Psychology — research methods

What research methods and key science skills are used in VCE Psychology?

CHAPTER 1 Research methods in psychology

UNCORRECTED PAGE PROOFS

UNCORRECTED PAGE PROOFS

1 RESEARCH METHODS IN PSYCHOLOGY

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

Source: © VCAA, VCE Psychology Study Design (June 2017 update), pp. 11–12.

CHAPTER CONTENT

Defining psychology and its subject matter 4
Mental processes and behaviour 5
Psychology as a scientific study 6
Steps in psychological research 7
Research methods 13
Sample and population 14
Research hypothesis 15
Experimental research 17
Advantages and limitations of experimental research 59
Cross-sectional studies 66
Advantages and limitations of cross-sectional studies 66
Case studies 69
Advantages and limitations of case studies 71
Observational studies 72
Advantages and limitations of observational studies 75
Self-reports 81
Advantages and limitations of self-reports 86
Types of data 89
Primary and secondary data 89
Qualitative and quantitative data 90
Organising, presenting and interpreting data 94
Descriptive statistics 94
Inferential statistics 103
Conclusions and generalisations 104
Reliability and validity in research 106
Reliability 106
Validity 106
Ethics in psychological research and reporting 110
National Statement on Ethical Conduct in Human Research 110
Role of ethics committees 112
Australian Privacy Principles 113
Role of the experimenter 113
Use of animals in psychological research 116
Reporting conventions 120
Written report 121
Poster report 122

UNCORRECTED PAGE PROOFS

The VCE Psychology study design prescribes a set of key science skills that is a core part of the study of psychology and applies across all areas of study in all units. These skills primarily involve research methods that may be used to undertake investigations and to evaluate the research of others. This chapter focuses on the research methods and the underlying attitudes and principles specified within the key science skills. We start with a brief description of the nature of contemporary psychology to clarify its subject matter and the overall context of VCE Psychology, its aims, areas of study, learning outcomes, key knowledge and key science skills.

DEFINING PSYCHOLOGY AND ITS SUBJECT MATTER

The term psychology originates from two Greek words: *psyche*, meaning soul or mind, and *logos*, which loosely translated means study or knowledge. Therefore, by its original definition, psychology was initially described as 'the study of the soul or mind'. By the late 19th century, when psychology became a distinguishable scientific discipline, it was commonly described as 'the science of mental life'. At this time, 'psychologists' studied the mind by asking their research participants to describe their mental experiences, using questions such as 'What are you thinking?' and 'What are you feeling?'

During the early 20th century, many psychologists adopted the view that a true science can study only overt behaviour. *Overt behaviour* is any action or response made by an organism (person or animal) that is clearly visible and therefore directly observable and readily measured. They rejected the study of mental experiences, as these are *covert*; that is, internal and hidden from view and therefore not directly observable or readily measured. This led many psychologists to change direction from studying mental experiences, such as aggression or forgetting, to studying the outward expressions of these experiences through observable behaviour, such as displays of anger or performance on a memory test. Consequently, in the 1920s, psychology was commonly defined as 'the scientific study of behaviour' (Sdorow, 1995).

This view was dominant until the 1970s, when interest in studying the mind returned, primarily as a result of the development of new technologies such as scanning devices that could capture images of the active human brain while participants engaged in various experimental tasks. This enabled psychologists to more effectively observe and measure the previously 'hidden' activity of the brain. As the new technologies were refined and the discipline of psychology matured, it became increasingly clear to most psychologists that they could not fully understand overt human behaviour without also understanding the mental and biological activity and processes underlying and sustaining externally expressed behaviour.

Currently, definitions of psychology refer to the discipline as involving the systematic study of mental processes and behaviour. This is also the definition adopted for VCE Psychology. However, in VCE Psychology, psychology is defined more precisely as the scientific study of human thoughts, feelings and behaviour. Note the key parts of the definition in Figure 1.1.

UNCORRECTED PAGE PROOFS

(a)

(b)

Figure 1.1 (a) Psychology may be defined as the systematic study of mental processes and behaviour. (b) The psychology symbol is a representation of the Greek letter *psi*.

MENTAL PROCESSES AND BEHAVIOUR

Psychologists usually distinguish mental processes from behaviour. The term **mental processes** generally refers to a person's thoughts and feelings, which are personal, or subjective, and cannot be directly observed. What you think about, how you think when problem solving, how you interpret relationships with others, how you learn, your choice of words in a conversation, your perceptions, how you experience happiness or sadness, your dreaming when asleep, your daydreaming when awake, your emotions and moods, and what motivates you to do something are all examples of mental processes.

These are private, internal experiences that cannot be observed by others in the way that we can see externally expressed actions such as smiling and crying. Consequently, psychologists rely on making inferences about mental processes on the basis of observable or measurable behaviour. An *inference* is a logical assumption, judgment or conclusion based on available evidence. For example, learning cannot be directly observed as it is a mental process that occurs within the individual. Instead, psychologists observe (and measure) performance, or what people do. Then, on the basis of what has been observed, they make inferences about the learning that may have taken place. Similarly, happiness, anger and frustration may be studied by analysing behaviour that accompanies them.

The term **behaviour** generally refers to any action a person (or animal) uses to adjust to the environment. Behaviour involves doing something and is therefore an active process. It is the means by which we can physically express our thoughts and feelings when interacting with the environment. It is also the means by which we (and other organisms) adapt to the environment. Talking, touching, running, perspiring, hugging, flirting, texting, Skyping, watching television, sleeping, socialising and reflexive responses such as blinking and automatically withdrawing your hand if you touch a very hot object are all examples of behaviour that is externally expressed and can therefore be directly observed as it occurs.

Human actions, however, are not necessarily limited to directly observable ones. Behaviour also includes actions through which we adjust to the environment but that cannot be readily observed, such as heart beat, breathing and glandular responses. These activities are hidden from view and can only be detected by inference or through the use of specialised devices and procedures. Consequently, behaviour may be referred to as involving either overt or covert actions and some psychologists may therefore also distinguish between *overt behaviour* and *covert behaviour*.

Although the definition of psychology distinguishes between mental processes and behaviour, these do not often occur independently of one another. They are interrelated and constantly interact. For example, how you overtly behave in a particular situation will be accompanied by underlying thoughts and feelings about that situation. Similarly, your thoughts and feelings about a situation can influence how you behave and the extent to which you will sustain that behaviour. For example, if you think you are being 'used' by one of your friends and feel angry about it, you may refuse to give them your class notes the next time they ask for them. Mental processes and behaviour are only considered separately for the purposes of scientific study, in courses such as VCE Psychology, and in textbooks such as this.

CORRECTED PAGE PROOFS

Figure 1.2 Psychologists may distinguish behaviour as (a) directly observable and therefore overt or (b) hidden from view and therefore covert.

