences refer to the unique combination of personal characteristics, abilities and backgrounds each participant brings to an experiment. These participant variables, as they are sometimes called, may be biological, psychological or social in nature. For example, some participants will be more or less easy going, anxious or motivated than others. They will also differ in a wide range of mental abilities such as intelligence, learning, memory, reading comprehension and problemsolving skills, as well as physical abilities such as strength, athleticism, eye–hand coordination and finger dexterity. Furthermore, they will differ in such variables as sex, age, educational background, social relationships, ethnicity, cultural experiences and religious beliefs.

Any one or more of these variables can affect how participants respond in an experiment. However, they are expected and the researcher focuses on minimising the influence of those participant variables other than the IV that could have a measurable effect on the DV if left uncontrolled.

Consider, for example, an experiment to test whether ignoring attention-seeking behaviour of children who misbehave in class will reduce the frequency of their attention-seeking behaviour. A reduction in the number of times attention-seeking behaviour has occurred after a month of ignoring this type of behaviour may not only be a result of ignoring the misbehaviour. Variables relating to the children or their respective personal experiences can impact on their changed behaviour. For example, a child's family situation may have become more or less unsettled or their behaviour may change irrespective of the researcher's experimental treatment. A child’s health, mood or peer relationships may also have an impact on whether or not they use attention-seeking behaviour and how often they may do so.

Consequently, the researcher will try to ensure that the influence of these specific variables is minimised and will do so before the experiment is conducted.

**Use of non-standardised instructions and procedures**

The instructions and procedures used by the researcher can also impact on how participants respond, and therefore on the results. For example, suppose that a researcher is interested in studying the effect of prior experience on food perception and eating. The researcher sets up an experiment in which participants must eat different foods that are ‘plated’ in novel ways but which they may refuse to eat. Figure 2.17 shows two examples of the ‘test materials’.

*Imagine* how the results could be affected if participants received different instructions on what the experiment is about, what eating actually means (as compared to tasting), how many foods there are, whether the foods are safe, whether they can refuse to eat, and so on. What if some participants present for the experiment just before they have eaten a meal and others just after? Or what if some participants complete the task alone while the rest have other participants present? What if the researcher laughs at or comments on the responses of some participants but not others?
Generally, procedures involve everything the researcher does to actually conduct their study, including:
- selection of participants
- instructions for participants in different groups
- interaction with participants
- use of materials or apparatus
- use of rooms or other experimental settings
- observation and measurement of variables
- data-recording techniques.

Procedures not only involve what the researcher does but also how the relevant research activities are conducted, including their sequence. When the research procedures (including instructions) are non-standardised, this means that they are not the same for all participants (except for exposure to the IV by participants in the experimental group). Even small variations in procedures may affect participants' responses in unforeseen ways.

An experiment that uses non-standardised instructions and procedures is not strictly controlling all of the possible extraneous and confounding variables that can influence the DV and therefore the results.

Order effects
In some experiments, participants may be required to perform the same type of task twice or even many times under different conditions. For example, in the experiment described earlier to determine the effects of alcohol on driving performance, the same group of participants may be exposed to the control condition for which they do not drink any alcohol before a driving test in a simulator. After a short break, the participants may then be exposed to the experimental condition for which they are given an alcoholic drink before completing the test.

It is possible that the order in which participants experience different conditions can be a problem in an experiment with this type of design.

An order effect occurs when performance, as measured by the DV, is influenced by the specific order in which the experimental tasks are presented rather than the IV. Performing one task affects the performance of the next task. Order effects may change the results so that the impact of the IV may appear to be greater or less than it really is. Two types of order effects that explain how this can occur are called practice effects and carry-over effects.

Practice effects are the influence on performance (the DV) that arises from repeating a task. For example, the participants' performance in the alcohol experiment may be influenced or partly determined by practice. Through repeated experience in the driving simulator, participants may get better at the driving task and perform better on the driving test due to greater familiarity with the simulator and its controls, or by anticipating events designed to cause driving errors that were presented during the first driving test.

Participants' responses can also be influenced by other practice effects. For example, performance may get worse as the experiment proceeds due to tiredness or fatigue. Similarly, their performance may be influenced by boredom due to repeating the same task, especially if the task takes a long time and does not change. Boredom is quite common in experiments in which participants are required to complete many trials or tests.

Carry-over effects are the influences that a particular task has on performance in a task that follows it. For example, if alcohol was given first in the driving simulator task and the task is then repeated without alcohol (in the control condition), a carry-over effect would occur if insufficient time was allowed for the effects of the alcohol in the first condition to wear off. Similarly, if a task (such as taking a test in a driving simulator) happens to be very easy, difficult, frustrating or even anxiety-provoking, the feeling may ‘carry over’, lowering performance the next time the task is completed (driving in the simulator again) in an unwanted way.

Experimenters' expectations about how participants will perform can also influence the results. Occasionally, these expectations can cause participants to behave in ways that are contrary to the conditions they are actually in. This is called the experimenter effect, also sometimes called experimenter bias or research bias, is an unwanted influence(s) on the results which is produced consciously or unconsciously by a person carrying out the research. In an experiment, the effect occurs when there is a change in a participant's response because of the experimenter's expectations, biases or actions, rather than the effect of the IV. A common type of experimenter effect is called experimenter expectancy.

Experimenter expectancy involves cues (‘hints’) the experimenter provides about the responses participants should make in the experiment. In particular, the experimenter's non-verbal communication (‘body language’) can produce a self-fulfilling prophecy — the experimenter obtains results that they expect to obtain. The results may therefore be attributable to behaviour associated with the experimenter's expectations rather than the IV.
Actions that can promote a self-fulfilling prophecy include:

- facial expressions, such as smiling at participants in one group but not at those in another
- mannerisms, such as shaking hands with participants in one group but not with those in another
- tone of voice, such as speaking in a monotone voice to participants in one group and in a more lively way to those in another.

German-born American psychologist Robert Rosenthal has demonstrated the experimenter effect in numerous studies. In one well-known study, 12 university students taking a course in experimental psychology unknowingly became participants themselves. The participants were asked to place rats in a maze. Some were deliberately told that their rats were specially bred to be ‘maze bright’ and would show ‘learning during the first day of running’ in the maze. The others were deliberately told that their rats were ‘maze dull’ and would ‘show very little evidence of learning’. In reality, the rats were all standard laboratory rats and were randomly allocated to each group (Rosenthal & Fode, 1963).

As evident in figure 2.19, the group of apparently ‘maze bright’ rats learned the maze significantly faster than the ‘maze dull’ rats (as measured by the number of errors in the maze).

The researchers concluded that the lower error rate had more to do with the participant’s expectations of their rats than the rats’ actual abilities. They suggested that ‘experimenter expectations’ about their rats’ capabilities caused participants to subtly alter their training and handling techniques, which in turn affected the animals’ learning. The participants were not cheating or purposefully manipulating their results. The participants were thought to have unintentionally and unconsciously influenced the performance of their rats, depending on what they had been told by the experimenter.

**FIGURE 2.19 Mean number of correct responses per rat per day**

**FIGURE 2.20 Example of a maze used for ‘maze running’ laboratory experiments with rats**

### Placebo effect

A *placebo* is an inactive substance or fake treatment. In an experiment, the *placebo effect* occurs when there is a change in a participant’s behaviour due to their belief that they are receiving some kind of experimental treatment and they respond in accordance with that belief, rather than to the effect of the IV. Essentially, the participants’ behaviour is influenced by their expectations of how they should behave.

For example, consider an experiment to test a hypothesis that drinking alcohol makes members of the opposite sex look more attractive. Participants aged over 18 are randomly allocated to either the experimental or control group. The experimental group are given drink containing orange juice with vodka and the control group drink orange juice alone. All participants are told whether their drink contains vodka. After enough time passes for a alcohol to take effect, participants are asked to rate the attractiveness of people in a set of colour photos.

Suppose that the experimental group rated the photos as significantly more attractive than did the control group. The researcher would like to conclude that alcohol caused the difference in perceived attractiveness. However, alcohol consumption may not have been the only variable that impacted on the DV.

Participants who drank alcohol also knew they drank alcohol, and those who did not drink alcohol knew they did not. The act of being given an alcoholic drink by a researcher might have promoted expectations in participants about how they should respond. For example, experimental group participants might have thought that they were given alcohol because they were expected to perceive more people as attractive, so they did.

Because the experimental group received the alcohol and the control group did not, only the experimental group experienced the placebo effect. This means that a confounding variable is present. Therefore the researcher cannot be certain whether it was the effect of alcohol or the placebo effect that caused the performance difference (Stangor, 2004).
Confounding variables can influence the results of the study to be conducted. Variables that are considered to be important are those that can influence the results of the study to be conducted. It must reflect its attributes (characteristics) of the sample more closely when working alone or when working in small groups. Because of a shortage of rooms, the participants working alone completed the problems in a small tutorial room with no windows in the corner of the school library. The participants working in small groups completed the problems in a large classroom with big windows on the first floor of the building (above the library). As hypothesised, the groups performed better than the individuals.

**Ways of minimising extraneous and confounding variables**

When planning an experiment, the researcher will consider potential extraneous and confounding variables. The extent to which these variables are anticipated and controlled will determine the quality of their experiment and its results. Ways of minimising extraneous and confounding variables include use of appropriate sampling procedures for selection and allocation of participants, counterbalancing, single- and double-blind procedures, placebos, standardised instructions and procedures, and use of an appropriate experimental research design.

**Participant selection**

A sample has to be selected in a scientific way so that the results obtained for the sample can be legitimately applied to the population from which it was selected. The process of selecting participants from a population of interest is called sampling.

A key goal of sampling is to ensure that the sample closely represents its population so that the results can be generalised to that larger group. It must reflect its population in all the personal characteristics of participants that are important in the research study. Participant variables that are considered to be important are those that can influence the results of the study to be conducted.

2 An experiment was conducted to investigate whether alcohol consumption increases errors when driving. Volunteer participants were given a drink that they were led to believe contained alcohol. It looked and tasted like an alcoholic drink but did not contain any alcohol. The participants were then given a test in a driving simulator with automatic transmission and the number of driving errors was recorded. The next day, at the same time, participants were given an alcoholic drink that looked and tasted like the non-alcoholic drink. After allowing sufficient time for the alcohol to take effect, the participants were given a test in a driving simulator with manual transmission and the number of driving errors was recorded. The results showed that more driving errors were made after having the alcoholic drink.

For example, in a study on how friendships form among adolescents, personal characteristics of participants such as their sex, age, type of school attended, family background and cultural background could be assumed to be important.

When a researcher selects a sample that represents its population, the sample is called a representative sample. A representative sample is a sample that is approximately the same as the population from which it is drawn in every important participant variable.

Larger samples also minimise the likelihood of a freak ‘sampling error’ resulting in a sample which does not represent its population well and would therefore make it difficult to apply the results to that population. Some researchers have described the law of large numbers in relation to sampling. The law of large numbers states that as sample size increases, the attributes (characteristics) of the sample more closely reflect the attributes of the population from which the sample was drawn. Basically, the more people who are selected, the more likely it is that they will reflect and therefore be representative of the population.

There are different ways of obtaining a sample. The most common sampling procedures are called random sampling, stratified sampling and convenience sampling. Convenience sampling is the simplest method but is the least likely to achieve a representative sample.

**Summary**

**LEARNING ACTIVITY 2.7**

**Summarising potential extraneous and confounding variables**

Complete the following table.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Example</th>
<th>Why a potential extraneous or confounding variable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual participant differences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-standardised instructions and procedures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimenter effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo effect</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**LEARNING ACTIVITY 2.8**

**Review questions**

Identify the IV(s), DV(s) and a potential extraneous or confounding variable in each of the following experiments. Explain your answers.

1. An experiment was conducted to investigate whether young adults performed better on maths problems when working alone or when working in small groups. Two groups of participants were used, ensuring an equal spread of mathematical ability and personal characteristics across both groups. Because of a shortage of rooms, the participants working alone completed the problems in a small tutorial room with no windows in the corner of the school library. The participants working in small groups completed the problems in a large classroom with big windows on the first floor of the building (above the library). As hypothesised, the groups performed better than the individuals.

2. An experiment was conducted to investigate whether alcohol consumption increases errors when driving. Volunteer participants were given a drink that they were led to believe contained alcohol. It looked and tasted like an alcoholic drink but did not contain any alcohol. The participants were then given a test in a driving simulator with automatic transmission and the number of driving errors was recorded. The next day, at the same time, participants were given an alcoholic drink that looked and tasted like the non-alcoholic drink. After allowing sufficient time for the alcohol to take effect, the participants were given a test in a driving simulator with manual transmission and the number of driving errors was recorded. The results showed that more driving errors were made after having the alcoholic drink.
Random sampling

The dictionary definition of the term ‘random’ is something which is haphazard, unpredictable or ‘hit-or-miss’. However, when the term random is used by researchers in relation to a sample, it has the opposite meaning. Random actually means using a planned, systematic procedure to obtain a sample.

Random sampling is a sampling procedure that ensures every member of the population of research interest has an equal chance of being selected to be part of the sample. This can be achieved in a number of different ways.

One way is to obtain a complete list of all the people in the population. This list is commonly called a sampling frame. For example, an electoral roll may be used as a sampling frame, or the telephone numbers of all the people in a relevant database may be used. If you were conducting a research study in your school, class rolls could be used, but only those with the names of students in the target population — the population of interest.

After the sampling frame is obtained, the researcher could obtain a random sample using a simple lottery procedure to select the required number of names. The lottery procedure could involve drawing names out of a box or tossing a coin. For example, if a sheet of paper had all the names of the people in the population on it, the sheet would be cut up into slips of paper equal in size, with one name on each slip of paper. The names would then be thoroughly mixed in the box to help ensure their distribution throughout the box. Then, names of sample members (or research participants) could be drawn out ‘blindly’, one at a time. As a result of this procedure, the likelihood that the sample is representative of the population is increased, and so is the ability of the researcher to generalise the results to the population.

When a large number of participants are required, researchers often use a digitally generated list of random numbers. Each participant in the sampling frame is given a number from 1 through to however many are in the population of interest. If the first number in the digitally generated list of random numbers is 22, then the twenty-second person in the sampling frame is included in the sample; if the second number in the digitally generated list is 93, then the ninety-third person in the sampling frame is selected, and so on until the required number of participants have been selected.

Sometimes a researcher may not find it necessary or even desirable to use a random sample that is fairly representative of the population of interest. For example, a researcher interested in the language development of children may intentionally undertake a case study of a child raised in a harsh, deprived environment where there is little or no opportunity to learn language, rather than studying a sample of ‘average’ children from a ‘normal’ home environment.

The most important advantage of random sampling is that it helps ensure a highly representative sample. The larger the sample, the more likely it is that this will occur, but there is no guarantee that the sample will be representative.
The main limitation of random sampling is that it can only be carried out if a complete list of the target population is available. If available, it may be difficult to gain access. If accessed, the process of random selection may be time-consuming.

**Stratified sampling**

In some research studies it is important to ensure that particular groups in a population of interest are represented in their known proportions in that population. For example, suppose that a researcher wanted to study the attitudes of adult Australians to arsonists who deliberately light bushfires. They could reasonably expect that attitudes may differ depending on whether someone lives in the city or in a rural community. Consequently, the researcher would want to ensure that each of these groups was represented in the sample in about the same proportions that they were known to exist in the adult population. This can be achieved by using the sampling procedure called stratified sampling.

**Stratified sampling** involves dividing the population to be sampled into different subgroups, or strata, then selecting a separate sample from each subgroup (called stratum) in the same proportions as they occur in the population of interest. Socio-cultural factors such as residential area, type of accommodation, age, sex, income level, income type, educational qualifications and cultural background are examples of personal characteristics that may be used as the basis of dividing a population into strata.

The stratified sampling procedure is commonly used to study behaviour and mental processes that tend to vary greatly among different subgroups of a population. For example, suppose you were going to undertake a research study on attitudes of students in your school towards teachers’ use of rewards and punishments. You expect that attitudes may differ among students in different year levels so you want to ensure each year level (stratum) is proportionally represented in your sample.

You could first obtain separate lists of the students in each year level and then randomly sample from each list. If, for example, about 10% of all students in your school are enrolled in year 12 and about 15% in year 11, then your sample would consist of about 10% year 12 students and about 15% year 11 students. This would ensure students from each year level are represented in about the same proportions in the sample as they are in the population (the school). Using a stratified random sampling procedure would ensure that the sample is highly representative of the population and therefore not biased in a way you consider to be important.

Figure 2.22 shows an example of a stratified sample that could be obtained for the attitudes study. If everyone in a target population does not have an equal chance of being selected as a participant, then a biased sample will be obtained.

The most important advantage of stratified sampling is that it enables the researcher to sample specific groups (strata) within populations for comparison purposes; for example, males vs females, or people of different ages and cultural backgrounds who have been diagnosed as having schizophrenia and will be exposed to a new type of treatment program for the disorder. In addition, this means that there can be greater precision in the study and its findings when compared to the simple random sample taken from one larger group.

A major limitation of stratified sampling is that, like random sampling, it can be carried out only if complete lists of the target populations (strata) are available and accessible. However, if accessed, a representative sample cannot be obtained unless stratified random sampling is used. Either way, stratified sampling can be a very time-consuming procedure, more so than standard random sampling.

**Convenience sampling**

For some research studies it is not convenient, suitable or possible to obtain a representative sample. In such cases, a convenience sample (also called an opportunity sample) may be used and the researcher may use anyone who is available or present.
Convenience sampling, or *opportunity sampling*, involves selecting participants who are readily available without any attempt to make the sample representative of a population. For example, a representative sample of illegal drug users or homeless teenagers is not often readily available. Consequently, the researcher may go to locations known to be frequented by the required participants and simply select the first individuals they meet who are in the target population and who are willing and available to participate.

Similarly, a researcher conducting a study on drivers who do not obey red traffic lights at a particular intersection at a particular time would be using convenience sampling. Psychology students often use convenience sampling; for example, when selecting participants they have the opportunity to study other students in their school, children at a local primary school, friends, parents or relatives.

In most cases, convenience sampling produces a biased sample because only those people available at the time and location of the study will have a chance of being included in the sample. If a researcher used convenience sampling at a local shopping centre, they may select only those shoppers who look as if they will be cooperative to be in the sample and ignore those who appear uncooperative. Shoppers left out of the sample might think, feel or behave differently from those who are selected in the sample, yet these thoughts, feelings and behaviours will not be represented in the sample. Since a convenience sample is not representative of the target population under investigation, the data obtained can be misleading and the results of the study cannot be legitimately applied (generalised) to the entire population.

Despite these limitations, convenience sampling is widely used in psychology. It is quick, easy and inexpensive. Convenience sampling can also be of considerable value when conducting research to pilot, or 'test', procedures or to gain a preliminary indication of possible responses before conducting the actual study. Many researchers regard convenience sampling as an adequate sampling procedure when investigating aspects of mental processes or behaviour that are assumed to be similar in all 'normal' individuals, despite individual differences. For example, all 'normal' adults are capable of reflecting on their personal experiences and using language to communicate what they think or feel. Similarly, all normal adults are capable of seeing, hearing and responding reflexively.

**FIGURE 2.23** Convenience sampling involves selecting participants who are readily available without any attempt to make the sample representative of a population.
### LEARNING ACTIVITY 2.9

#### Summarising sampling procedures for participant selection

<table>
<thead>
<tr>
<th>Sampling procedure</th>
<th>Description</th>
<th>Example for a within-school investigation</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stratified sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Explain the difference between a sample and a population.
2. What does sampling involve?
3. Distinguish between a biased sample and a representative sample.
4. What are two potential limitations of small sample size?
5. Explain how the type of sampling procedure used can minimise extraneous and confounding variables.
6. Complete the table below.

---

**Participant allocation**

The method of selecting the sample is important in ensuring it is unbiased and representative of the population being studied. Equally important is the way in which research participants are placed in either the experimental or control group in an experiment.

In an ideal research world, everything about the experimental and control groups would be identical except for the IV. In reality, however, it is to be expected that there will be individual participant differences that may become extraneous or confounding variables and make it difficult to isolate the effects of the IV on the DV. Consequently, it is important to ensure that participant variables that might affect the results of the experiment are evenly spread in the experimental and control groups.

**Random allocation**

One way of minimising differences in the composition or make-up of the experimental and control groups is to randomly allocate participants to these groups. Random allocation, also called random assignment, is a procedure used to place participants in groups so that they are as likely to be in one group as the other. This means that every participant has an equal chance of being selected for any of the groups used. Participants selected for the experiment are just as likely to be in the experimental group as the control group.

As with random selection, random allocation can be achieved using a lottery procedure in which chance alone will determine the group to which each participant will be assigned. For example, drawing ‘names out of a box’ or tossing a coin are also appropriate ways of randomly allocating participants to groups.