PSYCHOLOGY AS A SCIENTIFIC STUDY

Psychologists attempt to understand mental processes and behaviour by using the highly disciplined approach and methods of science to systematically collect, analyse and interpret data. This typically involves experimental research using carefully controlled observations and measurements. The reliance on a scientific approach and use of scientific research methods sets psychology apart from non-scientific disciplines such as history and politics.

Sciences share a common approach to studying their respective subject matter. For example, physics, chemistry, biology, environmental science and psychology differ in what they study, yet each discipline uses a scientific approach and method to achieve common goals of description, explanation, prediction and change, the latter by applying knowledge to prevent an unwanted outcome or bring about a beneficial outcome. While an environmental scientist might pursue these four goals in studying the effects of greenhouse gases on global warming and climate change, a psychologist might pursue them in studying thoughts, feelings and behaviour during sleep.

Table 1.1 shows the key science skills specified for study in VCE Psychology. You will be given the opportunity to develop then demonstrate these skills at a progressively higher level throughout Units 3 and 4.

TABLE 1.1 VCE Psychology Units 1–4 key science skills

Key science skill	VCE Psychology Units 1–4 skills
Develop aims and questions, formulate hypotheses and make predictions	• determine aims, research hypotheses, questions and predictions that can be tested • identify and operationalise independent and dependent variables
Plan and undertake investigations	• determine appropriate type of investigation: experiments (including use of control and experimental groups); case studies; observational studies; self-reports; questionnaires; interviews; rating scales; access secondary data, including data sourced through the internet that would otherwise be difficult to source as raw or primary data through fieldwork, a laboratory or a classroom • use an appropriate experimental research design including independent groups, matched participants, repeated measures and cross-sectional studies • select and use equipment, materials and procedures appropriate to the investigation • minimise confounding and extraneous variables by considering type of sampling procedures, type of experiment, counterbalancing, single and double blind procedures, placebos, and standardised instructions and procedures • select appropriate sampling procedures for selection and allocation of participants including random sampling, stratified sampling, convenience sampling and random allocation of participants to groups
Comply with safety and ethical guidelines	• understand the role of ethics committees in approving research • apply ethical principles when undertaking and reporting investigations, including consideration of the role of the experimenter, protection and security of participants' information, confidentiality, voluntary participation, withdrawal rights, informed consent procedures, use of deception in research, debriefing and use of animals in research • apply relevant occupational health and safety guidelines while undertaking practical investigations
Conduct investigations to collect and record data	• work independently and collaboratively as appropriate and within identified research constraints • systematically generate, collect, record and summarise both qualitative and quantitative data
Analyse and evaluate data, methods and scientific models	• process quantitative data using appropriate mathematical relationships and units • organise, present and interpret data using tables, bar charts, line graphs, percentages, calculations of mean as a measure of central tendency and understanding of standard deviation as a measure of variation around the mean • recognise the difference between statistics that describe a specific sample and the use of statistics to make inferences about the population from which the data were drawn • use basic principles of reliability and validity in evaluating research investigations undertaken • explain the merit of replicating procedures and the effects of sample sizes in obtaining reliable data • evaluate investigative procedures and possible sources of bias, and suggest improvements, with reference to identification of potential extraneous and confounding variables including individual participant differences, non-standardised instructions and procedures, order effects, experimenter effect and placebo effects • explain how models are used to organise and understand observed phenomena and concepts related to psychology, identifying limitations of the models • distinguish between scientific and non-scientific ideas

UNCORRECTED PAGE PROOFS

Key science skill	VCE Psychology Units 1–4 skills
Draw evidence-based conclusions	• determine to what extent evidence from an investigation supports the purpose of the investigation, and make recommendations, as appropriate, for modifying or extending the investigation • draw conclusions consistent with evidence and relevant to the question under investigation • identify, describe and explain the limitations of conclusions, including identification of further evidence required • critically evaluate various types of information related to psychology from journal articles, mass media and opinions presented in the public domain • discuss the implications of research findings and proposals
Communicate and explain scientific ideas	• use appropriate psychological terminology, representations and conventions for reporting research, including standard abbreviations, graphing conventions and the components of a scientific report with reference to inclusion of an abstract, an introduction and sections for method, results and discussion • discuss relevant psychological information, ideas, concepts, theories and models and the connections between them • identify and explain formal psychological terminology about investigations and concepts • use clear, coherent and concise expression • acknowledge sources of information and use standard scientific referencing conventions

Source: © VCAA, *VCE Psychology Study Design* (June 2017 update), pp. 11–12.

UNCORRECTED PAGE PROOFS

STEPS IN PSYCHOLOGICAL RESEARCH

Most of what psychologists know about mental processes and behaviour comes from psychological research studies ('investigations') that have been conducted in a scientific way.

Scientific research involves using an appropriate research method to collect data ('evidence') relating to a question of interest, then analysing the data and drawing valid conclusions from the results in relation to the hypothesis that was tested. The research is based on scientific assumptions, principles and procedures, and is planned, conducted and reported in a systematic way. This approach is also called *scientific method*.

Generally, a systematic, step-by-step procedure is followed when undertaking scientific research. There is not a single, unique set of steps or sequence that everyone follows. Although there are variations to the names of the steps and the order in which they may be followed, most are like those described below and summarised in Figure 1.5 on page 11. The steps do not guarantee accurate and justifiable conclusions, but they help ensure appropriate data are collected and minimise the chance for bias, errors or faulty results and conclusions that cannot be tested through replication of the research.

Step 1 Identify the research topic

The first step involves identifying the specific topic of research interest. The topic that is identified is often called the research 'question' or 'problem'. For example, a researcher might be interested in ways of improving memory. To do this, they may conduct a 'literature review' to find reports of research already conducted on their topic. For example, research studies on memory improvement may include reports on experiments that have:

- tested specific techniques for improving memory (called 'mnemonics') such as 'narrative chaining' and 'acrostics'
- compared 'maintenance rehearsal' involving simple, continuous repetition of the information to be remembered with 'elaborative rehearsal', which involves linking new information with information already in memory.

A literature review enables the researcher to become more familiar with their topic of interest, to refine their ideas and to propose a relevant research question that will provide an aim and therefore help focus their research activities. For example, the researcher might decide to investigate the question 'Is narrative chaining a more effective technique for improving memory than acrostics?'

Step 2 Formulate the research hypothesis

When the topic has been identified and refined to the level of a question, the next step involves construction ('formulation') of an appropriate hypothesis for the research. This is essentially a thoughtfully considered prediction about the outcome when the hypothesis is tested; for example, in a memory improvement study, a prediction about whether the use of a specific memory improvement technique when learning new information will lead to an improvement in memory of that information (when measured using a test of recall).

The research hypothesis is formulated before the study is conducted. It must be testable and written

as a very specific statement. For example, a research hypothesis for the memory improvement study could be 'students who use narrative chaining to learn new information will have a better memory of that information than students who learn the information using acrostics'. Note that within the hypothesis is a possible explanation of the outcome — why some students will have a better memory of what they learn.

There are several other important characteristics of the research hypothesis. These are described on page 15.

Step 3 Design the research

The next step involves determining the type of evidence required to test the hypothesis then devising a plan on how this is best achieved. The overall plan or 'blueprint' that is formulated to collect relevant data is commonly referred to as the *research design*. Generally, the research design will address how, when and where the required data will be collected and analysed.

When designing the research, the researcher must select the most appropriate type of research method (such as an experiment) for testing their hypothesis, and devise the specific procedures that will be used to collect the required data. The researcher must take account of their hypothesis, identify relevant variables and how they will be measured, decide which participants will be studied, how many there will be and how they will be selected and allocated to different groups that may be used in the study. The participants' responses provide the data that become the results for the research.

For the memory improvement study, the researcher may decide to conduct an experiment using Year 6 students from a nearby primary school as participants. For example, if the researcher uses a type of experimental research design called 'independent groups', one-third of the participants could be taught the narrative chaining technique, then they would use it to learn new information (such as a shopping list); one-third would be taught to use acrostics; and the other third would be used for comparison and would therefore not be taught or use narrative chaining or acrostics. Narrative chaining involves story creation using key words and acrostics involve sentence creation using the first letters of key words. The two groups who use a technique for improving memory would be called 'experimental groups', and the comparison group would be called the 'control group'.

There are also organisational matters to attend to; for example, preparation of participant information and consent forms, ensuring availability of participants and a suitable venue at specific times, preparation of instructional and test materials, and submission of the research proposal for review and approval by a human research ethics committee.

When designing the research, it is vital that relevant ethical standards and guidelines that regulate psychological research are considered to help 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 110–16. For example, since children will be used in the memory improvement experiment, the researcher will be required to obtain informed written consent from the parent(s) or guardian(s) of all children who will be selected as participants.

UNCORRECTED PAGE PROOFS

Figure 1.3 An experimental research design may use primary school students to investigate the effectiveness of different memory improvement techniques.

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. Data may be 'primary' or 'secondary' and 'quantitative' or 'qualitative'. In the memory improvement experiment, the data will be primary (collected 'first hand') and quantitative (in the form of 'numbers' such as test scores). These different types of data are described on pages 89–91.

It is vital that data collection and all other aspects of the research are conducted in an objective way. *Objectivity* involves taking steps to prevent personal factors from influencing any aspect of the research (or its reporting). It requires that data are collected and recorded free of bias, prejudice and other personal factors that may distort it in an inappropriate way. For example, an objective description of an event would simply describe what happened. The description might offer suggestions or beliefs about the motives or emotions of the people involved, but these would be considered *subjective* (or 'personal interpretations'), rather than objective observations.

Step 5 Analyse the data

The data collected by the researcher are initially referred to as *raw data* because they have not been processed or analysed. The next step is therefore to objectively summarise, organise and represent the raw data in a logical and meaningful way to help determine whether the hypothesis is supported and to draw other 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 down to 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. For example, examination ('analysis') of the data may identify patterns, trends, or relationships that support interpretation and evaluation of the results and may subsequently also help explain the results.

Step 6 Interpret and evaluate the results

Once the data have been analysed, they need to be interpreted and critically evaluated. This includes drawing a conclusion(s) about what the results mean and identifying, describing and explaining limitations such as possible bias, errors and uncertainties in procedures and results.

The data may be considered as evidence and all conclusions must be evidence-based. In much the same way as a crime investigator evaluates evidence, the researcher must consider the strengths and potential limitations of their evidence to help ensure their conclusions are justifiable on the basis of evidence and ultimately both valid ('accurate') and reliable ('stable').

One type of conclusion relates directly to the research hypothesis. This involves a judgment on whether or not the results support it. The researcher may also consider how widely the results can be applied. Most psychological research usually involves hypothesis testing with a relatively small number of participants who have been selected from a bigger group of interest; for example, 9 male and 9 female year 6 students from a primary school, rather than all year 6 students at the school or all year 6 students in the area or in general.

Of particular interest to the researcher is whether the findings obtained from a relatively small number of cases, such as the primary school students who were participants in the experiment, can be extended to apply to the bigger group of year 6 students at the school where they were selected, and possibly to other groups and situations; for example, to all years 6 students, to younger and/or older students or people in general.

With reference to the memory improvement experiment, if the results were found to clearly indicate that narrative chaining was a more effective technique than acrostics, then the researcher would interpret the results as supporting the hypothesis. They would conclude that narrative chaining is a more effective strategy for improving memory of new information than acrostics. However, the results might also indicate that using acrostics is also an effective strategy for improving memory (when compared with using no strategy). If so, the researcher would make this conclusion as well.

When drawing conclusions, the researcher examines patterns and trends in the data and bases their judgments strictly on what the results show. Statistical techniques are used to support the process. Their conclusions are restricted to the available data. They would also seek to identify, comment on and take account of any limitations of their study when drawing conclusions. For example, limitations might include:

- participant characteristics (including abilities) or events during the experiment that may have influenced the results in an unwanted way
- that the results may apply only to specific kinds of information (e.g. grocery lists, word lists)
- that mnemonic devices may be more or less suitable for other age, social or cultural groups.

They would also seek to identify further evidence that may be required, including how this could be obtained if the study were to be replicated.

UNCORRECTED PAGE PROOFS

Step 7 Report the research and findings

The final step 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. It is only on the basis of these 'peer reviews' that the research may be accepted for publication. In addition, a poster report may also be prepared for display and discussion at a conference or meeting with other researchers.

The report prepared for publication in a journal follows a strict format and typically describes in detail all aspects of the research, including relevant background information, how the research was conducted, the results and findings, possible explanations of what was studied and the results obtained, any limitations which may have impacted on the results, comments on the validity and reliability of the research and its data (including reference to their repeatability and/or reproducibility), and a list of references used in preparing for the research 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 in areas of common interest and the way scientific progress is achieved. It also enables the general public to benefit from the findings of research.

Additionally, 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 results and to consider alternative conclusions that may be valid. It also enables replication by other researchers.