With a sufficiently large number of participants, it is reasonable to assume that each group will end up with the same kind of spread of participant characteristics, abilities and backgrounds that may affect the DV and therefore the results. For example, consider the experiment on alcohol consumption and driving ability described previously. If the experimental group has a larger proportion of ‘bad’ drivers than the control group and the experimental group makes significantly more driving errors in the driving simulator, it will be difficult for the researcher to isolate the effect of alcohol (the IV) on driving ability (the DV).

The problem is that the participants in the experimental group may make more driving errors than the control group even when not under the influence of alcohol. Through random allocation of participants to the experimental and control groups, each group would be expected to end up with relatively even numbers of participants who are ‘good’ and ‘bad’ drivers.
The purpose of random allocation of participants is to obtain groups that are as alike as possible in terms of participant variables before introducing the IV. With random allocation of participants to the experimental and control groups, researchers can more confidently conclude that if two groups responded differently in the experiment in terms of the number of driving errors, then it most likely had something to do with the effect of the IV. Consequently, random allocation is a very important means of experimental control.

For a classroom experiment, placing all males in one group and all females in the other group would not be a random allocation procedure. Similarly, assigning the people seated in the front half of the room to one group and the people seated in the back half to the other group is not random allocation. There could be a difference in one or more personal characteristics of participants who prefer to sit at the front or back of the classroom.

Random allocation is different from random sampling. Random allocation is used to place participants in groups whereas random sampling is one of the methods that can be used to select participants for an experiment. Random sampling, however, is based on the same principle of ‘equal opportunity for all participants’.

![Diagram](image)

**FIGURE 2.24** A simple experimental design using random sampling to select participants and random allocation to assign them to either condition.

**LEARNING ACTIVITY 2.11**

**Review questions**

1. What are the random allocation procedures?
2. What does random allocation achieve in relation to groups selected for an experiment?
3. Why is random allocation considered to be a crucial feature of good experimental design?
4. Give an example of a random allocation procedure that could be used within a class experiment at school.

**Counterbalancing**

A counterbalancing procedure is commonly used to minimise order effects such as practice and carry-over. Counterbalancing involves systematically changing the order of treatments or tasks for participants in a ‘balanced’ way to ‘counter’ the unwanted effects on performance of any one order. By counterbalancing, the researcher recognises that an order effect is a potential extraneous or confounding variable and cannot be controlled or eliminated any other way.

There are different types of counterbalancing procedures. The most commonly used is called *between-participants counterbalancing*. This involves alternating the order in which the groups of participants are exposed to the experimental conditions. Each group of participants receives the treatments in a different order.

For example, if there were 20 participants in the alcohol and driving experiment, counterbalancing could require half the participants to undertake the driver test in the no alcohol condition first, followed by the test in the alcohol condition. The other half of the participants would undertake the driver test in the alcohol condition, followed by the no alcohol condition. Participants would also be randomly allocated to experience one condition or the other first or second.

![Diagram](image)

**FIGURE 2.25** A simple counterbalancing procedure for the alcohol and driving experiment.
The results for all participants are then combined across the entire experiment to achieve counterbalancing. In this way, any order effect that impacts on performance, as measured by the DV, is controlled.

**Single- and double-blind procedures**
Participants' expectations can influence the results of any investigation, in an experiment, so it is important that participants do not know whether they are in an experimental or a control group. In this case, the experiment is said to be using a single-blind procedure. It is called a single-blind procedure because the participants are not aware of (are 'blind' to) the condition of the experiment to which they have been allocated and therefore the experimental treatment (the IV).

To control possible experimenter effects, while also controlling participant expectations, researchers may use a procedure in which neither the participant nor the researcher interacting with the participants knows which participants are in the experimental or control groups. This is called the double-blind procedure because the participants and the researcher (or research assistant) directly involved with the participants are unaware of (are 'blind' to) the conditions to which the participants have been allocated. Only a researcher who is removed from the actual research situation knows which participants are in which condition (or groups).

The double-blind procedure has obvious value in experiments in which knowledge of the conditions might influence the expectations or behaviour of the researcher as well as the participants.

**Placebos**
In an experiment, participants in the experimental group are exposed to the IV 'treatment' and participants in the control group are not. Because only the experimental group receives the treatment, they may be influenced by their expectations about how they should behave. Therefore, there is a potential confounding variable — the experimental group may respond differently to the control group either because of the IV or because of their expectations of how they should behave.

In order to minimise the impact of this variable on the DV, the control group can be given a placebo — a fake treatment that is like the IV treatment used in the experimental group but which is actually neutral or has no known effect. In this way, control group participants should form the same expectations as the experimental group, thereby controlling the effects of this unwanted variable.

For example, consider an experiment to test a new herbal drink called Attendo Memoro that claims to improve concentration and memory. Participants in the experimental group could be given a daily drink of Attendo Memoro over a 6-week period and control group participants could be given a drink that looks, smells and tastes like Attendo Memoro but has no active ingredient. At the beginning and end of the study all participants could be given concentration and memory tests. Because a placebo was used in the control group, any difference in the results could not be said to be due to participant expectancy effects.

Similarly, when testing drugs (or new medical therapies), researchers give placebo pills or injections to the control group so that all participants experience the same procedure and form the same expectations. And in studies that require the experimental group to perform, for example, a physically or mentally demanding task prior to making a response, the researcher would have the control group perform a similar placebo task to eliminate differences between the groups in terms of motivation or fatigue (Heiman, 2002).

When a placebo is given to a control group, the group is sometimes referred to as the placebo control group or the placebo condition.
Standardised instructions and procedures

The instructions and procedures used by the researcher are a source of extraneous or confounding variables so their potential unwanted influences must be minimised. This is achieved by standardisation (‘consistency’) across the different conditions. Using standardised instructions and procedures means that instructions and procedures are the same for all participants (except for variations required for experimental group participants exposed to the IV).

The use of standardised instructions ensures that the directions and explanations given to all participants in each condition are identical in terms of what they state and how they are given. To help ensure this ‘sameness’, researchers usually read from a pre-prepared script in a ‘neutral’ voice. The script typically contains all the information about what the researcher says and does throughout the entire experiment.

It is also essential that all participants experience the same environment and procedures, with the only exception being exposure to the independent variable. Therefore, standardised procedures must be used — the techniques used for observing and measuring responses should be identical for all individual participants. All participants should be treated in the same way, as appropriate to the condition to which they have been assigned. For example, using standardised procedures, all participants would:

- participate in the experiment at the same time of day
- have the same amount of time
- learn the same amount of information
- complete the same activities (except for variations required for IV exposure).

How the researcher presents stimuli, obtains responses and records scores during DV measurement can be controlled through automation by using electronic or mechanical devices to present stimuli and to measure and record responses. Electronic timers, data projectors, video and audio recorders, and computers ensure controlled and consistent stimulus presentations (such as an image on a monitor to measure reaction time). Automating data collection ensures that the scoring system is consistently and accurately applied and provides for sensitive measurement of responses (such as the keystroke used to measure reaction time to a stimulus).

The use of standardised instructions and procedures minimises unwanted participant variables (including the placebo effect) because all participants have the same experience. It can also help control experimenter effects, as all the researchers involved will follow the same procedures. Consequently, when the results for experimental and control groups are compared, significant differences can be said to be due to the IV with confidence.

FIGURE 2.28 Automation ensures recording of participant’s responses is standardised across different experimental conditions.
Use of an appropriate experimental research design

Various experimental designs can be used to minimise the effects of potential extraneous and confounding variables. Three of these designs are the independent groups, repeated measures and matched participants designs.

**Independent groups**

In an experiment with an independent groups design, also called between participants, each participant is randomly allocated to one of two (or more) entirely separate ('independent') conditions ('groups').

The simplest independent groups design uses two groups — most often one group as the experimental group and the other as the control group. Many examples of experiments with this design have been given throughout this chapter.

Random allocation is an essential feature of the independent groups design in order to minimise individual participant differences. Random allocation to the different conditions will help ensure groups are well matched on participant variables and therefore fairly equivalent. The bigger the groups, the more likely it is that a uniform spread of characteristics and abilities will be achieved. Although random allocation does not guarantee that different conditions are entirely equivalent in the spread of participant variables, it does greatly reduce the likelihood of differences so that the effect(s) of the IV on the DV can be isolated.

The independent groups design is very common in experimental research. An advantage is that, unlike the repeated measures design, there is not often a need to spread out the time period between the different conditions. This means that the experiment can usually be completed on one occasion, which also helps ensure participant attrition ('dropout' rate) is negligible. There are also no order effects between conditions to control. However, there is less control over participant variables than in the repeated measures and matched participants designs, especially when a small sample is used.

**FIGURE 2.29** In an independent groups experiment, participants are randomly allocated to either the experimental or control group.
Repeated measures

In a **repeated measures** design, also called within **participants**, each participant is in both the experimental and control conditions. This means that the same participants are in both the experimental and control groups. The groups are therefore identical so individual participant differences are controlled.

For example, suppose a researcher is interested in investigating whether memory of a crime is better when hypnotised. Using the repeated measures design, all participants may be shown a video of two robberies that are reasonably alike in detail, once when hypnotised and once when not. For each condition they could be given a test of recall with 20 questions about the robbery. The number of correct answers given under each condition could then be compared.

This design would give the researcher strict control over all the possible participant variables that could influence memory, such as individual differences in attention, level of alertness, motivation, mood, visual perception, prior experience, and so on. Participant differences that may not have been identified by the researcher as potential extraneous or confounding variables will also be controlled because the participants in both conditions are identical in every respect.

When using repeated measures, the researcher has to consider order effects because they are more likely to arise for this type of design. Participants may perform better in the second condition because they have practised a task or gained other useful knowledge about the task or the experiment. Alternatively, participants’ performance may be impaired by an order effect such as boredom or fatigue, and they may not perform as well on the second occasion. In either case, order effect is a potential confounding variable because the researcher cannot be confident about whether the IV or order effect caused the change in the DV.

One way of dealing with order effects such as practice, fatigue, and boredom is to increase the time between measuring the DV in each condition. For example, participants might be in the experimental condition one day, then return a week later for the control condition. If this procedure is inappropriate, inconvenient or impractical, the researcher can use a counterbalancing procedure.

For example, as shown in table 2.3, half the participants would follow one order (view video when hypnotised in the experimental condition first, then when not hypnotised in the control condition). The other half would follow the reverse order (view video in the control condition first, then in the experimental condition).

<table>
<thead>
<tr>
<th>Participant</th>
<th>Experimental condition</th>
<th>Control condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>2</td>
<td>Second</td>
<td>First</td>
</tr>
<tr>
<td>3</td>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>4</td>
<td>Second</td>
<td>First</td>
</tr>
<tr>
<td>5</td>
<td>Second</td>
<td>First</td>
</tr>
<tr>
<td>6</td>
<td>First</td>
<td>Second</td>
</tr>
</tbody>
</table>

The main advantage of the repeated measures design is that any difference in performance on the DV in each condition of the experiments is unlikely to be due to individual participant differences because each participant is in every condition. This design also tends to require a relatively smaller number of participants when compared with other designs because the same participants are in all conditions.

**FIGURE 2.30** In a repeated measures experiment, the same participants are in both the experimental and control groups.
However, the repeated measures design also has limitations. Although this design keeps individual participant differences constant, it does not necessarily control all participant variables that can influence the results. For example, participants may guess what the experiment is about as they compare the two conditions, creating expectations and beliefs that lead to unnatural responses. Experiments using the repeated measures design also have greater participant attrition (loss) rates, especially when the experiment is conducted over several days to reduce participant fatigue, boredom or overload. Order effects are more likely to occur with this design, but counterbalancing can be used for control.

**Matched participants**

In a matched participants design, also called matched groups, each participant in one condition ‘matches’ a participant in the other condition(s) on one or more participant variables of relevance. For example, in the memory and hypnosis experiment, memory would be the most relevant variable.

To create two matched groups, one for the experimental condition and one for the control condition, all participants could be pre-tested on memory ability to obtain recall scores. Then, each participant would be paired with someone else with a similar score until all the participants had been matched on recall.

The participants would then be randomly allocated to a group on the basis of their test scores. For example, the two participants with the highest scores would be randomly allocated to the hypnosis and no hypnosis groups respectively. Then the two participants with the next highest scores would be allocated to the two groups, and so on. In this way, the two groups in the experiment would be matched in terms of memory, thereby controlling this potential extraneous or confounding variable.

Participants can be matched on more than one characteristic when using this experimental design. Matching for age, sex and mental abilities is quite common. Randomly allocating one member of each matched pair to a different group (condition) ensures that each group is fairly equivalent in terms of the spread of participant variables that can cause a change in the DV.

The advantage of matching participants is that it ensures that in every condition there is a participant with very similar or identical scores on the variable(s) the researcher seeks to control. This means that these variables are the same across the conditions, thereby eliminating them as potential extraneous or confounding variables. Participant attrition is less common than with the repeated measures design and there is not often a need to spread out the time period between the different conditions.

There are, however, limitations to the matched participants design. One potential problem is that it is often difficult and time consuming to actually recruit participants who are sufficiently alike in participant variables of research interest. There are also other practical problems. For example, to find matching participants, the researcher often has to pre-test many individuals and/or settle for a very small number of participants. This can be very time consuming, especially if there are more than two groups in the experiment. Pre-testing can also create order effects. And the loss of one participant through attrition means the loss of a whole pair, triplet and so on. Pre-testing is not required for the repeated measures design and is used in an independent groups design only if the researcher elects to do so.

Use of the matched participants design is often not necessary in experimental research. Random allocation is usually sufficient to control individual differences of participants as it ensures equivalence of the experimental and control groups. Consequently, the matched participants design is not used as often as the other designs.

**FIGURE 2.31** In a matched participants experiment, each participant in one condition ‘matches’ a participant in the other condition(s) on one or more participant variables of relevance.
Advantages and limitations of experimental research

A key feature of an experiment is the researcher's attempts to control the conditions in which a behaviour of interest or other event occurs, whether the experiment occurs in a laboratory setting or in a real-life, field setting. As well as controlling the IV, the researcher also attempts to minimise or eliminate the influence of unwanted extraneous variables to concentrate entirely on the effect the IV has on the DV. Elimination of all extraneous variables is not always possible, but control is usually greater than in other research methods, especially if the experiment is conducted in a laboratory setting. Consequently, the experiment has several advantages when compared to other research methods.

One advantage of the experiment is that the IV can be manipulated in order to observe the effect on the DV, therefore making it possible to test if there is a cause and effect relationship between the IV and DV. Furthermore, because controlled conditions are known conditions, the experimenter can set up the experiment a second time and repeat it to test (or 'check') the results.

Alternatively, the experimenter can report the conditions of an experiment in such a precise way that others can replicate the experiment and test the results. Replication is very important because when a study is repeated and similar results are obtained, there can be greater confidence in the reliability and validity of the results obtained.

Despite its precision, there are several limitations of the experiment. Although a field experiment occurs in a real-life setting and therefore has a relationship to the real world, it is often difficult to strictly control all variables because of the unpredictability of real-life settings. The ability to more strictly control variables is an advantage of the laboratory setting; however, it is often artificial and too dissimilar to real life. For example, bringing someone into the unfamiliar environment of a psychology laboratory can change their behaviour to the point where it is not appropriate to generalise or apply the observed behaviour to situations outside the laboratory.

Furthermore, some things cannot be measured in a laboratory. The researcher cannot break up families, for example, to measure the effects of family separation. Nor would the laboratory be the best setting for testing variables such as grief, hate or love. It may be difficult for participants to express these emotions naturally or very realistically in a laboratory setting.

### LEARNING ACTIVITY 2.13

#### Summarising three experimental designs

1. Complete the following table.

<table>
<thead>
<tr>
<th>Experimental design</th>
<th>Description</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeated measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matched participants</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. For each of the following potential extraneous or confounding variables, rate each of the three experimental designs from 0 to 10 to indicate how well it controls the variable, as compared to the other designs. A score of 0 indicates no control and a score of 10 indicates perfect control. Explain your choice of ratings.

   (a) Individual participant differences
   (b) Order effects
   (c) Experimenter effect
   (d) Placebo effect

3. 
   (a) Formulate a definition of experimental research.
   (b) Briefly describe two advantages and two limitations of experimental research.

#### LEARNING ACTIVITY 2.14

#### Identifying the experimental design

Name the type of experimental design used in each of the following.

1. To compare the effectiveness of different psychotherapies for treating spider phobia, participants with a spider phobia are allocated to one of two treatment conditions or a control condition.
2. To study the effects of an anxiety reduction medication, participants diagnosed with a phobia are tested before and after they are given the medication.
3. To compare the effects of inspirational message types A and B, participants listen to message A for one week then complete an assessment on their personal wellbeing. The next day they start listening to message B for two weeks after which they complete the wellbeing assessment.
4. To study whether meditation reduces stress, participants' blood pressure is measured before and after a period of meditation.
5. A study on whether males and females are persuaded differently by a female car salesperson.
6. For study 5, a requirement that for each male of a particular age there is a female of that age.
**BOX 2.1**

**Correlational research**

Sometimes, an experiment is impractical or inappropriate to use. For example, suppose a researcher wanted to find out how a severe psychological trauma in childhood affects school performance. It would be unethical to set up two similar groups of participants and expose one of these groups to some kind of traumatic event that would trigger a severe emotional reaction so that its effects on a measure of school performance could be observed. In such cases, researchers tend to rely on existing information to assess the ‘co-relationship’, or correlation, between the variables of interest.

Correlational research is used to study the type and strength of relationship between two or more variables. Unlike experimental research, there is no attempt to manipulate any variable. The researcher simply assesses the relationship between the variables of interest. This is usually done by applying a statistical technique to data that have been collected on each variable.

For example, to study the relationship between air temperature and occurrence of violent crimes, the researcher could obtain existing data on both the daily air temperature (such as maximum and minimum temperatures) over a period of time and violent crimes committed over this period of time, then determine the number of violent crimes committed on very hot days.

**FIGURE 2.32** A researcher would access existing data to study whether there is a correlation between air temperature and violent crime.

The term correlation is used to identify and describe how variables are ‘co-related’. It does not indicate whether one variable (such as air temperature) causes another (such as violent crimes). Rather, it indicates the direction of the relationship and the strength of the relationship.

**Direction of correlation**

For any two variables which are assessed in a correlational study, there are three possible relationships between them: positive, negative and zero (no relationship).

A **positive correlation** means that two variables vary (‘change’) in the same direction — as one variable increases, the other variable tends to increase (and vice versa). For example, as age increases, vocabulary tends to increase (and as vocabulary increases, age tends to increase, or the lower the age, the smaller the vocabulary).

A **negative correlation** means that two variables vary in opposite directions — as one variable increases, the other variable tends to decrease (and vice versa). A negative correlation is like a seesaw. For example, as self-esteem increases, sadness tends to decrease (and as sadness increases, self-esteem tends to decrease).

A **zero correlation** means that there is little or no relationship between two variables. For example, there is no relationship between intelligence and hair colour. These two variables can change entirely independently of each other.

A correlation is usually described by a number known as a correlation coefficient. This is expressed as a decimal number which can range from +1.00 to −1.00. The plus or minus sign describes the direction of the relationship between the two variables; that is, positive or negative.

A correlation coefficient with a plus sign indicates a positive correlation. This means that high scores for one variable tend to go with high scores on the other, middle scores with middle scores, and low scores with low. For example, consider the results of a correlational study on age and problem-solving ability. If there is a high positive correlation (say +0.75) between age and problem-solving ability, then older people tend to be good problem-solvers (they solved many problems in a 20-minute period) and younger people tend to be poor problem-solvers (they solved fewer problems in a 20-minute period).

A correlation coefficient preceded by a minus sign indicates a negative correlation. This means that when a score on one variable is high, the score on the other tends to be low, and middle scores tend to go with middle scores. For example, if there is a high negative correlation (say −0.75) between age and problem-solving ability, then older people would tend to be poor problem-solvers and younger people would tend to be good problem-solvers.

When reporting correlation coefficients for positive correlations, researchers usually omit the plus sign from the front of the score. However, the minus sign is always included for a negative correlation.

**Strength of correlation**

The decimal number of the correlation coefficient describes the strength of the relationship between the sets of scores for two variables; that is, whether the relationship is strong, moderate or weak. A correlation coefficient which is close to +1.00 indicates a very strong
positive correlation between two variables. A correlation coefficient which is close to –1.00 indicates a very strong negative correlation between two variables.

Correlation coefficients of 1.00 and –1.00 indicate perfect correlations but these rarely occur in psychology. A correlation coefficient which is close to 0.00 indicates little or no relationship between two variables. For example, 0.13 and –0.13 would be considered a weak positive and weak negative correlation respectively.

**Correlation and causation**

Correlations show the existence and extent of relationships between variables but they do not necessarily indicate that one variable causes the other. For example, people get older as the world rotates on its axis. There is an extremely strong correlation between these two factors, but it would be incorrect to assume that the Earth’s rotation causes people to age or that people’s ageing causes the Earth to rotate.