Replication refers to the reproducibility and repeatability of research and its results. This helps ensure the research procedures can be 'reproduced'

(a)

Examples of journals with research reports

- *American Journal of Psychology*
- *American Psychologist*
- *Behavioral and Brain Sciences*
- *Behavioral Neuroscience*
- *Brain Research*
- *British Journal of Psychology*
- *Developmental Psychology*
- *Health Psychology*
- *Journal of Abnormal Psychology*
- *Journal of Adolescence*
- *Journal of Affective Disorders*
- *Journal of Applied Psychology*
- *Journal of Comparative Psychology*
- *Journal of Consulting and Clinical Psychology*
- *Journal of Cross-Cultural Psychology*
- *Journal of Educational Psychology*
- *Journal of Experimental Psychology*
- *Journal of Personality*
- *Journal of Sleep Research*
- *Legal and Criminological Psychology*
- *Neuropathology and Applied Neurobiology*
- *Organizational Psychology*
- *Psychological Bulletin*
- *Psychological Science*
- *Psychotherapy*
- *Schizophrenia Bulletin*

(b)

Figure 1.4 (a) Some of the international journals in which psychological researchers from throughout the world seek to publish their research reports. Note that many are based on sub-disciplines or topic areas of Psychology. (b) One of the Australian journals in which psychologists may publish their research reports.

eBook*plus*

Weblinks
View examples of journal articles

UNCORRECTED PROOFS

under different conditions and that the data and conclusions are not one-off but are 'repeatable' if the research is conducted again under the same conditions (including the type of research method or design, the specific procedures, measures, materials and so on). When replicated, there should be close agreement with the original results. Replication also enables researchers to test the applicability or relevance of the research or results to other groups or situations. Overall, replication is an important test of the reliability of research. *Reliability* is examined in more detail on pages 106–9, as are *reproducability* and *repeatability*.

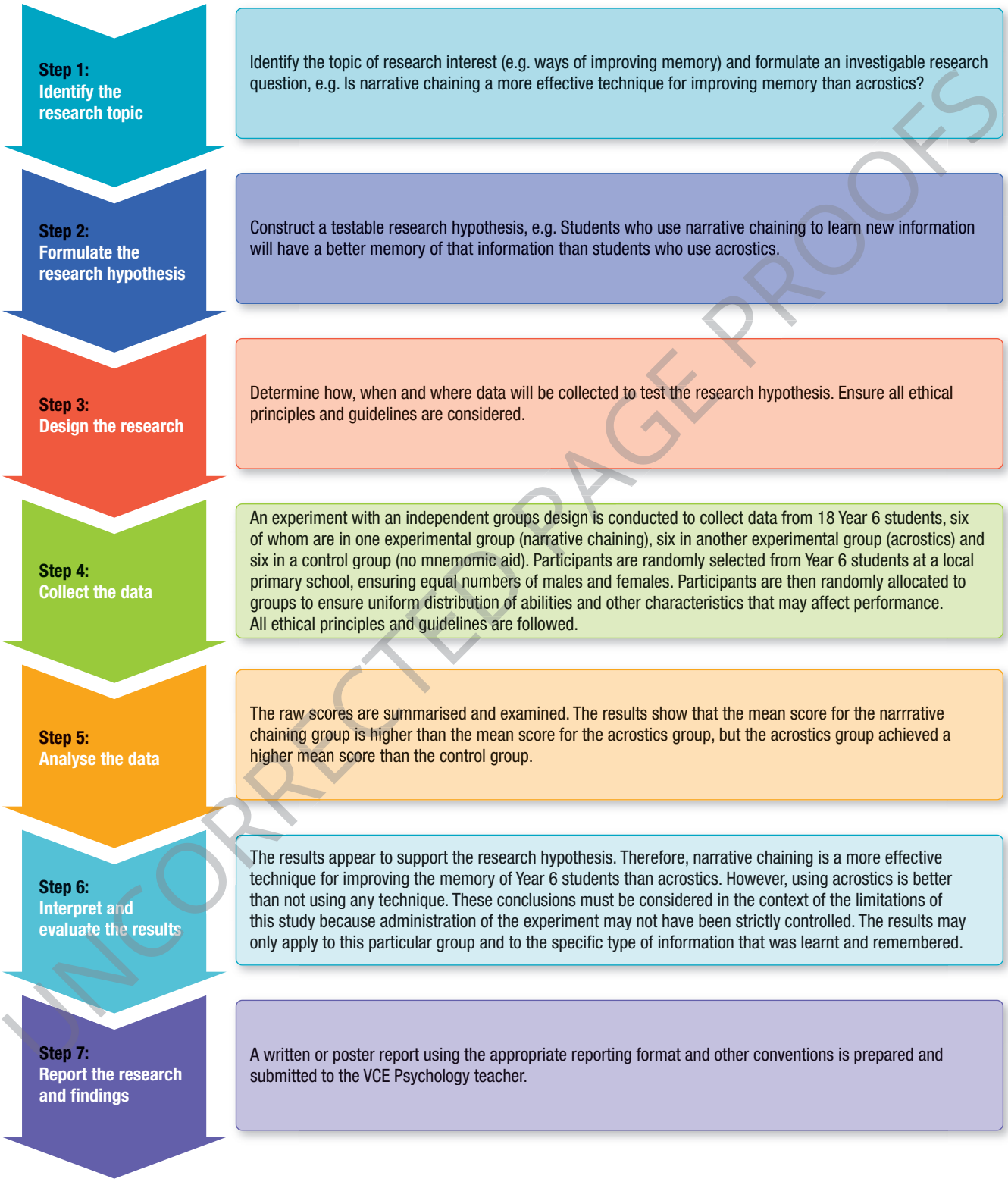


Figure 1.5 Flow chart of steps in psychological research. There are variations of these steps, but all are based on scientific principles and practices. The flow chart summarises the memory improvement study as if it were completed for VCE Psychology.

UNCORRECTED PAGE PROOFS

BOX 1.1 Designing and conducting experimental research for a practical investigation

The Unit 4 practical investigation SAC requires you to design (or adapt an existing design) and conduct your own experimental research on 'mental processes and psychological functioning'. This may be completed during Unit 3, Unit 4 or across both units. Although there is considerable scope for topic choice, you must use an appropriate experimental research design to collect primary data (see pages 58–60) and present the report using the VCAA poster format (see pages 122–3).

The following guidelines for the practical investigation are based on and summarise the 'Steps in psychological research' described previously. They assume you will work independently and can help ensure you successfully plan and conduct your investigation. The rest of this chapter includes detailed data collection and analysis information relevant to these requirements. Your teacher may provide additional and/or alternative advice of importance, including information about the task, the process, authentication and assessment requirements.

Step 1 Identify the research topic
Identify a suitable topic for investigation. It is desirable that the topic be of personal interest as well as relevant. At this stage, the topic may be broad (e.g. memory).

Step 2 Read relevant literature
Review and become familiar with the area of study and topic you intend to investigate. Refer to VCAA specifications and advice for the investigation (e.g. in the Psychology study design and 'Advice for teachers' available online), the relevant chapter(s) in this textbook, other psychology textbooks, your class notes, credible websites or even journal articles from a library or the internet. This will assist you to refine your ideas. It is important that you focus on arriving at a *specific* topic in the form of a relevant research question that can actually be investigated (such as effectiveness of certain mnemonic devices), rather than a general or broad topic (such as memory). It is also important that you be practical and realistic about what is achievable when deciding on your research question.

Step 3 Identify the research question
Formulate a research question that can be investigated through controlled experimental research or a cross-sectional study (which is like an experiment but lacking its control). Note that primary data must be collected.