There are also many instances when high correlations suggest a logical cause–effect relationship, and sometimes correlations really do represent causal relationships. The number of friends a person has may be closely related to how happy they are. But a very high correlation doesn’t necessarily mean that there is a cause–effect relationship because both variables may be correlated with a third variable.

For example, there is a high positive correlation between the number of permanent teeth in children and their ability to answer increasingly difficult questions on intelligence tests. It cannot be assumed, however, that having more teeth causes increased cognitive ability. The correlation is high because a third variable — increasing age — accounts for both new teeth and cognitive development.

When two variables are strongly correlated, this is not accepted by psychologists as evidence of causation in the absence of other research evidence. In such cases, researchers may test the possible cause–effect relationship by conducting a controlled experiment.

---

**Cross-sectional studies**

Cross-sectional studies are most commonly used in the study of human development, primarily to study how one or more aspects of our development changes over time and/or factors influencing change.

A **cross-sectional study** selects and compares groups of participants on one or more variables of interest at a single point in time. It is most commonly used to study age-related differences. For example, to study the use of rules in games played by children, groups of children representing each age group from three to seven years inclusive can be selected and observed at about the same time. Or, to study age differences in how much information can be held in short-term memory, groups of people selected at ten-year intervals from 10 to 80 year olds could be tested and the results compared.

It is also used to study differences between groups on other variables at a specific time; for example, samples may be selected on the basis of one or more mental abilities, personality type, family environment, mental health characteristics, sleep habits, cholesterol levels and so on.

A cross-sectional study uses an independent groups design and is sometimes called a **quasi-experiment**.

However, it is not a true experiment because participants cannot be randomly assigned to experimental and control groups. Instead, a cross-sectional study uses naturally formed or occurring groups. For example, in a cross-sectional study investigating age-related differences, the researcher can select participants from different age groups of interest but cannot randomly assign people to be a particular age. In addition, the researcher observes characteristics or events that already exist or occur naturally in a sample (or population), without manipulating any variables.

**Advantages and limitations of cross-sectional studies**

The major advantage of a cross-sectional study is that it can be conducted relatively quickly. Compared to other research methods, it tends to be easy to undertake, not too time-consuming and less expensive. For example, a researcher can study differences in one or more variables of interest in 5-, 10- and 15-year-olds at one time over a short period, instead of tracking them over 10 years to complete their investigation. In this way, a snapshot of age-related differences can be obtained without having to wait many years for the results.

A major limitation of cross-sectional studies is that a cause–effect relationship between variables cannot be tested or determined. In addition, when age differences are studied, variables other than age can influence the results and therefore be confounding variables.

Differences found between age groups may be due to factors other than age such as the particular backgrounds and life experiences of participants in each age group. For example, genetic make-up, number of siblings, family environment and schooling can cause differences in a cross-sectional study of language development in young children.

In particular, one participant variable that cannot be controlled in a cross-sectional study is called a **cohort effect**. This shows up when the researcher measures characteristics in groups of people who were born at significantly different times and members of each group share life experiences associated with the period in which they grew up or a particular event or situation to which they were all exposed. For example, people who are currently in their nineties experienced childhood during the 1930s depression. They may behave differently from 50- or 60-year-olds not because
of chronological age differences, but because of that particular life experience. Similarly, consider also the fact that you have grown up during a period marked by widespread access to the internet and use of new digital technologies and social media, which is a different experience from that of your parents and grandparents when they were growing up.

The larger the differences in age between different groups in a cross-sectional study, the greater the potential for a cohort effect — when age differences may be entangled with (confounded by) differences in participants’ life experiences associated with being born at a particular time and growing up as a member of a particular generation.

**BOX 2.2**

**Longitudinal studies**

A *longitudinal study* is a long-term investigation that tracks the same group (or groups) of people over an extended period of time, observing changes that occur in behaviour and/or mental processes at several points in time. Some longitudinal studies are relatively brief, lasting for one to two years; others can last a lifetime.

Usually, the same group(s) of participants is studied and re-studied at regular intervals. For example, *Growing up in Australia: the Longitudinal Study of Australian Children* is being conducted by the Australian Institute of Family Studies in conjunction with other organisations (2015).

The study commenced in 2004 and follows the development of 10 000 children and families from all parts of Australia. There are two groups with about 5000 in each — families with 4–5 year old children and families with infants aged 0–1 years.

The study is investigating the contribution of children’s social, economic and cultural environments to their adjustment and wellbeing. Parents, child care providers, teachers and the children themselves provide information. Families are visited for a face-to-face interview every two years. Various aspects of the children’s development are also measured, including their physical development, emotional wellbeing, and intellectual and social development.

The longitudinal method is particularly useful in the study of lifespan development. These studies provide information to help psychologists understand long-term changes in thoughts, feelings and behaviour.

**FIGURE 2.33** In the longitudinal study, the *same* participants are tested at different points in time over a number of years (e.g., 2000, 2005, 2010, 2015, 2020). In a cross-sectional study investigating age-related differences, *different* participants in different age groups are tested at the same time (e.g. 2020).

<table>
<thead>
<tr>
<th>Participant’s age</th>
<th>5 years</th>
<th>10 years</th>
<th>15 years</th>
<th>20 years</th>
<th>25 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of testing</td>
<td>2000</td>
<td>2005</td>
<td>2010</td>
<td>2015</td>
<td>2020</td>
</tr>
</tbody>
</table>

**LEARNING ACTIVITY 2.15**

**Review questions**

1. Explain what a cross-sectional study is, with reference to an example.
2. What is an advantage of the cross-sectional study (other than efficiency and expediency) when compared to other research methods?
3. (a) A cross-sectional study is sometimes referred to as a quasi-experiment. What does this mean? (b) Explain why a cross-sectional study cannot be used to investigate a causal relationship between age and another variable of interest.
4. (a) Suggest a definition for a cohort. (b) What is a cohort effect?

**Case studies**

Sometimes a researcher will collect detailed information on only a small number of people, perhaps an individual or a small group of two or three. When this is done, the research method used is likely to be a case study.

A *case study* is an intensive, in-depth investigation of some behaviour or event of interest in an individual, group, organisation or situation. For example, many of the early language researchers started out by keeping detailed diaries on the language development of only a few children. Social psychologists have also learnt about behaviour in small friendship groups by conducting case studies in which they observe and record social interactions within the same group of people in different situations over a period of time. An assumption is that patterns of behaviour observed...
within the group may apply to other friendship groups made up of people of similar ages and backgrounds. Such case studies can also suggest hypotheses that could be tested using other research methods.

Case studies are often used when large numbers of participants are not available for study; for example, to study individuals with a rare or unusual disorder or ability. The case study may involve a combination of data collection methods. For example, an individual may be interviewed at length. Information may also be collected through interviews of family members, friends and teachers or co-workers. The individual’s medical records and school reports may also be considered. Other sources of information can include extensive psychological testing and observations of the person’s behaviour.

Much of what was first known about the role of the brain in behaviour and mental processes has come from case studies. When used in a clinical setting for therapeutic (or ‘treatment’) purposes, a case study is often referred to as a case history or a clinical observation.

One early case study conducted by a neuropsychologist involved a rare disorder called face agnosia (Bodamer, 1947). Face agnosia is an impaired ability to recognise faces. People with this disorder have difficulty in recognising the faces of family and friends or famous personalities.

The case study involved a young female who was also unable to recognise her own face. She was referred to as S.T. in the case study report to protect her identity. Whenever S.T. looked in the mirror, she saw a reflection of a stranger. However, S.T. knew that she was the strange looking person because she was the only person in front of the mirror. In one series of tests, S.T. was asked to speak in front of the mirror and make gestures such as a nod or a shrug. S.T. often recognised her own voice and occasionally recognised gestures, but her face was always completely new to her. S.T. also had difficulty recognising animal faces. For example, she described a dog’s face as ‘a human face with funny hair’.

Although unable to recognise faces, S.T. knew what a face was and could recognise and name everyday objects such as furnishings, articles of clothing, trees, cars and so on without difficulty. This suggested that the area of the brain involved in facial recognition was different from that involved in recognising objects. Furthermore, the different brain areas probably interacted with language and memory in different ways.

When neuroimaging devices such as PET and fMRI scans became available, neuropsychologists were able to conduct experiments with individuals suffering from face agnosia. Such experiments over the past 30 years confirmed the conclusions of early case studies. They have also enabled researchers to pinpoint brain areas and structures that interact in facial recognition, object recognition, memory and language.

Advantages and limitations of case studies
Case studies provide a useful way of obtaining detailed information on behaviour and mental processes. There is usually no manipulation or control of variables. Consequently, case studies can avoid artificiality and provide a ‘snapshot’ of the actual experience of one or more individuals at a particular time in a particular situation. They can also provide insights into how others may think, feel or behave under similar circumstances.

Another advantage of case studies is that they can be a valuable source of hypotheses for further research. However, case studies cannot be used to actually test hypotheses unless combined with the results of other case studies of similar participants or another research method that is suitable for testing hypotheses.

A major limitation of case studies is their sample size. They are commonly based on the experiences of one or more individuals at a particular time in a particular situation.
of only one individual or a very limited number of individuals. The very small sample means that case studies can usually provide only weak support for drawing scientific conclusions. Furthermore, generalising or applying the results to others in a relevant population cannot be done with any certainty whatsoever. Generalising is a bigger problem when the case study involves someone with a rare or unusual disorder or ability.

Case studies also have the limitation of being susceptible to biased information from the participant or the researcher. This can influence the accuracy of the information that is obtained and conclusions that may be drawn. For example, case studies usually rely on the individuals under investigation to provide a great deal of the required information. Some participants may not remember clearly what they actually experienced, or they may intentionally change or omit information that they do not wish to reveal for personal reasons.

Similarly, case studies are usually conducted by one researcher. It is possible that the researcher sees or hears what they expect or hope to see or hear. Furthermore, the researcher is also responsible for deciding what to include in their descriptions and what to leave out. In writing up the case, the researcher may select information that supports key points or conclusions they wish to make, and omit other points that may be just as relevant and could have been included by another researcher interpreting the same information.

**LEARNING ACTIVITY 2.16**

**Review questions**

1. What is a case study?
2. Give three examples of research findings or theories that have been derived from case studies in psychology.
3. Could an experiment be conducted as part of a case study? Explain your answer.
4. Describe two advantages and two limitations of case studies when used for research purposes.

**Observational studies**

In our everyday lives we observe the behaviour of other people and draw conclusions about them from their actions. For example, if we notice that someone is always quiet in class, prefers to sit by themself and blushes when asked a question, we might conclude that the person is shy, lacking in confidence or withdrawn.

Researchers in psychology, however, use observation in a more precise and planned way. For example, they precisely describe the behaviour to be observed, and do not jump to conclusions about attitudes, personality characteristics, motives or other factors that may underlie the observed behaviour.

An **observational study** involves collection of data by carefully watching and recording behaviour as it occurs. Psychologists use observational studies to collect data in research when the behaviour under investigation is clearly visible and can be easily recorded.

Data collection may be:

- **structured**: a prepared system is used to guide and record observations; for example, a checklist of items to look for
- **unstructured**: observations are made without a predetermined format
- **semi-structured**: a part of the observational study involves use of a predetermined format.

Most observational studies conducted in psychology are structured and use systematic data collection techniques, such as the checklist with predetermined criteria shown in figure 2.35. A structured study typically involves operationalising the behaviour of interest and variables that are involved. For example, a researcher might elect to observe ‘laughing’ as an indicator of a participant’s level of enjoyment of a television comedy program. To do this, the precise facial and bodily responses to be measured must be specified to distinguish between ‘laughing’ and other similar behaviours, such as ‘smiling’ or ‘chuckling’.

Sometimes an observational study might resemble an experiment. For example, to investigate roles and hierarchies (‘pecking orders’) in groups, a researcher might ask the members of a friendship group to discuss a controversial issue, then observe and record who starts the discussion, who changes the topic, who speaks, how often and for how long, and so on. This study could occur in a controlled laboratory...
setting, or in a field setting such as a place where the group normally meets and interacts; for example, the school canteen or an area of the school grounds. The researcher might also observe roles and hierarchies in a group comprising strangers to make comparisons with the friendship group.

Although a particular observational study might use an independent groups design and all experiments actually involve observation of responses, an observational study is not a true experiment unless random allocation is used. Furthermore, an observational study can reveal a relationship between two variables (e.g. group type and roles adopted by group members), but only a true experiment can establish a cause–effect relationship.

**Natural and contrived settings**

Observations may be conducted within a participant's natural environment or in a contrived environment. In both settings, researchers would wait for the behaviour of interest to occur voluntarily and to unfold as it usually does.

When observations are conducted within the participant's natural environment, the method is commonly called naturalistic observation. In naturalistic observation, the researcher views behaviour in the natural, 'real-life' environment where it would ordinarily occur. This is a situation where behaviour in its genuine form would be most likely to be observed. In addition, the researcher conducts their observations in an inconspicuous or 'unnoticeable' manner so that their presence does not influence the behaviour of interest.

For example, in a study on the development of social behaviour, a researcher might observe children at play in a preschool centre's outside area at lunchtime. They would do so from the 'sidelines' so that the children are not aware that they are being observed to help ensure their presence does not interfere with the naturally occurring, voluntary play behaviour. The researcher may observe that younger children tend to play alongside other children but not actually interact with them, whereas older children tend to interact more in their play with other children. On the basis of these observations, the researcher may assume that there are different types of play in which children may engage and that these types of play are age-related or age-dependent.

A contrived environment is one that the researcher creates for the specific purpose of conducting an observational study. It is an artificial ('non-naturalistic') environment for the behaviour of interest and is sometimes referred to as a structured or laboratory environment because of the degree of control the researcher has over it or where the observations are made. For example, the researcher conducting the study on social behaviour may decide to observe children at play in a room set up for that purpose at a venue outside the preschool centre. Specific playthings may be made available and strategically located together with a table and chairs. Observations could then be made from behind a one-way mirror so that the children are not aware that they are being observed. The children's behaviour might also be video recorded so that researchers can also record observations to help ensure reliability of the data.

In a study on the development of social behaviour, a researcher might observe children at play in a playgroup situation from behind a one-way mirror so that the children are not aware that they are being

<table>
<thead>
<tr>
<th>Coder name</th>
<th>Olive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episode</td>
<td>Coding categories</td>
</tr>
<tr>
<td></td>
<td>Proximity</td>
</tr>
<tr>
<td>Mother and baby play alone</td>
<td>1</td>
</tr>
<tr>
<td>Mother puts baby down</td>
<td>4</td>
</tr>
<tr>
<td>Stranger enters room</td>
<td>1</td>
</tr>
<tr>
<td>Mother leaves room, stranger plays with baby</td>
<td>1</td>
</tr>
<tr>
<td>Mother re-enters, greets and may comfort baby, then leaves again</td>
<td>4</td>
</tr>
<tr>
<td>Stranger tries to play with baby</td>
<td>1</td>
</tr>
<tr>
<td>Mother re-enters and picks up baby</td>
<td>6</td>
</tr>
</tbody>
</table>

The coding categories are:

- **Proximity**: The baby moves toward, grasps or climbs on the adult.
- **Maintaining contact**: The baby resists being put down by the adult by crying or trying to climb back up.
- **Resistance**: The baby pushes, hits or squirms to be put down from the adult’s arms.
- **Avoidance**: The baby turns away or moves away from the adult.

**FIGURE 2.35** This checklist was used to observe and record the behaviour of 12-month-old infants in an observational study of attachment behaviour (as indicated by responses to strangers). The infants were observed playing in a room with two adults — the infant’s mother and a stranger. The infants’ behaviour in different situations ('episodes') was rated ('coded') on a 7-point scale according to fear categories. A rating of 1 meant ‘The infant makes no effort to engage in the behaviour’ and a rating of 7 meant ‘The infant makes an extreme effort to engage in the behaviour’ (Ainsworth et al., 1978).
observed. From the observations of each child's interactions, the researcher might make assumptions or inferences about children's social behaviour.

For example, the researcher may observe that younger children tend to play alongside other children but not actually interact with them, whereas older children tend to interact more in their play with other children. On the basis of these observations, the researcher may assume that there are different types of play in which children may engage and that these types of play are age-related or age-dependent.

**Participant and non-participant observation**

Sometimes, psychologists engage in **participant observation**. They actually participate in the activity being observed and may deliberately try to be mistaken by the participants as being part of the group or situation being observed.

In one study that used participant observation, the researchers had themselves admitted to several different psychiatric hospitals by imitating the symptoms of a severe mental illness. After they had been admitted, they kept records of their observations while in the hospital. Their record-keeping behaviour was regarded by the hospital staff as being a symptom of their mental illness (Rosenhan, 1973).

When researchers try to conceal their presence while making observations, it is called **non-participant observation**. When observations of behaviour are made in a field setting, that is, the usual or real world surroundings in which the behaviour occurs, psychologists will often conceal their presence by watching from the 'sidelines'. For example, a researcher might sit on a nearby bench pretending to be absorbed in a book in order to observe people's reactions to a group of 'street kids'. In other situations, psychologists might use a hidden video camera to record events.

**Advantages and limitations of observational studies**

Each type of observational study is useful under different circumstances and has advantages and limitations depending on the specific procedures used, particularly the degree of structure in the data collection technique and the observational setting.

The main advantage of observational studies, especially naturalistic observation, is that researchers can watch and record behaviour as it usually occurs, without the need for any manipulation or intervention. When people are observed in this way, they are not influenced by perceptions that can form in artificial, contrived environments and lead them to behave differently from how they normally do. Sometimes, merely being present in an artificial environment can cause an unnatural change in behaviour. Thus, naturalistic observation often enables researchers to gain more accurate information about the typical behaviours of people (and animals), both immediately and over a longer period, than do other research methods. When compared to research methods that involve asking people about their behaviour, the researcher can observe what people actually do (or say), rather than what they say they do.

In addition, structured observations through use of checklists and specific criteria enhance the accuracy of data collection and therefore the results obtained. This is a more likely outcome when the observational setting is strictly controlled, as in a contrived laboratory-type situation.

Another advantage of naturalistic observational studies is that some types of human behaviour can only be studied as they naturally occur because it would be unethical or impractical to study them in a laboratory situation. For example, it would be unethical to severely deprive children in their early life in order to observe the effect of deprivation on behaviour in the future. Similarly, some behaviours cannot be realistically reproduced in a laboratory. A researcher cannot, for example, study crowd behaviour in a laboratory. Nor could a researcher expect to obtain valid information about how people usually behave when they are in love by bringing a pair of participants into a laboratory situation and asking them to 'be in love' so that observations can be made. However, since the observer does not directly influence the behaviour being observed in an unobtrusive observational study, it sometimes requires a lot of time and patience to wait for the behaviour of interest to occur. Consequently, some observational studies can be very time-consuming.

A practical advantage naturalistic observation is that it does not require the co-operation of participants being observed. However, this raises the ethical issue of not obtaining informed consent, particularly if participant observation is required. When participant observation is used without informed consent, a person's expectation of privacy can be violated. This issue has to be weighed against the fact that the participants are not informed that they will be observed in some special way so that their observed behaviour is more likely to be true to life.

Another limitation of an observational study is that it can be difficult to determine the causes of the behaviour of interest that is observed, because many factors may influence that behaviour. This is especially the case in a natural environment. For example, the researcher observing aggressive behaviour from the sidelines outside a King Street night club will often not be able to determine with certainty why people become aggressive towards one another when a skirmish or fight breaks outs unless they intervene in some way; for example, by interviewing those involved. The true factors that control a particular behaviour could be ones of which the researcher is not immediately aware.
A potential limitation of any observation procedure is observer bias, which is a type of experimenter bias. It is possible, for example, that researchers sometimes unconsciously distort what they see so that it resembles what they hope to see, even when they are using structured formats. Researchers who collect the data must be trained to observe and record accurately in order to minimise the influence of their personal biases. Furthermore, when recording participant responses or making detailed notes as part of the observation process, the researcher may neglect to record certain behaviours that they either judge to be irrelevant or do not actually see. To overcome these limitations, researchers often use two or more observers for data collection and check for inter-rater (‘inter-observer’) reliability. This procedure usually results in a more complete and accurate set of data than one observer could obtain alone.

Another limitation of naturalistic observation in a field setting is that it can be difficult to determine the causes of observed behaviour, because there are many factors which may influence that behaviour. For example, a researcher could not determine through observation alone why some children become aggressive towards others in the school yard.

A potential limitation of any observational study is the experimenter effect involving observer bias. It is possible, for example, that researchers sometimes unconsciously distort what they see so that it resembles what they hope to see. This is especially a potential problem with participant observation because the researcher can unintentionally cause participants to behave in the expected way. However, the use of observation checklists or score sheet such as that in figure 2.35 can help minimise observer bias.

Finally, in making detailed records if required as part of the observation process, psychologists may neglect to record certain behaviours which they either judge to be irrelevant or do not actually see. To overcome this limitation, when using observation, researchers may use a team of trained observers, then compare their notes to identify common points or themes. This often results in a more complete and accurate set of data than one observer could obtain alone.