Step 4 Draft a research proposal
Draw up an initial plan for carrying out your investigation. Consider the following:

- formulation of your research hypothesis and how it could be tested with reference to the specific type of experimental design, the independent and dependent variables, and how these variables will be operationalised
- variables other than the independent and dependent variables that may have an unwanted influence on the results and therefore need to be controlled
- relevant ethical considerations and safety issues
- who the participants will be, their availability, time requirements for key stages and procedures, and materials/equipment requirements
- the type of primary data that will be collected (i.e. quantitative and/or qualitative) and how it will be summarised and interpreted
- any test or other measure used to collect data, e.g. list of words for a test of recall
- potential limitations of the research methodology and data and the extent to which these may impact on the conduct of the investigation and its results.

Submit your proposal to your teacher. Regardless of how careful you are, your teacher may find flaws in the research design and oversights in ethical or safety considerations that you, the original designer (or adapter), being so close to the research, might not notice. Your teacher will also check that the investigation can meet all SAC requirements. Furthermore, according to the VCE Psychology study design (pp. 8–9), your teacher is responsible for ensuring that your research is conducted safely and ethically.

Step 5 Refine the research design
Correct any problems with your initial proposal. Ensure your research hypothesis is clearly stated and testable using your proposed research design. Your choice of sample and the control of variables are important considerations. The sample size should be manageable. Prepare a revised research proposal.

Step 6 Obtain approval to conduct the investigation
Submit your research proposal to your teacher to obtain approval to conduct the investigation.

Step 7 Organise materials and arrangements for the investigation
Ensure compliance with relevant participant consent requirements. Schedule and confirm availability of participants, the research setting or venue, materials, equipment and any other relevant resources.

Step 8 Collect the data
Conduct the investigation to test your hypothesis. Ensure that you collect the data objectively and follow all relevant ethical and safety guidelines.

Step 9 Present the results
Summarise the raw data, organise it and present it in an appropriate form using relevant statistical measures such as percentages and means. Data is usually presented in tables and figures such as graphs and charts.

Step 10 Analyse the results
Evaluate and interpret the summarised data in relation to your research hypothesis. Identify patterns and relationships, including sources of error and limitations of the data, your research design and how you implemented the design.

UNCORRECTED PAGE PROOFS

Step 11 Draw a conclusion(s)
Having interpreted the data, state the major finding(s) of the investigation. Make a statement about whether the results support or do not support your hypothesis. Review your research design and comment on any limitations of the investigation (including potential influences of unwanted extraneous or confounding variables), ways to improve the design and the possible wider application (generalisation) of the results.

Step 12 Present a report
Finally, prepare a report using the poster format described on pages 122–3. Your teacher will advise you about the template that can be used to guide the format and presentation, as well as the degree of detail required. Keep in mind that your report will be assessed, and ensure that your report will meet all the assessment criteria.

eBook*plus*

Weblink
VCAA practical investigation specifications

RESEARCH METHODS

Psychologists can choose from a wide variety of research methods to scientifically collect data in order to test hypotheses and answer questions on topics of interest. A *research method* is a particular way of conducting an investigation to collect accurate and reliable information about mental processes and behaviour. For example, an experiment and an observational study are different research methods, each of which uses a specific set of procedures for data collection. Although they share some procedures, they are clearly distinguishable research methods.

Each type of research method has its specific purposes, procedures, advantages and limitations. The researcher's choice depends on which method is most appropriate for the specific topic of research interest and hypothesis being tested. This is not unlike the choice a carpenter makes of which tool to use. Whether a hammer or drill is selected depends on the work to be done. Each tool has a specific use and way of being used. Similarly, each research method has a particular logic underlying its use and how it is used. In some cases it may be appropriate to use a combination of research methods to investigate and collect data. For example, a researcher conducting an experiment on different memory aids used by VCE students when studying for an exam may also conduct a survey to find out what motivates them to study for an exam, where they study, when they study and how much time they spend.

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 these methods.

UNCORRECTED PAGE PROOFS

Figure 1.6 A carpenter chooses tools depending on the work to be done, just as the researcher chooses the most appropriate research method to conduct an investigation on their topic of research interest.

Sample and population

Psychologists mostly conduct research with people. Participants are the people who take part in the research. The responses (or 'reactions') of the participants provide the data and the results for the research. The participants who are selected and used in the research are said to form the sample. A **sample** is a group of participants that is a subset or part of a larger group chosen to be studied for research purposes.

The term **population** is used in psychological research to describe the larger group from which a sample is selected and to which the researcher will seek to apply (generalise) the results. In scientific research, the term 'population' does not refer to all the people in a particular country or to the whole human race. It refers to a particular group or an entire set of individuals who has one or more characteristics in common; for example, all VCE Psychology students enrolled at a particular school, all VCE students in all schools, all females, all females who have been diagnosed with schizophrenia and are patients in a hospital, all left-handed males, all registered nurses aged 25 to 30 years, all cigarette smokers, all twins, all four-year-old twins, or all four-year-old identical twins born at a particular hospital.

UNCORRECTED PROOFS

Figure 1.7 A sample is a subset of a population selected as participants for a research study. The population is the entire set of individuals to which generalisations will be made based on the sample.

The population used for research can also be a measurement; for example, all EEG (brain wave) recordings for an individual during a certain period of time, the IQ scores of all students in a particular school, all drug-related deaths reported by the coroner in the previous 12 months, all absences from a workplace in a ten-year period due to a stress-related problem, all the days on which the temperature exceeded 30°C, all the words in the English language, all public hospitals in Victoria providing adolescent mental health services, or all of any other specified data.

A sample is always a part of the population. Therefore, it is always smaller than a population. When studying people, psychologists can rarely be certain about any mental process or behaviour that occurs in a population because they can rarely study all its members — it's usually too large a group. Consequently, researchers draw a sample that is appropriate for testing their hypothesis and attempt to generalise the results obtained for the sample to the population from which it is drawn. This is why it is important that the sample accurately reflect, or be representative of, the entire population of interest, although this is not always possible.

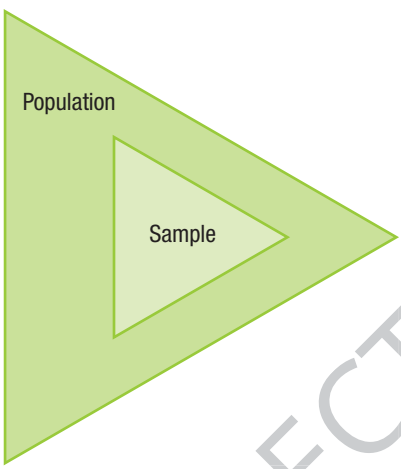


Figure 1.8 If a sample accurately reflects its population, the researcher is more able to generalise the results for the sample to its population.