**Self-reports**

A self-report is the participant’s written or spoken responses to questions, statements or instructions presented by the researcher. For example, a self-report may be responses to questions about bullying, romantic relationships or thoughts when daydreaming, to statements in rating scales measuring the extent of their agreement or disagreement on asylum seeker policy, or diary records on sleep habits or homework activity kept in response to a researcher’s specific request. In most cases, one person’s self-report is compared with those of others responding to the same questions, statements or instructions.

Assuming that the participants are honest, understand the questions, can accurately recall what they have been asked about and are able to give sufficiently detailed accounts of the thoughts, feelings or behaviour under investigation, self-reports can provide useful data on virtually any topic of research interest.

Questionnaires, interviews and rating scales are the most commonly used self-report methods. All use questions or statements requiring participant responses, but they are often distinguished in terms of how the questions or statements are asked and answered. For example, a questionnaire usually involves asking and answering questions in writing, whereas an interview usually involves asking and answering questions orally. However, this is not a fixed rule. Sometimes, a researcher may prefer to orally ask the questions in their questionnaire.

Although questionnaires, interviews and rating scales can be used exclusively or in combination to collect self-reports, they are also commonly used to collect additional data as a part of research studies using other method, such as experiments, case studies and observational studies.

**Questionnaires**

A questionnaire is a written set of questions designed to draw out self-report information from people on a topic of research interest. It has a structured format and the questions are usually answered by participants in writing, at their own pace and without supervision.

Questionnaires are most often used when responses are required from a large number of participants; for example, as part of a survey. They are an efficient way of collecting self-reports because a researcher can administer the questionnaire via surface mail, over the phone, the internet, or at the same time to a group who are located in the one place, such as in a school or workplace.

By guaranteeing anonymity to participants, written questionnaires can be a useful way of collecting self-report data that people are not willing to disclose publicly, such as ambitions, motivations, fantasies, sexual behaviour, gambling behaviour, addictive behaviour, socially unacceptable behaviour and illegal behaviour.

**LEARNING ACTIVITY 2.17**

**Review questions**

1. (a) What is an observational study?
   (b) Give an example of an observational study with an independent groups design, but not an example from the text.
   (c) Explain why this study would not be considered to be a true experiment.

2. Distinguish between each of the following.
   (a) A naturalistic and a contrived observational study
   (b) Structured and unstructured observations
   (c) Participant and non-participant observations
Interviews

An interview usually involves questions that are asked by the researcher with the aim of obtaining self-report information on a topic of research interest. Interviews are most often conducted with individuals, in a face-to-face meeting or sometimes by phone. They usually require spoken answers to questions and are rarely used with large samples as data collection would require a considerable amount of time. Unlike questionnaires, which are usually structured, interviews may be structured, unstructured or semi-structured.

In a structured interview, the participant is asked specific, pre-determined questions in a controlled manner. The most structured interview is when the interviewer simply reads closed-ended questions to participants and records their answers. The interviewer follows a script and the questions are read in a neutral manner with no comments or cues. This is done to ensure that all participants are treated in the same way and demand characteristics are minimised. A less structured interview may use open-ended questions, but the researcher will follow a script to ensure consistency across all participants.

In an unstructured interview, the researcher has an overall aim of what data should be collected but the questions asked can vary widely from participant to participant. There is also freedom of discussion and interaction between interviewer and participant. For example, the interviewer may ask additional questions to follow up on a participant’s response.

A goal of unstructured interviews is to allow people to describe their thoughts, feelings and behaviour in their own way using their own terms. This is different from structured interviews (and questionnaires) for which participants have to use the questioner’s terms and concepts to describe how they think, feel or behave. However, this also means that the data collected through unstructured interviews is much more detailed, has far less structure, and is therefore more difficult to analyse, summarise and describe for reporting purposes.

With this, however, comes a limitation. Answers to free-response questions are often difficult to summarise or score. This makes it harder for researchers to statistically analyse, describe and interpret the data obtained.

To avoid or overcome this limitation, researchers often ask fixed-response questions. Fixed-response (or closed) questions usually provide a respondent with a number of ‘fixed’ alternative answers from which they are required to choose.

Examples of such questions are ‘Have health warnings on cigarette packets led you to cut back on smoking: Yes, No, Undecided?’ and ‘How much time do you usually spend on Facebook each night: 0–30 minutes, 30–45 minutes, 45 minutes–1 hour, 1–2 hours, more than 2 hours?’

Answers to fixed-response questions are usually easier to interpret than are answers to free-response questions. In addition, because fixed-response questions provide specific alternatives from which the participant chooses, the researcher can accurately and concisely summarise

BOX 2.3

Free-response and fixed-response questions

When using a questionnaire or interview to collect self-report data, the researcher may choose to use free-response and/or fixed-response questions.

Free-response (or open-ended) questions require respondents to describe their thoughts, feelings or behaviour ‘freely’ in their own words. For example, the researcher might ask a question such as, ‘How do you feel when you are confronted by a barking dog?’ or ‘Why do you draw graffiti on trains?’ These kinds of open-ended questions enable participants to provide detailed responses without being restricted to giving answers that fit into pre-determined categories (such as those of fixed-response questions). Furthermore, free-response questions enable the researcher to ask questions of clarification or follow-up questions as participants give information about the thoughts, feelings or behaviour under investigation.
and describe the responses numerically. For example, a ‘0–30 minutes’ response to the Facebook question can be assigned a score of 1, ‘30–45 minutes’ a score of 2, and so on. Furthermore, the same scores can be reliably assigned to all other participants who give these responses and all responses can be efficiently analysed, described and interpreted using statistical procedures and tests.

**FIGURE 2.37** Answers to fixed-response questions about health warnings on cigarette packets will probably be easier to interpret than answers to free-response questions.

### Rating scales

A rating scale uses fixed-response questions or statements for which participants rank (‘rate’) each item by selecting from a number of choices. They may be used to get data on any behaviour or mental process about which a participant can provide information.

For example, participants may be asked to rate their level of happiness or sadness, how often they feel lonely, how much they like meeting new people, how easy it is to fall asleep, how tired they are when the alarm sounds in the morning, their level of stress when lining up to enter an exam, their confidence in an answer to an exam question, or their attitude to capital punishment. The questions or statements to which participants respond are usually related as they have been devised by the researcher for the topic or issue under investigation.

Responses are typically assigned numerical values which enable answers to be quantified (converted to numbers) for summary, analysis and interpretation. The rating scale is like a multiple choice test, but the answer options represent levels or degrees of a particular characteristic. The best known and most commonly used rating scale is the Likert scale. This consists of about 20 questions or statements to which the participant responds using a five-point scale. It is most commonly used to measure attitudes. For example, in a study on attitudes to war, a Likert scale statement could be ‘War is sometimes necessary to maintain justice’. Participants may then be required to rate their answers by selecting one response from five options ranging in strength, such as strongly agree, agree, neither agree nor disagree, disagree or strongly disagree.

Researchers have several choices in selecting how answers should be indicated on the five-point scale — for example, ticking or crossing a blank space, circling a number or underlining a response. Each of the responses has a numerical value (e.g., from 1 to 5) and the respondent’s attitude is defined as the sum (total) of these values. A Likert scale for measuring attitudes towards illegal drugs could include statements such as those shown in box 2.4.

When developing a Likert scale, half the attitude statements are worded in a positive way and half are worded negatively. For statements 1, 3 and 5, the answers would be scored as follows: SA = 1, A = 2, N = 3, D = 4 and SD = 5. For questions 2, 4 and 6, the answers would be scored in reverse: SA = 5, A = 4, N = 3, D = 2 and SD = 1.

When a respondent has completed the Likert scale, all of the responses are scored and a total is calculated. The result is a score on the attitude or whatever else was measured. Generally, the higher the score, the more favourable the attitude. Box 2.5 on page 54 describes how to construct a Likert scale.

**BOX 2.4**

**Sample items in a Likert scale for measuring attitudes towards illegal drugs**

Circle your response to each statement below.

1. The use of illegal drugs is a major social problem in Australia today.  
   SA A N D SD
2. There should be no restrictions on using illegal drugs as long as the individual using them does not harm anyone else.  
   SA A N D SD
3. Laws should be strictly enforced regarding the use of illegal drugs.  
   SA A N D SD
4. It is an invasion of privacy when law enforcement authorities search people suspected of carrying illegal drugs.  
   SA A N D SD
5. Individuals using illegal drugs should be punished severely.  
   SA A N D SD
6. In the privacy of their own homes, individuals should be allowed to use any illegal drug they desire.  
   SA A N D SD

<table>
<thead>
<tr>
<th>SA = Strongly agree</th>
<th>A = Agree</th>
<th>N = Neither agree nor disagree</th>
<th>D = Disagree</th>
<th>SD = Strongly disagree</th>
</tr>
</thead>
</table>

CHAPTER 2 Research methods in psychology 53
How to construct a Likert scale

The following steps enable you to construct a Likert scale to collect quantitative data for your own research on an attitude. Although your scale is likely to be a useful measure of an attitude, it will not be valid or reliable. This means that you will have to be careful with the conclusions you draw from the results obtained.

**Step 1**
Identify an attitude towards an object, group, issue or event of interest or importance to you.

**Step 2**
Write a list of different aspects of the attitude topic. For example, the Likert scale on illegal drugs in box 2.4 is based on aspects such as crime, punishment, civil liberties, privacy laws and impact on Australian society. If you have difficulties in generating a list, you may find it helpful to discuss your topic with others.

**Step 3**
Use your list to develop a group of attitude items (questions or statements) on the topic. Although Likert scales usually contain about 20 items, you should consider a scale based on about six or eight items. Generally, the list should consist of items which deal with different points of view on the topic. Consider the following guidelines for writing Likert scale items.
- Write items that are unlikely to be agreed with by everyone or no-one. About half of your items should be favourable towards the topic and the other half unfavourable. The more effective items will be those that tend to push respondents towards the strongly agree or strongly disagree ends of the scale. Try to avoid including items which are neutral and likely to cluster responses in the uncertain category (that is, ‘neither agree nor disagree’).
- Use simple, clear language that is suited to the experience, age, and cultural and educational background of the participants whose attitudes you are measuring.
- Write your items in such a way that they are unambiguous and only one interpretation is possible.
- Write each item so it contains only one complete idea.
- Avoid using words such as ‘all’, ‘always’, ‘none’ and ‘never’.

**Step 4**
When you have written your items, trial (‘test’) them with people who will not be a part of your sample but who have personal characteristics in common with those likely to be in your sample. This will assist you to identify problems with your items which you may not have noticed.
- Form your items into a list, with columns for respondents to indicate whether, and to what extent, they agree or disagree with each item. Randomly distribute positive and negative items in the list to avoid a pattern of responses.
- Present the items in a questionnaire format. The questionnaire should have a short introduction that includes instructions for respondents. For example: ‘Here is a list of statements about . . . Please read each statement quickly but carefully, then indicate whether you agree or disagree with each one by putting a circle around: SA = Strongly agree
A = Agree
N = Neither agree nor disagree
D = Disagree
SD = Strongly disagree.’

**Step 5**
- Make several copies of your questionnaire and test your questions again by asking two or three people with similar backgrounds to those in your sample to rate each response.
- Determine their scores for each response and then calculate their score for the entire scale. Score responses by allocating 1 for the most negative response, through to 5 for the most positive response for each item.
- Analyse the responses to determine which items you should include in the final scale. The best items are those that have a very high or very low relationship with the total score for all items. You may wish to rewrite or even replace items that seem to cluster responses in the neutral/unsure category.


Advantages and limitations of self-reports

Self-reports such as questionnaires, interviews and rating scales are widely regarded as useful techniques for collecting any type of data on how people think, feel and behave. In particular, they can be an efficient means of collecting data from a large number of people in a relatively short period of time.

By guaranteeing anonymity, questionnaires in particular provide a means of collecting self-report data on ‘sensitive’ or controversial topics that many people are not willing to disclose publicly, such as in an unstructured oral interview. However, like other self-reports, they rely on the assumptions that people are actually willing to answer all questions and that they will give accurate answers. We cannot always rely accurately or communicate information about how we think, feel or behave.

Another limitation of self-reports is social desirability. People may intentionally give false or misleading answers to create a favourable impression of themselves. For example, with socially sensitive issues such as attitudes to ‘boat people’, Aboriginal land rights, same-sex marriage and cruelty to animals, people sometimes give socially desirable responses instead of reporting their true attitudes. They want to appear likeable, to have a ‘social conscience’, or to
look good, so they present attitudes which encourage others to see them in a positive way.

Alternatively, the participants may be embarrassed to report their true attitudes or feelings, especially for very personal topics. Furthermore, in self-reports based on interviews, the interview situation may be a source of experimenter effect whereby the interviewer’s personal biases and prejudices influence how questions are asked and how the respondent answers them.

Even when researchers make careful use of random sampling, they need to consider the possibility of a type of sampling bias known as non-response bias. For example, if only a small percentage of randomly sampled people agree to respond to a questionnaire, it is quite likely that those who did respond will be different than those who refused or did not bother to participate.

Self-reports are language dependent so there are limitations when used with young children, adults with English speaking backgrounds but with weak literacy skills, people from non-English speaking backgrounds who have yet to learn English well (unless translated) and people with a severe intellectual disability. Generally, they are best used with people who have well-developed language skills, although interpreters and skilful interviewing can help overcome communication barriers.

When comparing the advantages and limitations of different self-reports, it is important to take account of the type of data that will be collected and the type of question used. Generally, questions that allow free, open-ended descriptive responses (a type of qualitative data) give answers that are richer in detail. However, these responses are often difficult to summarise and statistically analyse. Questions with scoreable fixed responses (a type of quantitative data) enable more precise and efficient statistical summaries and analyses.

### LEARNING ACTIVITY 2.18

#### Summarising three self-reports

Complete the following table.

<table>
<thead>
<tr>
<th>Self-report</th>
<th>Description</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaires</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviews</td>
<td>Structured</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unstructured</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rating scales</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### LEARNING ACTIVITY 2.19

#### Selecting an appropriate research method

For each of the following topics, decide which research method(s) — experiment, cross-sectional study, case study, observational study or questionnaire/interview — would be the most appropriate for undertaking the investigation and briefly explain why.

1. Investigating why someone with great potential gave up their career for a job they don’t really like so that they can spend more time with their family
2. Investigating differences in use of game apps by 2-, 3- and 4-year-olds
3. Investigating how young adolescents behave on their first date
4. Investigating whether people will obey a person in authority who ordered them to hurt another person
5. Investigating whether boys and girls in preschool have different preferences for book reading
6. Investigating differences in the diets of adolescents who jog regularly and those who do not jog at all

---

**FIGURE 2.38** Self-report methods of data collection provide useful information about human thoughts, feelings and behaviour. However, they rely on participants having well-developed language skills and being able to accurately recall and state the information required of them.
TYPES OF DATA

All psychological research involves collection of information. In research, the information which is collected is called data. The data is empirical evidence that will form the results of the study and be the basis of the conclusions that will be made.

Data can take different forms. The type of data collected is determined by the specific kind of research method used. For example, questionnaires and interviews often provide data in the form of words, whereas data collected in experiments is usually in the form of numbers.

There are many ways of classifying data. We consider the distinctions between primary and secondary data and qualitative and quantitative data in relation to psychological research.

Primary and secondary data

**Primary data** is data collected directly by the researcher (or through others) for their own purpose, usually to test a hypothesis. It is collected from the source and is sometimes described as ‘first hand’ data. For example, you will collect primary data when you conduct an experiment to test a research hypothesis for a practical activity. You will also be collecting primary data if you interview people as part of a survey. The primary data will be the participants’ responses, which have not been processed. Raw data is a type of primary data.

When you summarise your data as a table or convert it to percentages, it will still be primary data because you are the researcher who collected and processed it. You have also retained control over it.

When someone else accesses your primary data, you lose control over it because they can manipulate or use it in whatever way they want for their own purpose. It will be secondary data for the other person.

**Secondary data** is data that has been collected by someone other than the original user for their own purpose. It has been collected by some other individual, group or organisation and will not be used for the first time, which is why it is referred to as ‘secondary’ (like second-hand). For example, when you access data in a journal, book or at a website for your self-directed research investigation in Unit 1, then you will be using secondary data. The Australian Bureau of Statistics is a widely used source of secondary data, as are the results reported by researchers in journal articles.

The main difference between primary and secondary data is who collects the original data. Both types of data have their advantages and limitations.

Primary data offers tailored information sought by the researcher to test a hypothesis on a topic of their choosing. To the researcher, there is little doubt about the quality of the data collected. They are also responsible for the quality of their data, but it can be time-consuming to collect and process.

Secondary data tends to be readily available and can usually be accessed in less time, especially if you know where and how to look. There can be uncertainty about its quality because it was collected for another purpose and there is often a need to comb through it to find what you’re looking for.

Qualitative and quantitative data

Primary and secondary data may be qualitative or quantitative. The majority of studies referred to in this text use quantitative data. This reflects the preference for quantitative data in most psychological research. Qualitative data is information about the ‘qualities’ or characteristics of what is being studied. It may be in the form of descriptions, words, meanings, pictures and so on. It can describe any aspect of a person’s mental experiences or behaviour. More specifically, what something is like, or how something is experienced.

Qualitative data may be collected as written or verbal statements made by participants, or as descriptions of behaviour observed and recorded by the researcher. For example, a researcher studying self-esteem in young children may collect qualitative data through an interview by asking children open-ended questions related to their self-esteem. Likewise, a researcher interested in learning about the factors that enable some people to cope better than others with personal trauma may collect qualitative data through a questionnaire involving participants’ responses about how they felt in a specific traumatic situation and how they dealt with their feelings.

**Figure 2.39** Qualitative data may be collected through research on the life experiences of refugees and asylum seekers in a detention centre.
Quantitative data is numerical information on the ‘quantity’ or amount of what is being studied; that is, how much of something there is. It may be raw data that have not been analysed in any way, such as lengths or weights of prematurely born infants, or percentages of participants who respond with ‘Yes’ or ‘No’ to survey questions, or the mean reaction time of participants when a light is flashed onto a screen in an experiment, and so on.

All types of mental experiences and behaviours can be described in quantitative terms, as quantities or numbers. For example, in a survey, a question might ask participants to use a five-point scale to rate their feelings on issues such as compulsory school uniform or the persuasiveness of a particular advertisement.

Information about individuals' scores on a range of psychological tests such as intelligence tests, personality tests and various ability and interest tests are also provided as quantitative data. In addition, data collected during experiments are typically collected in a numerical form and are therefore usually quantitative.

The use of numerical data makes it easier to summarise and interpret information collected through research. This is why quantitative data is often preferred to qualitative data, although this does not mean that qualitative data is less important or less useful than quantitative data.

Experiments can produce qualitative data as well as numbers. For example, consider the Stanford Prison Experiment conducted by Zimbardo (1972) (see pages 397–9). Zimbardo observed, described and reported common and unusual behaviour and verbal responses of the ‘prisoners’ and ‘guards’, often referring to specific examples (qualitative data). However, Zimbardo also measured the extent to which the ‘prisoners’ were prepared to obey the demands of the ‘guards’ and reported his observations in the form of graphs and tables (quantitative data).

Although qualitative data is typically expressed in the form of words, it can be converted into a quantitative form. For example, participants’ responses to open-ended interview questions about their thoughts and feelings when they are anxious could be summarised as numbers based on the frequency (‘how often’) or intensity (‘how strong’) with which certain feelings are reported.

**Objective and subjective data**

*Objective data* is information that is observable, measurable, verifiable and free from the personal bias of the researcher. For example, the data can be seen, heard or touched (observable), counted or precisely described (measurable), could be confirmed by another researcher (verifiable) and is factual (free from personal bias). All scientifically conducted investigations target collection of objective data.

Data collected through a strictly controlled experiment in which observations and measurements are planned, precise and systematic is considered objective. So is data collected using an assessment device that yields a score, such as an intelligence or personality test.

Automated and mechanical devices can also be used to collect objective data. For example, an instrument that shows underlying physiological activity in measurable form, such as an EEG which records brain wave activity, provides objective data.

Sometimes researchers collect information about behaviour or mental processes that cannot be directly observed; for example, sexual behaviour or criminal acts. In these cases, researchers tend to rely on self-reports — participant responses to questions asked by the researcher. This information will be subjective.

*Subjective data* is information that is based on personal opinion, interpretation, point of view or judgment. Unlike objective data, this data is determined by the research participants and often cannot be verified by the researcher. It is often biased, can vary from person to person, day to day from the same person, and is not always entirely accurate.

When using subjective data, researchers assume that participants are honest, can accurately recall what they are asked to describe and are able to give detailed accounts about their thoughts, feelings or behaviour.