Research hypothesis

In psychology, as in other sciences, a research study is designed to test one or more hypotheses relevant to the specific topic or question of research interest. A **research hypothesis** is a testable prediction about the relationship between two or more variables (i.e. events or characteristics). It typically states the existence of a relationship between the variables of interest, the expected relationship between them and a possible explanation of the results. For example:

- Watching violent television programs will lead to an increase in aggressive behaviour in young children.
- VCE students perform better on exams when they study on a regular basis.
- Changing channels occurs more frequently when watching television alone than when watching with others.
- Mobile phone conversation while driving, whether hands-on or hands-off, impairs driving performance.

The research hypothesis is often described as an 'educated guess'. This is because it is based on what is already known about relevant behaviour and mental processes; for example, from:

- what has been learnt through prior experience, including observations in everyday life
- reading a psychological theory on the topic and/or
- findings from previous research studies.

The research hypothesis is formulated before the investigation is conducted and provides a very specific focus for the investigation and its report. A useful research hypothesis has the following characteristics:

- refers to events or characteristics that can be observed and measured and is therefore testable (e.g. a hypothesis that people see a bright light when they die is not testable because it is not possible to observe or measure human experiences after death)
- states the existence of a relationship between two or more variables (e.g. a change in one variable is caused by the presence of the other variable)
- states the expected relationship between the variables, sometimes referred to as the 'direction' of the relationship (e.g. how one variable will influence the other)
- explains the prediction or expected results (e.g. an increase in aggressive behaviour will be attributable to watching violent television programs)
- 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 specific population from which the sample was drawn and therefore the larger group about which the researcher intends to draw conclusions. The population, however, is most commonly described in the introduction to the report on the investigation.

There is no preferred writing style for a research hypothesis. Different writing styles can be equally valid. For example, some hypotheses use an 'if-then' style, such as '*if* a certain event occurs, *then* it will cause a certain response'. In relation to children's television viewing, this may be stated as: 'If children watch violent programs on television, then they will behave more aggressively'.

It is not always possible to be entirely certain about the accuracy of the prediction within a hypothesis, which is why it is often described as a *tentative* prediction or statement. This is mainly because the researcher does not necessarily know or can control the influence of the many different variables that may affect a person at any given time. Nonetheless, most researchers would probably consider it pointless to conduct an investigation when the outcome is certain.

UNCORRECTED PAGE PROOFS

Scientific predictions tend to be more accurate about a large group or people in general than about a specific person. For example, a car insurance company can more accurately predict the percentage of people in a particular age group who are likely to be involved in road accidents this year than it can predict whether any particular individual in that age group will have an accident. Similarly, a psychologist may be able to correctly predict that cigarette smokers will be more *likely* to suffer a heart attack, but they cannot predict with certainty whether a particular cigarette smoker will suffer a heart attack.

This situation is no different in other sciences, which can only make predictions with varying degrees of probability of being correct. For example, your doctor may prescribe an antibiotic that, based on medical research, is *usually* effective in treating pneumonia. Your doctor, however, cannot guarantee that it will cure *your* pneumonia. Similarly, seismologists know that cities lying along geological faults are more likely to experience earthquakes, but they cannot accurately predict the day, or even the year, when one of these cities will experience its next major earthquake (Sdorow, 1995).

Figure 1.9 Scientific predictions tend to be more accurate about a large group or people in general than about a specific person. For example, a psychologist may be able to correctly predict that cigarette smokers will be more *likely* to suffer a heart attack, but they cannot predict with certainty whether a particular cigarette smoker will suffer a heart attack.

Theory and model

A research hypothesis is different from a theory and model — it is a specific prediction that guides the collection, analysis, interpretation and evaluation of data to test it. In contrast, a **theory** or **model** is a general explanation of a set of observations or findings about behaviour and/or mental processes that seem to be related. Both are used to organise and understand observed phenomena ('events') and concepts ('ideas') that are relevant to some aspect of psychological functioning.

There are examples of theories and models throughout this text, as they are an important part of your study of psychology. These include:

- models of stress as a biological or psychological process
- models to explain learning
- Atkinson-Shiffrin multi-store model of memory
- theories of the purpose and function of sleep
- 4P factor model of mental health
- biopsychosocial model of mental health
- transtheoretical model of behaviour change.

The term model is used interchangeably with theory although many models tend to focus more on representing *how* some behaviour and/or mental process could, should or does occur. For example, a model is often supported by one or more diagrams with boxes and arrows to organise and show relationships between different concepts.

Theories and models vary in scope, complexity and detail. All have one or more advantages and limitations. 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 learning, remembering, personality, mental health, stress-related effects, phobia acquisition, child development, problem-solving, creative thinking 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, revised or rejected by further scientific research. Therefore, theories and models tend to not 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 either right nor wrong. Instead, it is simply regarded as more or less useful.

UNCORRECTED PAGE PROOFS

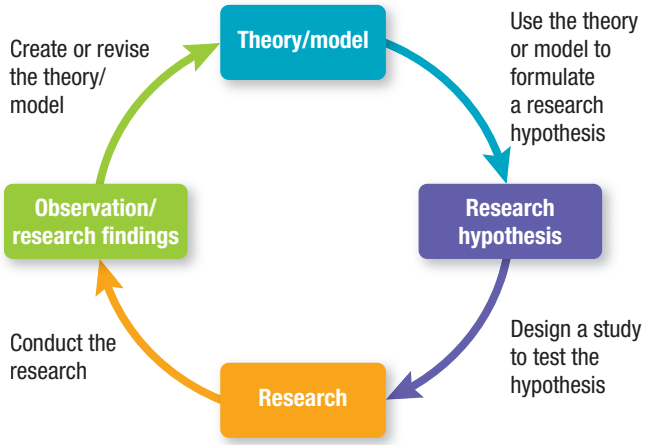


Figure 1.10 Theories and models are expanded and revised to reflect relevant research findings. New or revised theories and models lead to new observations or questions that stimulate new research.

LEARNING ACTIVITY 1.1

Review questions

1. (a) What is the difference between a sample and population in relation to research?
(b) Why is it important that a sample and its population are as alike as possible?

2. (a) Explain the meaning of hypothesis when used in a research study (investigation).
(b) Distinguish between a research hypothesis and a theory or model, with reference to an example of a psychological theory and/or model.
(c) Why can a research hypothesis be described as an 'educated guess'?

3. Construct a two-column table. In one column, list the characteristics of a useful research hypothesis. In the other column, summarise each characteristic using not more than three words.

4. Explain two possible limitations of the following question if it were to be used as a research hypothesis: *Do some people have an extrasensory perceptual ability to send and receive mental messages?*

5. Formulate a research hypothesis for each of the following topics or questions. Two of the hypotheses should be written using the *if-then* style. Note an example of a construction process for if-then hypotheses in the VCE Psychology *Advice for teachers* digital publication.
(a) Lack of attention causes forgetting.
(b) Crowding increases aggression.
(c) Positive thinking leads to success in a job interview.
(d) Does offering an incentive result in greater motivation to succeed?
(e) What is the effect of rote learning of information on a person's ability to recall the information when needed?
(f) Does being permitted to take a bottle of water into an exam improve performance on the exam?

eBook*plus*

Weblink
VCE Advice for teachers: Psychology

LEARNING ACTIVITY 1.2

Reflection

Someone has adopted a position that psychology is not a 'real' science like biology or physics. What two or three points could be made in support of psychology being a scientific discipline?

Experimental research

One of the most scientifically rigorous research methods used in psychology, as in other sciences, is the experimental method. This provides a means of collecting evidence to measure the effect of one variable on another. More specifically, an **experiment** is used to test a cause–effect relationship between variables under controlled conditions. For example, an experiment would be used to find out:

- if the variable of anxiety (a possible cause) has an effect on the variable of exam performance
- if the variable of playing music while studying (a possible cause) has an effect on the variable of learning
- if the variable of amount of sleep (a possible cause) has an effect on the variable of memory.

In any of these examples, a researcher (or 'experimenter') could conduct an experiment to find out if there is a causal (i.e. 'cause–effect') relationship between the different variables. When conducting an experiment, the researcher would do so in a planned, systematic way, controlling all variables that can influence the results, not just the independent variable (which is described below) but all other variables too. Generally, to 'control' variables means to regulate or 'keep constant' their influence (ideally minimising or eliminating), except those being systematically varied (i.e. the independent variable(s)).

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 the psychological experiment and why this particular research method can be used to investigate causes of behaviour and links between behaviour and mental processes.

Variables

In research, a **variable** is something that can change ('vary') in amount or type and can be measured. For example, sleep can change in both amount (e.g. number of hours) and type (e.g. with or without rapid eye movements) and is measurable (e.g. through recordings of bodily activities such as eye movements and brain wave patterns).

There is a virtually endless list of variables that may be studied in psychology. Examples include age, sex, intelligence, mood, problem solving, memory, state of consciousness, sociability, use of social media,

UNCORRECTED PAGE PROOFS

diet, exercise, media violence, drug-taking, risk-taking, family environment, religion, culture, work space, crowding, number of errors, and time taken to perform a task.

A variable may be a personal characteristic, either physical or psychological, or an event that can have a specific influence on how an individual may think, feel or behave. Personal characteristics such as biological sex, blood type, genetic make-up and racial or ethnic background are all inborn and therefore 'fixed' and ordinarily unchanging within a person. However, in psychological research they are still considered 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. Although a researcher cannot actually change a participant's sex or blood type, they can 'manipulate' (or 'change') them by allocating males and females and/or people with different blood types to different groups used in their experiment in order to make comparisons. Similarly, age and intelligence are 'fixed'. However, the researcher can manipulate them by comparing naturally occurring variations; for example, the performance of old and young participants, or, more or less intelligent participants (as indicated by IQ scores).

Independent and dependent variables

Every experiment includes at least one independent variable and one dependent variable. Whatever the type of true experimental design, the independent variable will be manipulated to determine the effect on the dependent variable.

The **independent variable (IV)** is the variable that is systematically manipulated or changed in some way by the researcher in order to measure its effect on the dependent variable. For example, in an experiment on whether watching a violent television program increases aggressive behaviour, the IV will be exposure to a violent television program. Its 'manipulation' will involve exposing or not exposing participants to violence in a television program to observe the consequential changes in the dependent variable (aggressive behaviour). The IV is sometimes referred to as the 'treatment' variable or condition to which participants in an experiment may be exposed. The independent variable is assumed to have a direct effect on the dependent variable so it is also assumed that any measurable change in the dependent variable will be due to the effect of the independent variable.

As indicated in the example above, a 'treatment' may also be withheld; for example, to compare its effect(s) on participants who are exposed to it with those who are not. This is why an IV may also be referred to as having different values (or levels) — the value of the IV is systematically 'manipulated'. In the simplest type of experiment, the IV has two values, such as exposure or non-exposure to violence in a television program.

More complex experiments have three or more values of the IV; for example, non-exposure to a violent TV program, exposure to one violent TV program and exposure to two violent TV programs. In this case, the IV has been manipulated 'quantitatively' by varying its 'amount' — exposure to one or two violent TV programs. Alternatively, the IV may be manipulated 'qualitatively' by varying its 'type'. In this case, the values of the IV may involve exposure or non-exposure to specific types of violent TV programs; for example, to compare the influence of violence by cartoon characters with violence by people in a movie.

An even more complex experiment may compare responses of male and female participants to various types of violent and non-violent male and female characters or people in different types of TV programs (qualitative variables) following different periods of exposure, such as 15 minutes, 30 minutes, 45 minutes, 1 hour, and so on (quantitative variables).

An experiment can also have more than one IV. For example, a researcher might test a hypothesis that a child will behave aggressively after watching a violent TV program only if other children are present. In this case, both the violent TV program and the presence of other children would be IVs. A third IV could be drinking a high sugar content cordial during the program. Of course, one or more of these IVs could also have different values.

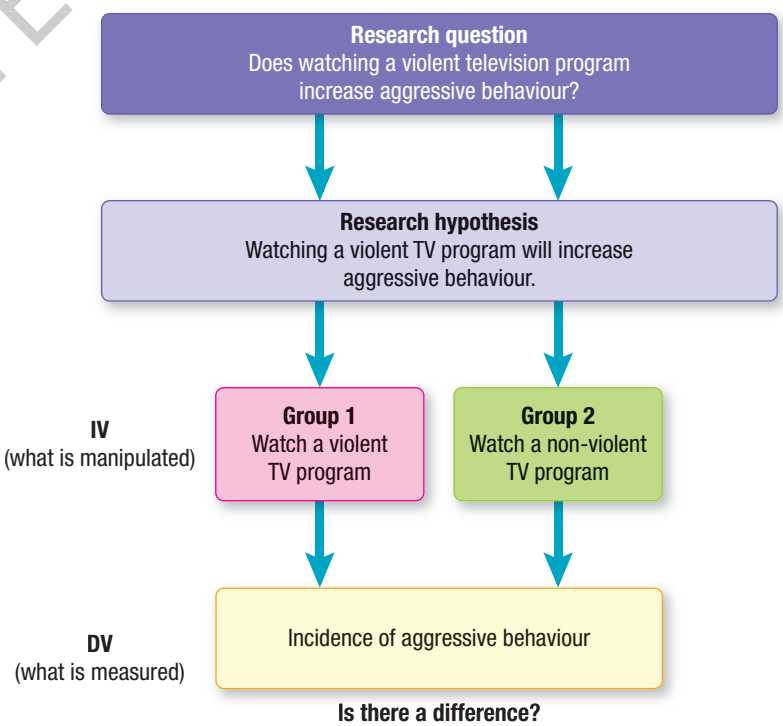


Figure 1.11 The IV and DV in an experiment on whether watching a violent television program increases aggressive behaviour

UNCORRECTED PAGE PROOFS

The **dependent variable (DV)** is the variable that is used to observe and measure the effects of the independent variable. It is the aspect of a participant's behaviour or experience that is assumed and expected to change as a result of the manipulation of the independent variable. The DV is often the responses made by participants and usually has a numerical (quantitative) value. For example, a behaviour such as aggression in young children might be observed and measured by the number of times physical contact is made with another person in a 5-minute period immediately after a child has been exposed to a violent or a non-violent TV program. Aggressive behaviour is the dependent variable, because the participants' responses are believed to be influenced by, or 'dependent on', the effects of the independent variable. It is sometimes referred to as the 'measurement' variable, because it provides a 'measure' of the participants' responses to the independent variable.