Although subjective data may be more detailed than that available from more scientifically rigorous methods under controlled conditions, it tends to be difficult to interpret accurately when compared with objective data (which is usually quantitative).
LEARNING ACTIVITY 2.20

Review questions

1. (a) Distinguish between primary and secondary data with reference to an example of each data type.
   (b) Define the terms qualitative data and quantitative data with reference to examples that are not used in the text.

2. (a) Indicate whether the data collected in each of the following research studies is primary or secondary data. Explain your answers.
   (i) Audio recordings of a student’s description of the effect of background noise on their ability to learn previously unseen material
   (ii) A student’s ratings on a seven-point rating scale used to assess how much background noise affected their ability to learn previously unseen material
   (iii) A newspaper article reporting psychological research
   (iv) A documentary on a controversial psychological experiment narrated by the experimenter
   (v) Records of whether people who wear glasses can read more quickly than people who do not wear glasses
   (b) Indicate whether each of the following research designs will obtain qualitative or quantitative data, or both.
      (i) Conducting an experiment to investigate whether having regular rest breaks during a prolonged study session improves performance on a test
      (ii) Observing the social interactions of students in a study group using pre-determined items on an observation checklist
      (iii) Organising a small number of participants into a discussion group to study the experience of sexual discrimination in the workplace
      (iv) Using a written questionnaire with closed-ended questions (e.g. Yes/No) to survey a large number of bushfire victims who may be experiencing post-traumatic stress disorder
      (v) Observing the effects of using a treat as a reward to teach a dog to sit on command

LEARNING ACTIVITY 2.21

Media analysis of psychological research reporting

Locate a newspaper, magazine or internet article that reports psychological research. Make a copy of the article so that it can be presented on an A3 sheet of paper or within PowerPoint. If required, reduce the size of the article but ensure it is still legible.

Using point form and lines, arrows or shapes to or around relevant information, complete the following tasks on the article.

1. Identify the research method used to collect data.
2. Suggest a possible research hypothesis if not stated.
3. Identify the data type(s) collected e.g. qualitative, quantitative.
4. Identify the sample and sample selection procedure (if stated).
5. (a) Outline the main finding(s) of the study reported in the article.
   (b) Comment on information that should have been included in the report to enable the reader to judge the accuracy of the reported findings.
6. Suggest a potential limitation of the research, taking account of possible sources of bias and potential extraneous and confounding variables.

ORGANISING, PRESENTING AND INTERPRETING DATA

When data have been collected to test a hypothesis, the researcher must decide whether the results support or do not support the hypothesis. The researcher must also draw a conclusion(s) relating to the hypothesis. This conclusion(s) must be based on the results obtained and limitations of the conclusions should be identified, described and explained. Reasons must be suggested about why the particular results were obtained and what they mean, including whether they can be applied to other groups or situations. In addition, suggestions for further research and evidence are often made.

To support all of these requirements, researchers use statistics to analyse (‘break down’) and describe the data they collect. They also use statistics to help organise, present and interpret (‘make sense of’) the results obtained from their research.

Two main kinds of statistics are used in psychology — descriptive statistics and inferential statistics. Descriptive statistics are used for analysing, organising, summarising and presenting results. They include calculations such as percentages and mean scores, and preparation of tables and graphs. Inferential statistics are used for interpreting and giving meaning to results. Like descriptive statistics, inferential statistics involve the use of mathematical procedures. However, unlike descriptive statistics, inferential statistics involve judgments about the results.
Descriptive statistics

Suppose that a researcher is interested in studying whether body image (a person’s ‘view’ of their body) changes during puberty and adolescence. The researcher might give a body image rating scale to ten 10-year-olds, ten 12-year-olds, ten 14-year-olds, ten 16-year-olds, ten 18-year-olds, ten 20-year-olds and ten 22-year-olds. Each research participant would be required to make a judgment about their physical appearance using a rating scale ranging from 1 to 10, with 1 being equivalent to very unattractive, 5 to neither attractive nor unattractive and 10 to very attractive. In all, there would be 70 bits of data (i.e., ratings) about the body image of participants in different age groups. How can the researcher make sense of all these different bits of information so that meaningful conclusions about body image and age can be drawn?

The first step would be to use descriptive statistics to analyse, organise and summarise the data so that it can be described and interpreted. It is difficult to draw conclusions about whether (and if so, how) body image changes with age by looking at 70 individual ratings. Thus, in order to compare the body image ratings of the seven different age groups to determine whether there is a change with age, the data for each group could be summarised and presented in a table.

Tables

A table is an orderly arrangement and display of data in columns and rows. The columns and rows are usually identified by names (or ‘headers’) that assist in making comparisons. Some conventions (standards) for tables used in psychology are:

- All tables should be consecutively numbered e.g. Table 1, Table 2.
- Each table should have an individual title that is italicised (except the word ‘Table’ and its number) and has each word in the title capitalised (except words such as ‘for’, ‘of’, ‘in’, ‘and’, ‘with’).
- The title should be a clear statement which explains what the table is about without being too long; for example, Mean Body Image Ratings of Each Age Group.
- The table number and title should be on separate lines with the table number above the title. For example:
  
  Table 1.
  Mean Body Image Ratings of Each Age Group
  - Each column should be identified using a descriptive header.
  - The first letter of each header in the table should be capitalised.
  - The reader should be able to quickly work out what the table is about and comparisons of data should be easy to make.
  - In the research report, essay or other document, the word table is capitalised whenever referring to it, e.g. ‘as shown in Table 1’.

Table 2.4 below provides some order to the data on body image ratings by organising the ratings into different age groups.

Table 2.4 Raw data — participant ratings

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Participant ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>22</td>
<td>8</td>
</tr>
</tbody>
</table>

However, comparison of ratings across the age groups is still difficult because the data have been inadequately summarised. To enable the ratings of different age groups to be compared, a single number that summarises all the data for each age group would be calculated.

For this investigation, the researcher could calculate the mean rating for each age group. The mean scores could be used to describe the ‘average’ body image rating for each age group and would enable the researcher to compare the different age groups. This is shown in Table 2.5. The mean is another type of descriptive statistic.

Table 2.5 Mean body image ratings of each age group

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Mean scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5.9</td>
</tr>
<tr>
<td>12</td>
<td>4.7</td>
</tr>
<tr>
<td>14</td>
<td>4.6</td>
</tr>
<tr>
<td>16</td>
<td>4.6</td>
</tr>
<tr>
<td>18</td>
<td>4.5</td>
</tr>
<tr>
<td>20</td>
<td>6.9</td>
</tr>
<tr>
<td>22</td>
<td>7.0</td>
</tr>
</tbody>
</table>

In this section we consider the descriptive statistics specified for study in VCE Psychology. Note that this textbook uses its publisher’s conventions for tables (and graphs), not the conventions used in psychology for formal research reports.

Graphs

A graph is a pictorial representation of data. Graphing or plotting data typically involves the use of two lines (axes) drawn at right angles to one another. The horizontal line is the X axis and the vertical line is the Y axis. The point where the axes intersect is called the origin (0). Generally, the frequency (for example, the number of cases or amount of something) is plotted on the Y axis. The unit of measurement (for example, time, weight) is plotted along the X axis.

Graphs are best used to show patterns or trends in the data collected; for example, how often a response
is made, how aspects of behaviour change over time or as a research participant's experience changes.

Various types of graphs display data in different ways. The kind of graph used depends mainly on the type of data collected and what needs to be displayed. Among the more commonly used graphs in psychology are bar charts and line graphs. In psychology, graphs are more formally referred to as 'figures' (along with drawings, photos and any type of illustration).

As with tables, there are conventions for presenting graphs. These include:
- All graphs should be consecutively numbered e.g. Figure 1, Figure 2.
- Each graph should have an individual title. The title is not italicised, but the word Figure and its number are e.g. Figure 1. Reaction time of each age group. The title should be a clear statement that explains what the graph is about without being too long.
- The number and title are both on the same line and shown below the graph.
- Both the horizontal and vertical axes must be labelled clearly and indicate what is plotted.
- The reader should be able to quickly work out what the graph is about.

**Bar charts**

One type of graph is the bar chart. A **bar chart** is a graph which uses a series of discrete (separate) bars or rectangles adjacent (next) to, but not touching one another, to enable comparisons of different categories of data. The bars can be positioned horizontally or vertically. One axis is used to show the types of categories (e.g. age, sex, type of response) and the other category is used to show the frequency with which each category occurs (e.g. how often, how much).

One important feature of a bar chart is that each of the categories shown in the graph is separate and there is no continuation between one category and the next; for example, there would be separate bars for data about female participants' responses and male participants' responses. Each bar is the same width and has a small space between it and the next bar.

Figure 2.41 shows an example of a bar chart. Researchers who studied the type of play in which four- to five-year-old children engaged recorded the type and amount of time children spent participating in each type of play at a kindergarten over a one-week period.

The type of play in which children engaged was categorised according to American psychologist Mildred Parten's (1932) system for classifying play behaviour. Parten described four main types of play: solitary play, when the child plays alone and independently; parallel play, when the child plays alone and independently alongside, but not with, other children; associative play, when the child plays with other children in a similar activity, but in their own way; and cooperative play, when the child plays with other children at the same activity.

The researchers were testing the relevance of Parten's theory among children today. They added a further category called unoccupied play which was observed when the child did not engage in any play at all for a period of time.

<table>
<thead>
<tr>
<th>Types of play</th>
<th>Amount of time (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solitary play</td>
<td>15</td>
</tr>
<tr>
<td>Parallel play</td>
<td>10</td>
</tr>
<tr>
<td>Associative play</td>
<td>12</td>
</tr>
<tr>
<td>Cooperative play</td>
<td>15</td>
</tr>
<tr>
<td>Unoccupied play</td>
<td>5</td>
</tr>
</tbody>
</table>

**FIGURE 2.41** Example of a bar chart showing results of a study on play types.

Sometimes a bar chart is used to present values or scores for two different categories within each bar. For example, figure 2.42 shows mean scores on a test of memory (recall) obtained by males and females of different ages.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>10–14</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>15–19</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>70–74</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>75+</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

**FIGURE 2.42** Mean scores on a test of recall obtained by males and females of different ages.
LEARNING ACTIVITY 2.22

Representing data using a bar chart
A researcher obtained data from a group of university students on the relaxation techniques they found to be most effective for minimising anxiety experienced prior to exams. The raw data are described in table 2.6 below. Present these data: (a) in another table that summarises the raw data (b) in a bar chart.

**TABLE 2.6** Types of relaxation techniques used

<table>
<thead>
<tr>
<th>Participant</th>
<th>Relaxation technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Meditation</td>
</tr>
<tr>
<td>2</td>
<td>Drinking coffee</td>
</tr>
<tr>
<td>3</td>
<td>Drinking coffee</td>
</tr>
<tr>
<td>4</td>
<td>Listening to music</td>
</tr>
<tr>
<td>5</td>
<td>Exercise</td>
</tr>
<tr>
<td>6</td>
<td>Meditation</td>
</tr>
<tr>
<td>7</td>
<td>Sleeping</td>
</tr>
<tr>
<td>8</td>
<td>Listening to music</td>
</tr>
<tr>
<td>9</td>
<td>Exercise</td>
</tr>
<tr>
<td>10</td>
<td>Listening to music</td>
</tr>
<tr>
<td>11</td>
<td>Exercise</td>
</tr>
<tr>
<td>12</td>
<td>Sleeping</td>
</tr>
<tr>
<td>13</td>
<td>Meditation</td>
</tr>
<tr>
<td>14</td>
<td>Drinking coffee</td>
</tr>
<tr>
<td>15</td>
<td>Exercise</td>
</tr>
<tr>
<td>16</td>
<td>Exercise</td>
</tr>
<tr>
<td>17</td>
<td>Meditation</td>
</tr>
<tr>
<td>18</td>
<td>Sleeping</td>
</tr>
<tr>
<td>19</td>
<td>Sleeping</td>
</tr>
<tr>
<td>20</td>
<td>Listening to music</td>
</tr>
</tbody>
</table>

---

**Line graphs**

A **line graph** uses points connected by lines to show how one variable changes (e.g. reaction time) as another variable changes (e.g. a person’s age, group size and time taken to complete a task) as shown in figure 2.43.

When used to show the results of an experiment, the horizontal, or X, axis usually has the independent variable plotted on it, with the numerical value of the data increasing from left to right along the axis. A line graph that describes the relationship between group size and time taken to complete a task would list the group size in terms of the number of members of the group on the X axis, in intervals; for example, beginning at two, then three, four people and so on. One important feature of a line graph is that the variable plotted on the X axis is continuous; that is, there is a series of progressively increasing values that can be listed.

The vertical, or Y, axis usually has the dependent variable (the measure of performance) plotted along it. A line graph that described the data from the experiment on group size and time taken to complete a task would record the amount of time taken along the Y axis, in intervals; for example, beginning at zero (which is a convention or ‘rule’ for graphs), then one, two, three, four and five minutes where five minutes is slightly higher than the maximum time ever taken by any group to complete the task (figure 2.43).

Various points on a line graph represent the score on one axis that corresponds with a value on the other axis. The intersecting point can represent a corresponding IV/DV score on the two variables by one research participant, or the mean score of a group of participants.

---

**Percentages**

Suppose you conduct an observational study to find out whether boys are more aggressive than girls during lunch time in the prep area of the school grounds at a local primary school. You want to obtain quantitative data, so you work out a list of observable behaviours that you consider to be aggressive, such as pretend fighting and intentional pushing or shoving.

Whenever you see a boy or girl demonstrating one of the aggressive behaviours on your list, you record your observation with a tick and shift your attention...
to another child. Of the 25 boys you observe, six use an aggressive act and are therefore judged as aggressive, and four of 16 girls observed are judged as aggressive.

On the basis of these results, more boys than girls were aggressive. However, more boys than girls were also observed. In order to reach a conclusion, you need to work out whether \( \frac{6}{25} \) is more than or less than \( \frac{4}{16} \). This can be achieved by calculating the percentages of boys and girls who were aggressive, then making a comparison.

A percentage is a statistic that expresses a number as a proportion (or fraction) of 100. The term per cent means ‘per hundred’, or ‘for every hundred’. It is shown using the per cent sign (%). For example, 65% is equal to \( \frac{65}{100} \) and means 65 parts out of 100; 100% of something means all of it. A percentage is calculated using the formula

\[
\% = \frac{\text{subtotal}}{\text{total}} \times \frac{100}{1}
\]

It is easy to calculate a percentage when the original amount is 100. For example, if you complete a 100 item speed and accuracy test and correctly answer 90 items within the time limit, then your percentage score is:

\[
\frac{90}{100} \times 100 = 90\%
\]

For the data obtained in the observational study of aggressive behaviour:

- boys: \( \frac{6}{25} \times 100 = \frac{6 \times 100}{25} = \frac{600}{25} = 24\% \)
- girls: \( \frac{4}{16} \times 100 = \frac{4 \times 100}{16} = \frac{400}{16} = 25\% \)

This means that the proportion of boys (calculated ‘out of 100’) who were aggressive in the school grounds is slightly less than the proportion of girls. The main problem in making a comparison of the boys and girls based on the raw data is that the two groups were of unequal size. Calculating a percentage for each group overcame this problem and enabled comparison of the scores for boys and girls.

Percentages are commonly used in psychology to describe data; for example, scores on a test, categories of scores, changes or trends in scores, the percentage of people who respond in a particular way (such as correct or incorrect, agree or disagree, do something or do not do something) and the percentage of people in a socio-cultural group (such as gender, age, income level, educational qualifications and ethnicity).

### LEARNING ACTIVITY 2.23

#### Calculating percentages

1. Calculate percentages for the following raw data. Round your answer to the nearest whole number.

   (a) Sixteen out of 62 participants observed in the library broke a rule at least once during a 10 minute observation period. What percentage of participants broke a rule at least once?

   (b) Results of an online survey show that 52 out of 75 students watch TV before school. What percentage of students watch TV before school? What percentage do not watch TV before school?

2. A researcher gave parents a 50 item questionnaire on child-rearing practices. There were 28 fathers and 44 mothers in the sample. The raw data were first organised in a table (shown below) to enable comparison of scores achieved by fathers and mothers. Complete the table by calculating each percentage to the nearest whole number.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Fathers</th>
<th>Mothers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw score</td>
<td>Per cent (%)</td>
<td>Raw score</td>
</tr>
<tr>
<td>0–8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10–20</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>21–30</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>31–40</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>41–50</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>100%</td>
</tr>
</tbody>
</table>

#### Mean as a measure of central tendency

A measure of central tendency is a score that indicates the ‘central’ or ‘average’ value of a set of scores. When a measure of central tendency is calculated, it often provides a ‘typical’ score for a set of scores.

Suppose you collected data for a practical activity which involved comparing males and females on a test for speed and accuracy of visual perception. The research participants are in five year 7 classes, each of which has 25 students. Data for each of the 125 students — 65 girls and 60 boys — are obtained. To help determine which group performed best, a measure of central tendency could be calculated. This would provide a single score for girls and a single score for boys. Scores could then be compared to estimate which group of participants, boys or girls, performed best on the visual perception test. The mean is the most commonly used measure of central tendency.
The mean is the arithmetical average of all the individual scores (or measures) in a set of scores. It is calculated by adding all the scores together and dividing the total by the number of scores. For example, 10 four-year-olds were required to complete a seven-piece jigsaw puzzle. The length of time (in seconds) it took each child to complete the puzzle is listed below:

26, 17, 21, 18, 12, 17, 18, 24, 25, 17

The mean for the group is calculated by adding the scores together (195), then dividing the total by the number of scores (10). The mean is 19.5 seconds. The formula for calculating the mean is shown as:

$$\bar{x} = \frac{\sum (\text{sum or total of all scores})}{N(\text{number of scores})}$$

In this example, the mean provides the most exact measure of central tendency. However, in other sets of data, the mean may not always provide the most accurate measure of central tendency of a set of scores, especially if the scores cluster at the extreme ends of the set of possible scores.

For example, if a set of scores consisted of 140, 140, 140, 140, 180, 180, 180, 180, the mean would be 160. Suppose that these data referred to the height (in centimetres) of players in a girls’ netball team. A manufacturer of netball skirts would be surprised when the players attended for fitting of their skirts, having been informed that the mean height is 160 centimetres. The skirts would not fit any of the players — they would be either too short or too long. Thus, when a mean is provided for a set of data, it doesn’t necessarily follow that any of the individual scores will be the same value as the mean or even approximate it.

Often the mean is calculated to several decimal places. In many instances this does not create a problem; however, sometimes the mean score may become meaningless in real life. For example, if the mean number of children per family in Australia is 1.75, it is difficult to imagine what 0.75 of a child means.

**FIGURE 2.44** The mean does not always give a completely accurate picture of scores.

---

**LEARNING ACTIVITY 2.24**

**Calculating a mean**

1. (a) What is a measure of central tendency and what does it indicate?
   (b) When would a researcher use a measure of central tendency to describe data in preference to a summary of the data in a table or graph?

2. A researcher studied the effects of stress during pregnancy on the birth weight of infants. The following data on the birth weight (in kilograms) of infants whose mothers stated they were stressed for the majority of their pregnancy were obtained:
   2.8, 3.3, 2.5, 3.0, 2.9, 4.4, 3.5, 2.7, 3.2, 4.4.

3. The birth weight (in kilograms) of infants whose mothers stated they were not stressed during their pregnancy were also obtained. These data were:
   3.8, 5.4, 4.3, 4.2, 3.5, 4.1, 4.0, 4.4, 3.9, 5.4.

   (a) What is the mean weight of the sample of infants?
   (b) Is the mean weight an accurate representation of the infants’ weights? Explain your answer.
   (c) What is the median weight of the infants?
Standard deviation as a measure of variation around the mean

Suppose that two psychology teachers discussed the abilities of their respective classes. The teacher of Class A explained that the mean of her students’ results for a test was 78%. The teacher of Class B replied that the mean of his students’ results for the same test was 68% and that his students must therefore be less capable than his colleague’s. ‘But how do you know I’m not just an easy marker? One of my students got 97%.’ Then again, another student got 18%,’ responded the Class A teacher. The Class B teacher was surprised: ‘The lowest mark in my class was 53%, but my highest mark was only 81%,’ he said, ‘so how do we know which class has the better abilities?’

The discussion between the teachers indicates that a mean, on its own, doesn’t provide the complete description of the data. The mean describes the ‘central’ value of a set of scores. In order to more accurately represent the data, a second kind of descriptive statistic is often used — a measure of variation.

A measure of variation, also called variability, indicates how widely scores are distributed or spread around the central point. The sets of scores in figure 2.45 both have the same mean, but they differ in variation; that is, how far the scores are either side of the mean. The distribution of Class A scores shows that it is tightly packed around the mean, indicating low variation (or variability). The distribution of Class B scores is more widely spread from the mean, indicating high variation, or variability.

The standard deviation summarises how far scores within a set of scores spread out, or deviate, from the mean for those scores. If all the scores in a set of scores were the same, there would be no variation and the standard deviation would be zero because none of the scores would be spread out from the mean. A low standard deviation indicates that there is little variation in the scores and that most scores are clustered around the mean. In this case, the mean is a representative descriptive statistic (figure 2.46, curve C). The higher the standard deviation, the greater the variation there is among the scores (figure 2.46, curve A).

The standard deviation is a particularly useful descriptive statistic in that it provides a point of comparison between two different sets of scores. For example, suppose a replacement teacher comes to a new school hoping for an easy day’s work. The replacement teacher is offered either of two classes, both of which have a mean IQ score of 100. There appears to be no difference between the two classes. The teacher is then informed that the standard deviation of IQs in one class is 1 and the standard deviation in the other is 3. Since a higher standard deviation means more variability, the class with the standard deviation of three may take more effort to teach because students vary more in ability.

**FIGURE 2.45** Class A and Class B both have the same means. The distribution of scores for Class A (yellow) shows low variation, as indicated by the clustering of scores around the mean. The distribution of scores for Class B (blue) shows high variation, as indicated by a greater spread of scores from the mean.

**FIGURE 2.46** This graph shows three distributions of scores, each with a different standard deviation. The purple curve has the highest standard deviation and the orange curve has the lowest standard deviation.
**LEARNING ACTIVITY 2.25**

**Review questions**

1. What does a measure of variation (variability) indicate?

2. What information does the standard deviation provide about the distribution of scores?

3. (a) Two classes sat the same practice Psychology exam. The following descriptive statistics were calculated from the students’ results in each class:

   - Class A: mean 70%.
   - Class B: mean 70%.

   On the basis of the mean scores alone, what might teachers of these classes conclude about the knowledge of students in each Psychology class? Explain your answer.

   (b) Suppose the teachers then calculated the standard deviations for their respective classes and obtained the following results:

   - Class A: mean 70%; standard deviation: 0.5
   - Class B: mean 70%; standard deviation: 2.3.

   On the basis of this additional information, what conclusions might the teachers now draw about the knowledge of the students in each Psychology class? Explain your answer.

**Inferential statistics**

Descriptive statistics are very useful for analysing, organising, summarising and presenting results. However, in order to interpret the data and find out whether the results are meaningful, inferential statistics are used. Of particular importance to the researcher is to determine whether the results support the hypothesis that was tested and whether the results can be applied to the population from which the sample was drawn. Inferential statistics are used to help interpret the results and make these types of judgments.

In most experiments, there will usually be a difference in the mean scores of the experimental and control groups. For example, the mean score on a task requiring concentration (time taken to thread a needle) achieved by the experimental group with a loud noise present (the IV) may be 15.7 seconds and the mean score on the same task the control group without noise may be 12.6 seconds. Can it be concluded that the difference in mean scores is due to the independent variable?

Perhaps the difference was due to chance factors in that the experimental and control groups were slightly different in their composition of relevant participant variables despite the use of random allocation. Perhaps there were a few individuals with slightly better eye–hand coordination in the control group.

**Standard deviations in a normal distribution**

When standard deviations are represented on the X (horizontal) axis of a normal distribution curve, the percentage of scores falling between the mean and any given point on the axis is always the same.

For example, 68.26% of the scores will fall within one standard deviation either side of the mean; 95.44% of the scores will fall within two standard deviations either side of the mean. These percentages apply consistently in a normal distribution curve, irrespective of the size of the standard deviation.
Perhaps the light was slightly better when the control group participants performed the needle-threading task. Perhaps the experimenter was tired and slightly impatient when instructing the experimental group participants. Perhaps the effects of these three apparently controlled variables ‘added up’ and combined in such a way as to cause the difference.

How big does the difference in the mean scores between two groups need to be before it can be said that the difference is due to the independent variable rather than chance factors? What is an acceptable difference between the mean scores? Is a difference such as 3.1 seconds (15.7 minus 12.6) ‘big enough’? Is 2.5 seconds a ‘big enough’ difference? 4.5 seconds?

One way to find out if the results of an experiment are due to the IV that was tested rather than chance factors is to repeat these studies several times in exactly the same way with the same participants to see if the results are about the same each time the study is replicated. This would be very time consuming, inconvenient and possibly impractical because participants may not be continually available. However, it is usually unnecessary to undertake these replications. A more efficient way of measuring the reliability of the results is to use inferential statistics by applying a test of statistical significance to determine the extent to which chance factors may account for the results.

**FIGURE 2.48** Is loud noise a distraction when threading a needle?

**BOX 2.8**

**Statistical significance and p values**

Tests of statistical significance can be used to determine the extent to which chance operated in an experiment and whether it is at an acceptable level. The tests enable a precise mathematical value to be obtained that will indicate the probability (likelihood) that if the same experiment were repeated, the results would be similar or different. If the likelihood of the difference occurring by chance is at an acceptable level, then it is said that the difference is statistically significant. In general, psychologists accept a given result, such as the difference in mean scores, as statistically significant if it is found that the probability or likelihood that the result might be due to chance is 5 or fewer times ≤ 5 in 100, or a 1 in 20 chance, if the same study were to be repeated 100 times. The way of saying this is that the result is significant at the 0.05 level; that is \( p < 0.05 \).

A significance level of \( p \leq 0.06 \) (less than or equal to 0.06) would indicate that there was a 6% (or 6 or less in 100) chance that the result obtained was most likely due to chance and this would generally be viewed as unacceptable. It would then be said that the results are not significant and therefore do not support the research hypothesis.

The significance level of any difference is called a p-value, with ‘p’ standing for probability. An acceptable p value for results is established before the experiment is conducted.

In some cases, a stricter probability level than \( p \leq 0.05 \) is used, such as \( p \leq 0.01 \) (less than or equal to 1 in 100) and \( p \leq 0.001 \) (less than or equal to 1 in 1000). Such a probability level would be used when the findings of the research are so important that the researcher wants to be extremely confident of the results; for example, when the research hypothesis being tested involves a radical new way of treating depression or if it contradicts a research finding or theory that is widely accepted.

If research is being undertaken in an area that is likely to be of immense benefit to the community, or if it involves a treatment that carries with it some chance of harm, then replication of the study is still likely to occur.

In some other cases, a researcher might be prepared to accept a more lenient level of significance than \( p \leq 0.05 \). For example, a researcher may conduct a pilot, or ‘trial’, study on a research topic of interest to see if it is worthwhile carrying out a full-scale research study. The researcher may set a significance level of \( p \leq 0.1 \) (10%). This would indicate that there may be a significant difference in the mean scores obtained. Therefore, it is worth continuing with further research, perhaps with refinements to the procedures.
Conclusions and generalisations

When the results have been evaluated, evidence-based conclusions need to be drawn. A conclusion is a decision about what the results obtained from a research study mean. All conclusions must be based on evidence (the results), be consistent with the evidence and relevant to what was actually investigated.

One type of conclusion relates to whether the hypothesis is supported on the basis of the results obtained. This requires careful examination of the results so that an objective (‘unbiased’) decision can be made. Although the results alone may indicate that the hypothesis is supported, the DV and therefore the results may have been influenced in a significant way by one or more variables other than the IV (or in addition to the IV). Therefore, uncontrolled extraneous variables and potential confounding variables also need to be considered when drawing a conclusion. The researcher must be confident that any change in the DV was due to the IV alone and not any other variable.

The conclusion about the hypothesis is expressed as a statement in the research report. In psychology, a hypothesis may be supported or it may be refuted (rejected), but it cannot be ‘proven’ true. This is because no matter how much evidence a researcher finds to support their hypothesis, there may still be one or more alternative explanations, some of which are not yet known or even thought of, that could better explain the results.

Another type of conclusion that can be made is called a generalisation. A generalisation is a decision about how widely the findings of a research study can be applied, particularly to other members of the population from which the sample was drawn. Because a study usually tests a sample from the population of interest rather than the whole population, making a generalisation is a process of forming an idea about whether findings obtained from a limited number of cases (the sample) can be extended to apply to an entire class of objects, events or people (the population).

In experimental research, generalising the results from the sample to the population is risky if the sample is not representative of the population of interest. Like any other conclusion, a generalisation must also be based on the results obtained and must consider the potential extraneous and confounding variables, as well as any other problems with the study.

When drawing conclusions, researchers try to avoid making errors or overstating what the results mean. For example, they attempt to ensure that:
- all conclusions are consistent with the results
  - any influential extraneous variables or potential confounding variables have not been overlooked
- analysis and interpretation of the results enables an accurate finding about whether or not the hypothesis is supported
  - any gaps in the results and further evidence that may be required are identified

• limitations of the sample used in the study have been considered.
• any generalisations are reasonable
• the explanation of the findings is reasonable and supported by the results.

Reliability and validity in research

An important goal of research is to obtain results that are both reliable and valid. This will mean that the results are consistent and accurate. It also means that other researchers are more likely to accept the findings. Reliability and validity are not necessarily ‘present-or-absent’ features of a research design. Instead, they are considered to vary in degree on a scale ranging from low to high.

Reliability

Reliability refers to the extent to which results obtained from a research study are consistent, dependable and stable. For example, if you measured your blood alcohol level on a breathalyser and then decided to double check it straight away, you should expect to get the same result.

Similarly, if you conducted an experiment on a group of participants and repeated it again with a similar group under the same conditions, you should expect the results to be very similar on each occasion the experiment is conducted.

Because conducting an experiment is a more complicated process than measuring your blood alcohol level, it is not likely or expected that the results will be identical each time the experiment is conducted. The main reason is because of individual differences within another sample. However, if the results are to be considered reliable, then they should be similar each time the experiment is repeated.

FIGURE 2.49 The general idea behind research reliability is that significant results must be more than a ‘one-off’ finding and be replicable (repeatable).
A researcher always sets out to obtain reliable results. However, when their study is repeated, it may be found that the results are not reliable. This is more likely to occur if the study is not repeated in exactly the same way in which it was first conducted; for example, if there are differences in important personal characteristics of participants, or if the conditions under which the study was first conducted are significantly different in some way.

Validity

Validity refers to the extent to which the procedure(s) used for a research study measures what it intended to measure. If valid, this means that the research study has produced results that accurately measure the behaviour or mental process that it claims to have measured. For example, if you measured your biceps with a cloth tape measure that had been left outside in the open weather for a long time and had become inaccurate through stretching, the result would not be a valid measure of your true bicep size.

The inaccurate cloth tape measure, however, is reliable as it will give you the same result each time it is used. This means that a measure can be reliable even though it is not valid, but a measure cannot be valid unless it is reliable.

Another type of validity relates to the conclusions (including any generalisation) the researcher makes about a study. In this case, the results are valid if the conclusion(s) drawn by the researcher is (are) correct. This means that the conclusion is specifically based on those variables that the study was investigating and the data obtained from the study.

For example, if a researcher concludes that a new drug they tested in an experiment reduces symptoms of depression, or that participants in a taste-preference study preferred Coca-Cola™ over Pepsi™, the research is valid only if the new drug really works or if the participants really did prefer Coca-Cola™ (Stangor, 1998).

As with seeking reliability, researchers always attempt to conduct valid research; that is, they attempt to draw accurate conclusions from their data. Yet often, despite a researcher’s best intentions, their research is invalid or not as valid as it could have been. This can occur for a number of different reasons.

Sometimes a researcher may draw a conclusion from their data that cannot actually be drawn; that is, the data do not actually justify, support or ‘back up’ the conclusion. Another reason that research and its results may be invalid is because one or more extraneous variables have not been adequately controlled, have become confounding variables, and have therefore influenced the results in an important way.

For example, in an experiment, a confounding variable and the IV may both affect the results. When this happens, the researcher will find it difficult to separate the effects of the IV and the confounding variable and therefore cannot be certain whether it was the IV or the confounding variable that caused the change in the DV.

Internal and external validity

Researchers often distinguish between the internal and external validity of their studies. They consider both internal and external validity in judging the overall validity of a study. These types of validity are not necessarily ‘present-or-absent’ features of a research design. Instead, they vary on a scale ranging from low to high.

Internal validity refers to the extent to which the results obtained for a study are actually due to the variable(s) that was tested or measured and not some other factor. For example, in an experiment, the researcher needs to be confident that the change in the DV was produced solely by the IV and not by any extraneous or confounding variable, nor due to chance.

If a study has gaps or ‘flaws’ in its procedures, such as the use of a sampling method that resulted in an unrepresentative sample, then it may be considered as lacking in internal validity.

Similarly, if participants were required to rate facial attractiveness, then the researcher needs to be confident that the scores (ratings) actually and only measured facial attractiveness. Internal validity may be lost if participants did not understand the rating procedure or their ratings partially reflected the style of dress worn by each person in a photo.

External validity refers to the extent to which the results obtained for a study can be generalised to the population from which the sample was drawn or to other people in other settings. For example, suppose that a researcher conducted a laboratory experiment on the effects of stress on behaviour using a relatively small sample of participants. If the experiment has high external validity, this means that the conclusions can more confidently be generalised to apply to the population from which the sample was drawn and to situations outside, or ‘external’ to, the laboratory.

Generally, the more representative a sample is of its population, the more confident the researcher can be in generalising from the sample to the population.

LEARNING ACTIVITY 2.26

Review questions

1. What can be achieved with inferential statistics that is not possible with descriptive statistics alone?
2. Explain the meaning of chance factors in relation to research with reference to an example.
3. (a) What kind of decision is made about the research hypothesis?
   (b) On what is this decision based?
4. (a) What is a generalisation?
   (b) What are the important considerations in making generalisations and other conclusions from the results of a study?
5. What does reliability mean in relation to research?
6. (a) What does validity mean in relation to research?
   (b) Distinguish between internal and external validity.
ETHICS IN PSYCHOLOGICAL RESEARCH AND REPORTING

Is it appropriate for a researcher to inflict pain on a person in order to study mental experiences associated with pain? Does your answer depend on the amount of pain, or is any amount of pain unacceptable? Does it matter if the pain is psychological rather than physical? Should a person know exactly what the research will involve before they participate; for example, the specific experimental procedures to which they will be exposed? Should a participant be allowed to opt out of a research study whenever they want to, regardless of their reason? What if the researcher has gone to great expense to conduct the research? What if the research has important benefits for humankind? Such questions raise important ethical issues that need to be considered by researchers.

The term ethics refers to standards that guide individuals to identify good, desirable or acceptable conduct. Essentially, ethical standards help us to make judgments about which behaviours are appropriate (‘right’) and inappropriate (‘wrong’). ‘Ethical conduct’ is more than simply doing the right thing. It involves acting in the right way out of respect and concern for one’s fellow creatures (NHMRC, 2007).

All societies and cultural groups have ethical standards that guide the behaviour of their members. In addition to these standards, most professions have their own standards of ethical conduct that must be followed. For example, just as it would be considered unethical for a medical doctor to discuss a patient’s condition with anyone apart from the patient or people legally responsible for the patient, so too would it be unethical for a psychologist to reveal information discussed in a counselling session or the results of a psychological test to anyone apart from the client (or the guardians of the client if the client is a child under a guardian’s care).

Ethical standards and considerations also apply to any type of research or data collection method involving people (or animals). These help ensure that the wellbeing and rights of research participants are respected and protected before, during and following their involvement in the research. In addition, ethical standards help prevent unnecessary research and promote research that is or will be of benefit to the wider community or humankind in general.

The Australian Psychological Society (APS) has a Code Of Ethics (2007) which provides standards and guidelines for all psychological research (and other areas of professional practice). The Code of Ethics has been devised with reference to a national set of standards and guidelines in a document called the National Statement on Ethical Conduct in Human Research 2007 (Updated March 2014). This is simply referred to as the National Statement.

**National Statement on Ethical Conduct in Human Research**

The National Statement has been jointly developed by the National Health and Medical Research Council (NHMRC), the Australian Research Council and the Australian Vice-Chancellors’ Committee. The NHMRC is the Australian government’s expert body for providing advice on research.

The purpose of the National Statement is to ‘promote ethically good human research’. It is organised around four values — research merit and integrity, beneficence, justice and respect for human rights. The design, review and conduct of all research with people as the participants must reflect each of these values.

1. **Research merit and integrity:** Research that has merit is worthwhile and conducted appropriately to achieve the aims. It has the potential to contribute to knowledge and understanding, and to improve social welfare and individual wellbeing. It must be properly designed and undertaken by people with suitable expertise. Research that is conducted with integrity is carried out with a commitment to the search for knowledge and understanding, to following recognised principles for conducting research and to

**FIGURE 2.50** Ethical standards for human research ensure all participants are given the respect and protection due to them, irrespective of who they are.
the honest conduct of research, including accurate and responsible reporting of findings, whether the results are favourable or unfavourable.

2. **Beneficence**: Research beneficence refers to the likely benefits to participants or the wider community. The researcher must consider and maximise all possible good outcomes while minimising the risks of harm to participants and to the community in general. The potential benefits must justify any risk or harm or discomfort to participants.

3. **Justice**: Research that is *just* has a concern for the use of fair procedures and fair distribution of costs and benefits. The process of recruiting and selecting participants should be fair so a researcher must avoid imposing on particular groups an unfair burden of participation in their research. Similarly, the benefits of the research should be distributed fairly between the participants and the wider community.

4. **Respect for human beings**: *Respect* is demonstrated when the researcher recognises and takes account of the rights, beliefs, perceptions and cultural backgrounds of all participants. In particular, all participants have the rights to privacy, confidentiality and to make informed decisions about matters that affect them. People must be protected and empowered if they are vulnerable or their capacity to make informed decisions is impaired; for example, children and intellectually disabled people who depend on others.

All four values apply to all research conducted with or about people, including experiments, cross-sectional studies, questionnaires, interviews, observational studies, psychological testing or treatment, and analysis of personal documents or other materials with information about participants.

### Role of ethics committees

The National Statement requires that all research that carries more than a low level of risk to human participants must first be reviewed and approved by an ethics committee. This type of committee is formally called a **Human Research Ethics Committee (HREC)**. A HREC has a minimum of eight members, with a mix of researchers and non-researchers (including community members). Its main purpose is to assess research proposals for approval purposes, and then monitor the conduct of the research (if approved) to ensure all relevant ethical standards are *adopted* and followed.

Generally, the roles and responsibilities of the HREC include:

- deciding whether a research proposal meets all the requirements of the National Statement and is therefore ethically acceptable
- deciding whether the researcher(s) is adequately experienced and qualified (or the researcher is supervised by a qualified person if there are concerns about their experience and qualifications)
- monitoring approved research (e.g. through progress reports, random inspections of research sites, interviews with participants)
- handling complaints (e.g. from participants, the wider community)
- ensuring accountability of the researcher (e.g. the researcher understands, accepts and maintains responsibility for all aspects of their research).

If the committee is satisfied that all ethical questions and issues raised by the research have been dealt with satisfactorily, approval will be given for the research to proceed. If the committee
has concerns about some aspects, it can highlight these and return the application to the researcher so the concerns can be addressed, possibly with suggestions on how. If the proposal has ethical issues that cannot be addressed, then the research will not be allowed to proceed.

HRECs are usually established by organisations (public, not-for-profit or private) which conduct a considerable amount of research involving humans. Universities and hospitals are the most common of these organisations. Not all organisations which conduct human research, however, have their own HREC. Some organisations and individual researchers use the services of HRECs within another organisation (NHMRC, 2007).

Human research considered to be at a low level of risk, where the only foreseeable risk is one of discomfort, does not have to be submitted to a HREC. In such cases, a research proposal may be reviewed by ‘a competent person or group’ familiar with the National Statement and other relevant ethical standards.

The NHMRC also requires the use of ethics committees for research involving animals. These are called Animal Ethics Committees (AECs) and members have roles and responsibilities similar to those of HRECs.

**Australian Privacy Principles**

The *Privacy Act 1988* is an Australian law which regulates the handling of personal information about individuals. This includes the collection, use, storage and disclosure of personal information, and access to and correction of that information.

Personal information is information or an opinion about any individual who can be identified; for example, information about someone’s racial or ethnic origin; health; genetics; political opinions; religious beliefs and sexual orientation or practices (Office of The Australian Information Commissioner, 2015).

The Privacy Act includes 13 *Australian Privacy Principles* (APPs) which set out standards, rights and obligations for the handling of personal information, some of which apply to psychology research. The APPs include requirements such as:

- **Open and transparent information management** — how personal information will be handled must be clearly expressed and made available
- **Anonymity** — ensure individual participants cannot be personally identified
- **Data collection** — collect personal information only if necessary; ensure informed consent
- **Data use** — use only for the purposes specified
Role of the experimenter

The experimenter (or researcher) must ensure their research is ethically appropriate so that participants are given the respect and protection that is due to them. They must observe all ethical standards, guidelines, codes or legislation such as the National Statement, the APS Code of Ethics and the Privacy Act. These will help them to meet their responsibilities to identify and properly address all ethical issues (APS, 2007; NHMRC, 2007).

Protection and security of participants’ information

The researcher must ensure that personal information is secure and protected from

- misuse, interference and loss; and
- unauthorised access, modification or disclosure.

In addition, the researcher must make provisions for maintaining confidentiality in the collection, recording, accessing, storage, dissemination and disposal of personal information. If personal information about an individual is no longer needed, then the information should be destroyed or de-identified.

Confidentiality

Confidentiality refers to the obligation of the researcher not to use or disclose private information for any purpose other than that for which it was given to them. Participants have a right to privacy, so the researcher must avoid undue invasion of privacy by collecting only information that is needed. In addition, any information that may identify an individual or their involvement in research, such as personal data or test results, cannot be revealed unless consent has been obtained.

The right to privacy and procedures for establishing and maintaining confidentiality must be explained to participants before the study commences. As mentioned above, the confidentiality requirement applies to the collection, recording, accessing, storage, dissemination and disposal of personal information.

Voluntary participation

The researcher must ensure participants voluntarily consent to be involved in a study. For example, participants must not be forced or pressured to take part in a study. The researcher must also ensure that prospective participants do not experience negative consequences if they choose not to be involved in a study.

Withdrawal rights

Participants have an unconditional right to withdraw from a study at any time without giving a reason for doing so. This includes withdrawing their data after the study is finished regardless of the effect this may have on the overall results. Withdrawal rights must be explained to participants before the study commences and the researcher must ensure that participants suffer no negative consequences as a result of withdrawing from the study.

Informed consent procedures

Consent is a voluntary choice for participants and must be based on sufficient information and adequate understanding of both the proposed research and the consequences of participation in it. In order for this to be achieved, information should be given about the purpose, methods, demands, risks and potential benefits of the research.

This information must be presented in ways suitable for each participant; for example, it should be in plain language (with the least possible technical jargon) and the researcher should take account of personal characteristics such as age, educational background, cultural background and any other possible barriers to understanding the information. There should be an opportunity for prospective participants to ask questions about the research.

For participants who are legally unable to give informed consent (for example, children and individuals with an intellectual disability), the researcher must obtain appropriate consent from the persons who are legally responsible for participants’ wellbeing (i.e. parent or guardian).

Consent may be expressed orally, in writing or by some other means that indicates consent (for example, return of a questionnaire), depending on:
(a) the nature, complexity and level of risk of the research; and (b) the participant’s personal and cultural circumstances.

Often, researchers obtain informed consent using a document like that in figure 2.54 on page 74. Two copies are made so that one can be kept by the researcher and one by the participant.

Use of deception

Deception occurs when participants are deliberately misled or not fully informed about the aim or some other aspect of the research. This is sometimes necessary to avoid unduly influencing their responses during the study and consequently the
accuracy of the results. By its nature, deception violates the ethical requirement of informed consent, but it is acceptable if the potential benefits of the research justify its use and there is no feasible alternative to its use. Whenever deception is used, it is essential that all participants are debriefed at the conclusion of the study.

**Debriefing**

Debriefing involves clarifying each participant’s understanding of the nature of the research as soon as possible after it has been conducted. This includes explaining the true purpose of the research and why it was necessary to deceive them, correcting any mistaken ideas and impressions participants may have, and providing an opportunity for questions about any aspect of the study, including the need for deception. Another important requirement of debriefing is to check the wellbeing of the participant and address any harm that may have resulted from their participation in the study; for example, providing information about counselling services and how to access them to help treat any distress resulting from the study.

**Figure 2.53** All ethical standards and guidelines must be followed when conducting psychological research.
CONSENT FORM TO PARTICIPATE IN RESEARCH

TITLE OF RESEARCH: .................................................................

DESCRIPTION OF RESEARCH: Insert an outline of the research and other relevant information. Include:
- aim/purpose/ reasons for the investigation
- method used to collect data
- how the data will be analysed, described and presented
- what the participants will need to do and time commitment
- how confidentiality will be maintained
- whether the participant will have a chance to see and comment on the final report
- what will happen to the final report
- who will read the report and have access to it
- withdrawal right
- name(s) of researcher(s), supervisor/teacher and school
- status of the researcher(s).

I, ............................................................................................, consent to taking part in the research investigation described above. I understand my rights as a participant in this research. The aim and procedures of the study have been explained to me and I understand them.

[Where deception is used a clause such as the following should be included.]

I understand that it is sometimes essential for the validity of research results not to reveal the true purpose of the research to participants. If this occurs, I understand that I will be debriefed as soon as is possible after my participation and, at that time, given the opportunity to withdraw from the research and have records of my participation deleted.

I have been advised the results of the research will be presented in a formal written report but that my personal details will remain confidential.

I voluntarily consent to participate but I understand that I may discontinue participation from the study at any time without giving a reason.

If you have any questions, comments or complaints to make on this research, please contact [insert the researcher’s name and/or the Psychology teacher’s name] at [insert the researcher’s and/or the Psychology teacher’s contact details, including phone number(s)].

Name of Participant: ............................................................
Signature: .................................................................

Name of Researcher: ...........................................................
Signature: .................................................................
Date: .................................................................

FIGURE 2.54 An example of a document for obtaining written consent to participate in research. Researchers often separate the study information from the consent form by using two separate documents — an Information Sheet and a Consent Form. In addition, a copy of the signed consent form is often given to the participant.
Ethical practices and conduct in VCE Psychology

The VCAA Psychology study design has advice on ethical conduct that must be followed by VCE Psychology students and teachers. This advice includes the following:

**Ethical conduct of experimental investigations**
As part of this study teachers and students will be involved in teaching and learning activities that include experimental investigations using human subjects. Teachers and schools have a legal and moral responsibility to ensure that students follow ethical principles at all times when undertaking such investigations. Teachers should refer to the following documents for detailed advice.

- The National Statement on Ethical Conduct in Human Research (2007), issued by the National Health and Medical Research Council (NHMRC) in accordance with the NHMRC Act 1992 (Cwlth), www.nhmrc.gov.au/publications/synopses/eh29syn.htm
- The Code of Ethics of the Australian Psychological Society (APS), www.psychology.org.au
- It is not expected that animals will be used in the teaching of this study. If using animals in teaching, schools must comply with the current legislation including:

**Safety and wellbeing**
This study may include potentially sensitive topics. Teachers should ensure that students have opportunities to consider topics systematically and objectively, and to become aware of the diversity of views held on such matters. Students should not be asked to disclose personal information about their own or others’ health status and behaviours.

When dealing with sensitive mental health matters, students should be specifically advised that they are neither trained nor equipped to diagnose problems or offer any counselling or therapy. Students should be given information as appropriate about sourcing available treatment services within and outside school.

As part of this study teachers and students may consider different psychological assessments, including standardised psychological tests which are designed to be administered only by trained psychologists. Teachers must limit access to such tests and ensure that students understand that such tests should only be administered by a qualified psychologist.

It is the responsibility of the school to ensure that duty of care is exercised in relation to the health and safety of all students undertaking the study. Teachers and students should observe appropriate safety precautions when undertaking practical work. All laboratory work should be supervised by the teacher. It is the responsibility of schools to ensure that they comply with health and safety requirements.

Relevant acts and regulations include:
- Occupational Health and Safety Act 2004
- Occupational Health and Safety Regulations 2007
- Occupational Health and Safety Management Systems (AS/NZ 4801)
- Dangerous Goods (Storage and Handling) Regulations 2012
- Dangerous Goods Storage and Handling Code of Practice 2000
- Hazardous Substances Code of Practice 2000

**Legislative compliance**
When collecting and using information, the provisions of privacy and copyright legislation, such as the Victorian Privacy and Data Protection Act 2014 and Health Records Act 2001, and the federal Privacy Act 1988 and Copyright Act 1968, must be met.

**USE OF ANIMALS IN PSYCHOLOGICAL RESEARCH**

Although psychology is primarily interested in people, about 7–8% of psychological research involves the use of animals. About 90% of the animals used have been rodents and birds, mostly rats, mice and pigeons. About 5% of the animals are monkeys and other primates. Use of dogs and cats is rare (American Psychological Association, (APA), 2015).

Research with animals has and continues to have an important role in psychology. Discoveries through animal research have advanced understanding of human behaviour and mental processes in a diverse range of areas; for example, behavioural and bodily changes that occur when stressed; basic learning processes; the neurobiology of learning and memory; processes of recovery after neural damage; brain plasticity; mechanisms that control hunger and thirst; behavioural and psychological effects of medications used in the treatment of various mental disorders; addiction to illegal drugs; how the senses function and physiological influences on perception; the critical role of early experience in development; attachment; aggression; emotion and cognition (APA, 2015; Bennett, 2012).
- Animals have practical advantages over people for use as research participants. For example, studying the effects of ageing from birth through to ‘old age’ is not generally practical in humans because most people live more than 75 years, compared with rats which have an average life expectancy of two years, or monkeys which live for 15–20 years. Another advantage is that some animal species breed a lot faster than humans. For instance, rats produce a new generation every three months and can be used to study the development of certain behaviours over successive generations within a relatively short period of time. Animals can also be kept for long periods of time in captivity in laboratories and it is easier to observe their behaviour under these conditions.

- The behaviour of animals can usually be controlled to an extent not possible with human participants. For example, a rat can be raised from birth in a cage. The rat can then be used in a learning experiment and the psychologist will have a good idea of what it has already learned before the experiment is conducted.

- When certain experiments require large numbers of participants who have, for example, the same genetic background, animals are more easily obtained than humans.

- Participant expectations can influence the results of an experiment; however, animals don’t usually have expectations and they are not able to guess the purpose of an experiment.

Many arguments have been presented against the use of animals in psychological research. One argument is that it is not possible to apply (generalise) the results of animal studies to humans because the species are not the same even though there may appear to be similarities. An issue for researchers is how far they can generalise about human mental experiences and behaviour from the results of animal studies. If laboratory animals die after prolonged sleep loss, would humans? If a drug causes a brain disorder in animals, should it be banned for human use? Another argument is that humans should respect animals and protect them from harm rather than use them in research. It is also suggested that humans do not have the right to dominate other species.

In order to ensure that all reasonable steps are taken to minimise the discomfort, illness and pain to animals used in research, ethical guidelines have also been established for the use of animals in research. The use and care of laboratory animals must be directly supervised by a person competent to ensure their comfort, health and humane treatment. The care and use of animals in research must follow the NHMRC Guidelines to promote the wellbeing of animals used for scientific purposes (2008).

The main reasons animals are used in psychological research to achieve the kinds of benefits described are:

- Some studies cannot be conducted with humans due to the risk of psychological and/or physical harm that may be caused, or because suitable human participants are unavailable. Various examples are included throughout this text.

- Bodily systems and/or behaviours of some animals are similar to those of humans; therefore, using animals can be a ‘starting point’ for learning more about human behaviour.

**FIGURE 2.55** About 7–8% of psychological research involves the use of animals.
According to these guidelines, *any* research with animals, including research activities in schools, can be performed only if the research can be justified. Justification involves weighing the predicted scientific or educational value of the research against the potential effects on the welfare of the animals. If an animal is to be subjected to pain, stress or deprivation (e.g. food, social interaction, sensory stimuli), research may only occur if no other alternative is available. If surgery is to occur, the animals must be given the appropriate anaesthesia so they do not experience pain. When an animal's life is to be terminated, it must be done quickly and painlessly.

**LEARNING ACTIVITY 2.27**

Review questions

1. Define the meaning of ethics in relation to research.
2. What is the main purpose of ethical standards and guidelines for psychological research with human participants?
3. Name and describe the four values that should be reflected in all human research.
4. What are three essential informed consent procedures?
5. What is the ethical responsibility of a researcher who conducts research with human participants, but does not fully inform them of the true purpose of the research before the study begins because it may influence the participants' behaviour?
6. If a research participant became distressed during the research, what should occur?

7. Explain the ethical relevance of the Australian Privacy Principles.
8. (a) What is an ethics committee?
   (b) Give three examples of its roles or responsibilities.
9. Briefly summarise three ethical guidelines that must be followed when planning to use animals in psychological research.
10. (a) Give two reasons for the use of animals in psychological research.
    (b) Give two arguments against the use of animals in psychological research.
    (c) What is the main purpose of ethical guidelines for the use of animals in research?

**LEARNING ACTIVITY 2.28**

**Applying ethical values**

Which ethical research value — *merit, integrity, beneficence, justice and respect for human beings* — is relevant to each of the following?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Ethical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The process of recruiting participants is fair.</td>
<td></td>
</tr>
<tr>
<td>2. The researcher does not ‘make fun’ of a participant’s unexpected responses.</td>
<td></td>
</tr>
<tr>
<td>3. The researcher has a commitment to following all relevant ethical standards.</td>
<td></td>
</tr>
<tr>
<td>4. The researcher does not put pressure on a participant to consent to study participation.</td>
<td></td>
</tr>
<tr>
<td>5. The researcher is certain that what is likely to be learnt from their study justifies the risks of harm or discomfort to participants.</td>
<td></td>
</tr>
</tbody>
</table>
Applying ethical standards to research
Suppose you have been asked to sit on an ethics committee. The following proposals for human research have been presented to your committee for approval. Your task is to evaluate the proposals in terms of whether they meet the standards, then write your recommendations, commenting on:
(a) whether the committee approves or rejects the proposal as it is presented
(b) if the proposal is rejected, on the basis of which ethical principle(s) it is rejected.

Proposal 1
Danielle Foster is a clinical psychologist who is interested in how parents cope with the death of a young child. She proposes to obtain qualitative data through research on grieving parents’ use of support available through the internet.
Foster is particularly interested in chat rooms dedicated to parents who have lost a young child. In order to obtain realistic qualitative data, she intends to pose as a parent who has recently lost a child and participate in discussions in several chat rooms. In the course of her chat room participation, she will raise issues for discussion and make judgments about the quality and usefulness of chat room support.

Proposal 2
Dr Amir is interested in the effect of stress on performance on the McCord IQ Test. He feels that the test, which is widely used in schools, gives misleadingly low scores to students under stress.
He wants to divide his participants (VCE students) into two groups, with 20 participants in each group. All participants will take a fake pretest and will be given their ‘results’. The experimental group will be told that they failed the test and that it is surprising that they were able to do well enough at secondary school to make it through to VCE. The control group will be told that they passed the test with flying colours. All of the students will then be given the real McCord IQ test.
Dr Amir hypothesises that the experimental group will not do as well on the IQ test as the control group. At the end of the experiment, all students will be debriefed and told that the pretest was not real, nor was the feedback following pre-testing. In addition, the true purpose of the study will be explained.


REPORTING CONVENTIONS
The final and very important stage in the research process involves the following preparation of a report on the research and its findings. This is done for detailed written reasons:
• to communicate or ‘share’ the results with others, particularly other researchers interested in what was studied, and
• to enable replication of the study to test the validity and reliability of its results.
When reporting research, psychologists provide a detailed description of the study and its findings. The written report has two important characteristics:
• there is enough information to enable close examination of all stages of the investigation (including the results), and, if required, to replicate the study; and
• reporting conventions are used.
Reporting conventions are well-established and widely recognised standards, or ‘rules’, about how a report is written and presented. Reporting conventions determine aspects of the report such as writing style, its structure and organisation, headings, presentation of tables and graphs, and formats for referencing.
For example, the writing style used in a psychological research report is like that of all scientific reports. The language is formal, clear, concise, written in the past tense, in the third person and using the passive voice. Appropriate phrases that meet these language standards are: ‘An experiment was conducted to test . . .’, ‘Each participant was . . .’, ‘The results show . . .’, ‘It can be concluded that . . .’. Scientific reports are not written using the first person, for example, ‘I did . . .’, ‘We asked . . .’, ‘In my opinion . . .’, ‘I believe that . . .’, ‘. . . and then we asked the participants to . . .’.
Conventions for psychological research reports are based on those described in the Publication Manual of the American Psychological Association, Sixth Edition (2010). This manual is commonly called ‘the APA manual’ and its conventions are commonly referred to as ‘APA format’.
The APA format is widely recognised and used by psychologists throughout the world to guide their preparation and presentation of written and poster research reports. These conventions are also used by psychology students for reporting conducted as part of their university studies.
The following guidelines for writing a research report and referencing are based on the APA format. Guidelines for preparing a poster report are then considered. This type of report is less detailed and is commonly used to provide a summary of the key features of a research study for presentation and display at conferences or meetings with other psychologists likely to be interested in the research and its findings.
Abstract
This is a brief, comprehensive summary (about 120 words) of the entire report, usually presented as a single paragraph on a separate page.

It should include a description of what was investigated (in one sentence if possible), participants (specifying relevant personal characteristics such as age and sex), the key features of the research method, the main result(s) and the conclusion(s).

The abstract is best written last.

Introduction
This section gives the background to the research. It often summarises theory and results of other relevant research. If you are unable to find background information, or it is not required by your teacher, then you should explain the rationale (‘reasoning’) for conducting the research.

The introduction is often written in a way that leads the reader to a statement of the aim (‘purpose’) of the research and the hypothesis which was tested. The hypothesis is usually included in the last paragraph of the introduction and should be expressed as a specific statement which, for an experiment, refers to the independent and dependent variables.

In formal journal articles, the introduction does not have a heading because it is clearly identified by its position in a report.

Method
This section clearly describes how the research was conducted. There should be enough details for the reader to know exactly what was done so that the research could be replicated exactly in order to test the results.

The method is often divided into three sub-sections, each with the relevant heading — participants (or subjects), measures and procedure.

Participants
Includes details on how many participants were used, important characteristics that might have influenced the results (such as age, sex, educational background), the population (i.e., the larger group) from which they were drawn, and how the participants were selected (i.e., the sampling procedure) and allocated to groups. Details of the participants are often presented as a table.

Measures
Describes the test or other means used to collect data. A description of any questionnaire, observation checklist, test items, word lists and so on which were used in conducting the research should be included. For example, you may state that a 10-item questionnaire was used to measure attitudes towards violence in cartoons. Any evidence of the measure’s validity and reliability should also be stated.

Examples or more detailed information about any measure should be included in an appendix at the end of the report.

Written report
A standard research report is presented in sections that follow a set order. However, the structure of the report and organisation of the sections may sometimes be modified to suit the particular investigation.

Generally, the report is presented in a logical sequence that describes:
• what was done
• why it was done
• how it was done
• what was found
• what the findings probably mean.

Although the different sections of the report (described below) are usually presented in the order shown, they do not have to be prepared in that order. For example, the abstract which summarises the investigation appears first in the report but is usually easier to write last.

Title
This should be brief (usually one sentence) and indicate clearly what the research was about e.g., ‘Effect of practice on reading speed’. Quite often, researchers use a statement based on the hypothesis for a title. Words that have no useful purpose should be avoided e.g., ‘An experiment on . . .’

The title should be centred and positioned in the upper half of a cover page. The author’s name is written under the title and centred on the page.
**Procedure**
A detailed description of all the data collection procedures. This information should be presented in a way that another researcher could conduct the same study just by reading your description. Include information about the roles of the researcher, how participants were recruited, what they were asked to do, and so on.

**Results**
This section has a summary of the main results. There should be sufficient detail to justify the conclusion(s). All results should be accurate and displayed clearly. Tables, graphs, charts and other figures are used, depending what suits the type of data collected. The reader should be able to understand any table or figure without referring to another section of the report.

Only summary data should be presented in the results section. If relevant, raw data could be included in an appendix. Detailed comments on the results are included in the discussion.

**Discussion**
This is where the results are examined, interpreted and explained, especially with reference to the hypothesis. It is also where you draw conclusions from the results.

The section usually starts with a clear statement about whether the hypothesis is supported or rejected on the basis of the results obtained. If the results do not support the hypothesis, explain why.

The general relevance of the results to the population from which the sample was drawn, and similarities and differences between your results and theory or other research (referred to in the introduction), is also described in this section.

In drawing conclusions, consider and explain sources of potential bias and other limitations or weaknesses of the research. Then, suggest ways of effectively addressing these if the study were to be replicated.

Finally, try to explain the practical applications of the findings to the real world. Often, this section ends with suggestions for future research.

**References**
This section has a list of all sources cited in the report (but no others). Every quotation or summary of information from another source which is used in the report must be substantiated with a reference.

The list of references should be presented in alphabetical order based on the surname of the first named author of a source. The formats for referencing in psychology are described in box 2.10.

**Appendices (if any)**
This is where materials which do not fit into the other sections of the report and are easily presented in print format are placed. There should be a different appendix for each set or category of materials. Each appendix should be numbered and have a title (e.g. *Appendix 1*. Test items for visual perception skills) and be presented on a separate page.

Materials included in an appendix should be referred to in the body of the report (e.g., *Test items for visual perception skills (see Appendix 1)*).

**Poster report**
A poster is another format for reporting research. In psychology, it is most commonly used for display and discussion purposes. It may also be used for reporting research conducted by students in psychology courses.

Poster formats and their specific headers can vary. Generally, a well-constructed poster is less detailed than the written report, covers the key features of the research and is self-explanatory.

The VCE Psychology study design (p. 13) includes a VCAA template (‘format’) to guide the headings, organisation and content of a *poster report*. This is shown below.

In Unit 4, one of the SACs is a student-directed practical investigation and the VCAA poster template must be used for the report. The poster may be produced electronically or in hard copy and should not exceed 1000 words. The production quality of the poster is not assessed.

The VCAA poster template may be used in Units 1 and 2 but specific requirements can be varied by the teacher. For example, an additional ‘Section’ (such as an *Abstract*) may be included and/or some of the ‘Content and activities’ may be excluded. The maximum word length may also be varied.

<table>
<thead>
<tr>
<th>Section</th>
<th>Content and activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Question under investigation is the title</td>
</tr>
<tr>
<td>Introduction</td>
<td>Explanation or reason for undertaking the investigation, including a clear aim, a hypothesis and/or prediction and relevant background psychological concepts</td>
</tr>
<tr>
<td>Methodology</td>
<td>Summary that outlines the methodology used in the investigation and is authenticated by logbook entries</td>
</tr>
<tr>
<td></td>
<td>Identification and management of relevant risks, including the relevant health, safety and ethical guidelines followed in the investigation</td>
</tr>
<tr>
<td>Results</td>
<td>Presentation of collected data/evidence in appropriate format to illustrate trends, patterns and/or relationships</td>
</tr>
<tr>
<td>Discussion</td>
<td>Analysis and evaluation of primary data</td>
</tr>
<tr>
<td></td>
<td>Identification of outliers and their subsequent treatment</td>
</tr>
<tr>
<td></td>
<td>Identification of limitations in data and methods, and suggested improvements</td>
</tr>
<tr>
<td></td>
<td>Linking of results to relevant psychological concepts</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Conclusion that provides a response to the question</td>
</tr>
<tr>
<td>References and acknowledgements</td>
<td>Referencing and acknowledgement of all quotations and sourced content as they appear in the poster</td>
</tr>
</tbody>
</table>
Referencing in psychology
The APA manual also describes conventions for citing and referencing sources of information used in a research report, essay or other psychological document. The conventions described in the APA manual are based on referencing styles commonly known as the ‘author–date’ or ‘Harvard’ system. There are numerous examples of the APA method within this text and in the references at the back.

Citations
Whenever another source is used to present evidence, give an example, develop an argument and so on, the source must be cited. This procedure helps the reader distinguish between your ‘ideas’ and ‘work’ and those of another person(s).

When writing a research report (or another document), it is sometimes necessary to cite within a sentence, and at other times at the end of a sentence (or paragraph). Examples of how this is done are:

Within a sentence:
One author: In a study by Trotter (2010), participants were required . . .
Two authors: A similar result was reported by Trinh and Jones (2012), who found that . . .
Three authors: Black, White, and Yellow (2013) studied the effects of . . .
Four authors: Samir, Vlahos, Smith, and Chan (2014) concluded that . . .
Five authors: Costas, Sim, Crow, Iona, and Ling (2015) predicted that . . .
Six or more authors: Franklin et al., (2016) proposed that . . .

Three or more authors cited again: Black et al., (2013) concluded that . . .

Note that ‘et al.’ is a short form of ‘et alia’, which is Latin for ‘and others’. In this text we prefer to use et al. for citations from journal articles or texts with four or more authors.

At the end of a sentence:
One author: The sex of the person in need of help was a factor that influenced whether or not the research participants would provide help (Willow, 1980).
Two authors: Our attitudes do not always match our behaviour (Pine & Chan, 2004).
Three to five authors: This behaviour is not unique to humans. It has also been observed in primates such as apes and gorillas (Cole, Schnell & Koumaki, 2005).
Six or more authors: . . . therefore, visual perception is fallible (Black et al., 2005) . . . apes and gorillas (Cole et al., 2005).

Citing references within sources
Sometimes you need to cite a source that was referred to by another author; for example, when you read about a study or research finding that was summarised and cited in a textbook. In this case, you would cite the source as follows:
Watson (as cited in White, 2015) replicated the study using . . .

Quoting from a source
If you copy (word for word) information from another source instead of summarising the information using your own words, you should use quotation marks at the start and end of the quotation, use an ellipsis ( . . . ) when you omit words, and provide the reference and page number.

For example:
Tanaka and Young (2001) explained the observation in terms of ‘ . . . the inability of a three-month-old child to recognise themself’ (p. 18).

The reference list
The reference list includes all references used in compiling the report or other document. The references are presented in alphabetical order based on the surname of the first author (if there is more than one) using the formats in the following chart. There are some minor variations to suit VCE Psychology.

<table>
<thead>
<tr>
<th>Type</th>
<th>Format</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>1. Author. (Surname of author then their initials. If more than one author, all names are presented in the order they appear on the title page of the book.)</td>
<td>Book with one author: Carr-Gregg, M. (2014). Beyond cyber bullying. Camberwell, Vic: Penguin.</td>
</tr>
<tr>
<td></td>
<td>2. Year of publication. (Enclosed in brackets, followed by a full stop)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Edition: (If a second or subsequent edition, ‘edition’ is abbreviated, enclosed in brackets and followed by a full stop.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5. City of publication (and state if city is not well known), followed by a colon</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6. Name of publisher. (Followed by a full stop. Omit terms such as Publishers, Co., and Inc. Retain the words Books or Press.)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Type</th>
<th>Format</th>
<th>Example</th>
</tr>
</thead>
</table>
| Chapter or article in an edited book | 1. **Author of chapter/article.** (Surname of the author then their initials. If more than one author, all names are presented in the order they appear in the title of the chapter/article.)  
2. **Year of publication.** (Enclosed in brackets, followed by a full stop.)  
3. **Title of chapter/article.** (Not italicised, followed by a full stop and the word In.)  
4. **Author of book.** (Initials of author followed by their surname and Ed. in brackets. If more than one author, all names are presented in the order they appear on the title of the chapter/article and followed by Eds. in brackets.)  
5. **Title of the book.** (Italicised and followed by the page numbers in brackets, then a full stop.)  
6. **City of publication** (and state if city is not well known, followed by a colon.)  
**Online:** If available only online, the URL takes the place of the publisher location and name, but it is not underlined, active or followed by a full stop, e.g. Kostas, M. (2016). Train your brain. In I. Smith & S. Battista (Eds.), Advances in neuropsychology (pp. 103–115). Retrieved from http://www.store.elsevier.com/Academic-Press/IMP_5 |
| Journal article               | 1. **Author.** (Surname of author then their initials. If more than one author, the names are presented in the order they appear in the article.)  
2. **Year of publication.** (Enclosed in brackets, followed by a full stop)  
3. **Title of article.** (Followed by a full stop)  
4. **Title of journal.** (Underlined or italicised and followed by a comma)  
5. **Volume number of journal.** (Italicised and followed by the issue number (if there is one) in brackets and not italicised, then a comma.)  
| Pamphlet/brochure/fact sheet  | 1. **Author.** (Surname of author then their initials, or the organisation name, followed by a full stop.)  
2. **Year of publication.** (Enclosed in brackets, followed by a full stop.)  
3. **Title (italicised).**  
4. **Type.** Identify as a pamphlet, brochure or fact sheet. (In brackets, followed by a full stop.)  
5. **City of publication** (and state if city is not well known, followed by a colon.)  
6. **Name of the publisher.** (Followed by a full stop.) | Print copy  
**Online:** If available only online, the URL takes the place of the publisher location and name, but it is not underlined, active or followed by a full stop, e.g. Kostas, M. (2016). Train your brain. In I. Smith & S. Battista (Eds.), Advances in neuropsychology (pp. 103–115). Retrieved from http://www.store.elsevier.com/Academic-Press/IMP_5 |
| Newspaper or magazine article | 1. **Author.** (Surname of author then their initials. If more than one author, the names are presented in the order they appear in the article.)  
2. **Date of publication.** (Enclosed in brackets, with the year before the month and day, followed by a full stop.)  
3. **Title of newspaper/magazine.** (Italicised and followed by a comma)  
If you do not know the author: Personality types in the workplace. (2016, January 12). Herald Sun, p. 32.  
**Online:** If accessed online, the URL takes the place of the page number(s) , but it is not underlined, active or followed by a full stop e.g. Smith, K. (2015, July 15). Study shows busy minds good for health. The Age. Retrieved from http://www.theage.com.au |
| Internet (including YouTube)  | 1. **Author.** (Surname of author then their initials, or the organisation name, followed by a full stop.)  
2. **Date of website publication.** (If available and enclosed in brackets, followed by a full stop)  
3. **Title of article.** (If specified)  
4. **When retrieved if date of publication is not available.** (Year, month, date followed by a comma and the word ‘from’)  
5. **URL.** (Not underlined, active or followed by a full stop) | When publication date is known:  
When publication date is not known:  
YouTube:  
The Royal Children’s Hospital Melbourne (2013, December 18). Having a CT scan of your head (video file). Retrieved from https://www.youtube.com/watch?v=VaDIL97CLM |
## Examples of conventions used by psychologists for referencing

### Motion picture (movie), TV program, DVD, audio
1. **Main contributors.** (Surname first, and, in brackets, the role of the main contributors, usually the director and/or writer)
2. **Year or date released.** (Enclosed in brackets and followed by a full stop)
3. **Title.** (Italicised)
4. **Type.** (Identify as a motion picture or other media type, in brackets and followed by a full stop.)
5. **Origin.** (Give the place of origin, where it was primarily made and the name of the production company)

### Personal communication or interview
1. **Initials and surname of communicator**
2. **Type of communication** (in brackets)
3. **Date of communication** (in brackets and followed by full stop)

**Note:** Personal communications are not included in an APA-type reference list because they are usually not recoverable/accessible.

### Example 1. Journal article (print copy)


### Example 2. Web site


**Figure 2.58** Examples of conventions used by psychologists for referencing. Note the use of commas, full stops and italics. Colour is used here for illustration purposes and is not a convention.
LEARNING ACTIVITY 2.30

Review questions
1. Explain the meaning of the term reporting convention.
2. What is a potential benefit of following conventions when reporting research?
3. List, in their correct order, the main sections of a report on psychological research.
4. Consider box 2.10 on referencing in psychology. Rewrite the following references correctly for inclusion in the reference list for a psychological research report or essay.
   4. An article called Violence Reigns Supreme in King Street on page 22 of the Age on Wednesday 20th July 2009; written by Felicity Mohammad.
CHAPTER 2 TEST

SECTION A — Multiple-choice questions

Choose the response that is correct or that best answers the question. A correct answer scores 1, an incorrect answer scores 0. Marks will not be deducted for incorrect answers. No marks will be given if more than one answer is completed for any question.

**Question 1**
The main purpose of ethical standards for research is to
A. ensure validity and reliability of the results.
B. ensure that the research proceeds scientifically.
C. safeguard the rights and wellbeing of participants.
D. keep problems with participants to a minimum.

**Question 2**
A researcher studied differences in the behaviour of newborn babies who are breast-fed and newborn babies who are bottle-fed. The psychologist conducted the research with 20 mothers and their newborn infants at the Royal Women’s Hospital (RWH). The 20 mothers (and infants) were selected from a group of 45 mothers at the RWH who had all volunteered to participate in the experiment. There were another 50 mothers with newborn infants at the hospital, but these mothers did not volunteer to be in the experiment.

In this experiment, there were _____ mothers (and their infants) in the sample, and _____ mothers (and their infants) in the population.
A. 20; 95
B. 45; 50
C. 20; 45
D. 45; 95

**Question 3**
A researcher interested in the effects of anxiety on exam performance asked research participants to describe how they feel during an exam when they come across a question they know they will get wrong. The participants’ responses were tape-recorded so that they could be analysed at a later time.

The type of data obtained by the researcher is best described as
A. secondary data.
B. quantitative data.
C. numerical data.
D. qualitative data.

**Question 4**
When the researcher replayed the tape of participants’ responses to the questions asked in the experiment described in question 3, it was decided to summarise the data using a table which showed the number of times certain anxiety-related words, such as ‘worried’ and ‘scared’, and non anxiety-related words such as ‘nothing’ and ‘didn’t care’, were used. This type of data is best described as
A. secondary data.
B. quantitative data.
C. qualitative data.
D. quantitative and qualitative data.

**Question 5**
To test the notion that ‘two heads are better than one’, a psychologist measures how long it takes people working either in groups of two or working alone to solve a problem. The independent variable is
A. the problem.
B. the number of people working on the problem.
C. the time it takes to solve the problem.
D. whether or not the problem is solved.

**Question 6**
Minimising unwanted effects of variables other than the independent variable involves controlling
A. participant responses in the experimental condition.
B. participant responses in the control condition.
C. extraneous variables.
D. dependent variables.

**Question 7**
Which of the following procedures would be considered to be unethical when conducting research?
A. choosing only volunteers as participants in an experiment
B. disclosing a participant’s extraordinary test results to the media without obtaining written consent to do so from the research participant
C. testing a child’s ability to do algebra even though the child’s ability to do algebra is already known
D. allowing a participant to discontinue being in the experiment, even though the experiment has started
Question 8
A research hypothesis is
A. a prediction about the results to be obtained for a study.
B. theory or findings from prior research of interest to the researcher.
C. a statement about whether the results apply to the population of research interest.
D. a statement about the accuracy of the results of a study.

Question 9
In an experiment, the group of participants used for comparison purposes in order to measure any change caused by the IV is called a/an
A. independent variable group.
B. placebo group.
C. experimental group.
D. control group.

Question 10
A researcher collected data for a study on the amount of vacation time employees had and their happiness at work. The data were presented in the following graph.

![Graph showing happiness at work vs. number of weeks vacation]

The type of graph they used to show the results is called a
A. line graph.
B. bar graph.
C. chart.
D. measure of variability.

Question 11
The most important feature of a table is that
A. percentages have been calculated.
B. the data are displayed in an orderly arrangement of rows and columns.
C. means have been calculated.
D. all raw data are included and accurately reflect participants’ responses.

Question 12
Which of the following series of steps is the most appropriate sequence for conducting psychological research using scientific method?
A. design the research, collect data, formulate hypothesis, analyse data, interpret data, report findings
B. formulate hypothesis, design the research, collect data, analyse data, interpret data, report findings
C. design the research, collect data, analyse data, interpret data, formulate hypothesis, report findings
D. formulate hypothesis, collect data, design the research, interpret data, analyse data, report findings.

Question 13
To generalise from the results of a research study means
A. overstating the results.
B. stating whether the results can be replicated.
C. restricting the conclusion(s) to the results.
D. applying the results to the sampled population.

Question 14
Validity in research means that
A. the participants knew what they had to do.
B. the researchers knew what they were doing.
C. the research study produced results that accurately measured the behaviour or event that it claimed to measure.
D. the researchers obtained results that were consistent and dependable.

Question 15
A research hypothesis with operationalised variables states
A. whether the results are valid and reliable.
B. the sample from which the population was drawn.
C. whether the IV and DV can be controlled.
D. how the IV and DV will be manipulated and measured.

Question 16
One way of controlling an experimenter effect is to ensure that
A. the experimenter is unaware of the experimental conditions to which research participants have been allocated.
B. the experimenter explains the instructions clearly.
C. the experimenter is informed about which participants are in which groups.
D. only the research assistants know which participants are in the different experimental groups.

Question 17
The internal validity of research study involves its ____, whereas its external validity involves its _____.
A. design; generalisability
B. design; replicability
C. generalisability; replicability
D. generalisability; design
A researcher conducted an experiment on the effectiveness of a new study technique for memorising Chinese characters used in reading and writing. Two groups of participants were used.

Group A were given instructions on how to use the new study technique, then required to memorise a list of 20 Chinese characters. Group B were simply asked to memorise the list of 20 Chinese characters. All participants in each group were given a test of recall on the list of characters.

A mean score of 14 correct responses was obtained for Group A and a mean score of nine was obtained for Group B.

In this experiment, the independent variable is _____; whereas the dependent variable is _____.
A. use of the new study technique; the number of Chinese characters correctly recalled
B. Group A; Group B
C. the number of Chinese characters correctly recalled; use of the new study technique
D. Group B; Group A

The results of a research study are said to be reliable when
A. the researcher has conducted an experiment that established a cause–effect relationship.
B. the results are consistent and dependable.
C. the researcher has drawn conclusions that are accurate.
D. an appropriate research method is used.

A researcher is interested in studying why some people willingly give up their personal time to help others. She has seen people operating a hot soup outlet for homeless people after midnight and decides to survey the operators.

The researcher’s sampling procedure is best described as ____ sampling.
A. random
B. convenience
C. stratified
D. situational
SECTION B — Short-answer questions

Answer all questions in the spaces provided. Write using black or blue pen.

**Question 1** (4 marks)
(a) What is a case study? 1 mark

(b) In what way is a cross-sectional study different from a case study? 1 mark

(c) Describe one limitation of a case study. 1 mark

(d) Describe one limitation of a cross-sectional study. 1 mark

**Question 2** (1 mark)
Explain a potential limitation of small sample size when drawing conclusions.

**Question 3** (5 marks)
(a) What is an order effect? 1 mark

(b) In which type of experimental research design is an order effect more likely to occur? 1 mark

(c) Explain your answer to (b) above. 2 marks

(d) Name the procedure used to minimise or control this effect. 1 mark

**Question 4** (2 marks)
(a) What is the placebo effect? 1 mark

(b) How is this effect best controlled? 1 mark

**Question 5** (2 marks)
(a) What is experimenter expectancy? 1 mark

(b) Explain how experimenter expectancy can produce results due to a self-fulfilling prophecy. 1 mark
Question 6 (5 marks)
A researcher wanted to find out whether the presence of nicotine in the bloodstream is linked to sleep loss, specifically sleep loss resulting from spending more time trying to fall asleep.
To investigate this issue, one group of 15 volunteer students who were smokers and enrolled in the first year of the Psychology course at a Victorian regional university (Group 1) were required to attend the university’s gym at 9.00 pm on Tuesday evening, smoke ten 8 mg cigarettes during a 90-minute period while listening to classical music, and then go to sleep as quickly as they could in one of the standard single beds at the other end of the gymnasium.
On Friday evening later that week, the procedure was repeated with another group of 12 volunteer smokers who were also enrolled in the first year Psychology course (Group 2). However, participants in this group were not permitted to smoke any cigarettes in the one-hour period before being asked to go to sleep in one of the beds.
Both groups were carefully observed from outside the gym by two research assistants through a monitor hooked up to infra-red cameras. The research assistants recorded the precise time when each participant was observed to fall asleep.
The results are presented in the graph below. The researcher concluded that neither smoking nor the presence of nicotine in the bloodstream cause sleep loss.

(a) Identify the operationalised independent and dependent variables. 2 marks

(b) Explain whether the conclusion made by the researcher is justified. 2 marks

(c) Describe one limitation of the research design used for the study. 1 mark

Question 7 (2 marks)
(a) Name the procedure used to control individual participant differences in an independent groups experiment. 1 mark

(b) Explain how this procedure minimises any unwanted influence of this variable. 1 mark
CHAPTER 2 Research methods in psychology

Question 8 (2 marks)
(a) What does it mean to standardise procedures for a research study? 1 mark

(b) Why are the research procedures standardised? 1 mark

Question 9 (1 mark)
Name a data collection method for obtaining self-reports.

Question 10 (1 mark)
Explain why volunteer participants are required to give informed consent.

SECTION C — Research scenario

A psychologist conducted an experiment to investigate a new online method of teaching algebra called Algebratics. She wanted to find out whether teaching young children using Algebratics is more effective than the standard textbook approach to teaching algebra in many schools. The psychologist predicts that children who use Algebratics will score higher on an algebra test.

To test her hypothesis, one class of year 6 students in a local primary school was taught a new algebra topic through Algebratics. This took place for one hour at 9 am on the first four days of the school week. Their teacher was trained in Algebratics to ensure the method was used correctly. Another year 6 class at the same school was taught the same topic, at the same time by their teacher through the standard approach. Both classes had 24 students, each with 11 males and 13 females. Informed consent was obtained for all participants.

On Friday at 9 am, all participants were given a 30-minute test on the topic. Their classroom teacher administered the test under strictly controlled test conditions. The results are shown in figure 1 below.

![Figure 1: Test scores for algebra teaching methods](image)
Question 1 (1 mark)
Identify the experimental and control groups.
experimental group: ____________________________
control group: ________________________________

Question 2 (2 marks)
Operationalise the independent and dependent variables in the experiment.
independent variable: __________________________
dependent variable: ____________________________

Question 3 (1 mark)
Write a research hypothesis for the experiment.

Question 4 (1 mark)
Name the experimental research design.

Question 5 (10 marks)
Construct a discussion containing
- a description and analysis of the results
- the conclusion(s) based on the hypothesis
- an explanation of whether the results can be generalised with reference to external validity
- a description of a limitation or weakness of the experiment.

eBookplus
The answers to the multiple-choice questions are in the answer section at the back of this book and in eBookPLUS.
The answers to the short-answer questions are in eBookPLUS.
The answers to the research scenario are in eBookPLUS.