In the examples of possible experiments given earlier in this section on page 17, anxiety would be the IV and exam performance the DV; playing music while studying would be the IV and learning the DV; and amount of sleep would be the IV and memory the DV.

In terms of a cause–effect relationship, the independent variable is viewed as the possible *cause*, and change in the dependent variable is the possible *effect*. In experimental research, the research hypothesis states the causal relationship between the independent and dependent variables to be tested; that is, that the IV(s) will cause the DV(s) to change in a particular way. When the results of an experiment are shown on a graph, the IV(s) is plotted on the horizontal axis (e.g. Hours spent studying) and the DV(s) on the vertical axis (e.g. Score on a Psychology Unit Test).

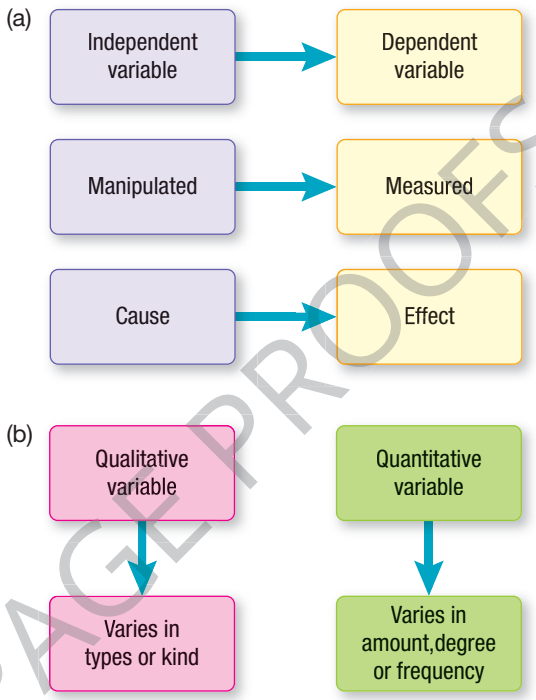


Figure 1.12 Distinguishing between (a) an IV and DV, and (b) a qualitative variable and quantitative variable

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.3

Review questions

1. What is an experiment?

2. Explain the meaning of control in relation to an experiment.

3. Write a definition of each of the following terms as they apply to an experiment:
- variable
- independent variable
- dependent variable
- quantitative variable
- qualitative variable

4. (a) What do researchers expect to happen to DVs when they manipulate IVs?
(b) What does 'manipulation' of an IV actually involve?
(c) Give an example of how each of the following could be manipulated quantitatively and qualitatively if it is an IV in an experiment:
(i) sleep
(ii) lighting
(iii) room temperature
(iv) a word
(v) body weight

5. Identify the IV and DV in each of the following if experimental research were to be conducted.
(a) Receiving a reward for studying will increase the amount of time students engage in studying.
(b) People who are in love perceive each other more positively than other people perceive them.
(c) Recall of information presented early in a list is better than recall of information presented later in a list.
(d) People react faster to sounds than to visual stimuli.
(e) Using adult language when talking to infants improves their vocabulary.
(f) People change their pitch of voice when lying.
(g) Daydreaming occurs more frequently during simple tasks than during complex tasks.
(h) Workers on an assembly line are more productive when working alone than in a small group.
(i) Heart rate and blood pressure increase when viewing a violent move clip as compared to a non-violent movie clip.
(j) Infants pay attention to a complex stimulus for a longer period of time when compared with a simple stimulus.

LEARNING ACTIVITY 1.4

Identifying IVs and DVs in experiments

A. A laboratory experiment

A researcher noticed that some of her laboratory rats stood on their hind legs for a moment whenever their food was brought into the laboratory. She decided to test whether she could teach the rats to stand on their hind legs when she rang a bell.

First she measured the exact amount of time the rats spent standing when the food was brought in. Then she rang a bell just before each meal. The rats eventually started to stand on their hind legs when they heard the bell.

1. What two IVs are being manipulated in the experiment?
2. What is the DV and how is it measured?
3. Why did the researcher measure the rats' movements before introducing the ringing bell?

B. A field experiment

A researcher observed helping behaviour in a real-world setting. An actor pretending to be either drunk or blind was required to collapse on a Melbourne underground train platform.

Sometimes the actor was an Aboriginal male and sometimes the actor was an Anglo-Saxon male of about the same age.

The researcher then recorded how long it took for help to be given.

1. What is the IV and how many values or levels does it have?
2. How is the IV manipulated?
3. What is the DV and how is it measured?

Operationalising independent and dependent variables

Operationalising the independent and dependent variables involves defining and explaining them in terms of the specific procedures ('operations') used to measure them in a particular experiment. Stating *how* the IV and DV will be defined and measured is an important step because many of the behaviours and mental processes psychologists investigate can have different meanings and can therefore be specifically defined and measured in more than one way.

For example, consider an experiment to investigate whether exercise provides relief from depression. 'Exercise', which is the IV, might be operationalised as 'walking at a particular pace for a specified period of time on an automated treadmill'. 'Depression', which is the DV to be measured, might be operationalised as the number of negative words used in writing a creative story, as it has been found through previous research studies to be related to the severity of depression. Similarly, consider the way in which each of the following potential variables that can have multiple meanings might be operationalised for the purpose of experimental research.

- intelligence — a score on a standardised intelligence test
- memory — a score on a test of free recall
- learning — the reduction in the number of errors when performing an unfamiliar task
- anger — changes in blood pressure, heart rate and respiration rate
- physical attraction — the number of times someone touches another person
- love — the frequency of expressions of affection such as kissing, touching and cuddling

Operationalising the IV(s) and DV(s) ensures that these variables are precisely defined and explained. The resulting definitions are sometimes called *operational definitions* (see Box 1.2 on the next page).

There are several important benefits of variables being defined precisely through operationalisation. These include:

- It helps ensure the independent and dependent variables are testable and therefore that the research hypothesis is testable.
- All researchers involved in conducting the experiment know exactly what is being observed and measured and how this will occur. This helps avoid experimenter biases and differences that can affect the results in an unwanted way.
- When the variables are defined in a very precise way, another researcher interested in the results, or perhaps even doubting them, will be more able to replicate 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 under the same conditions using a similar sample and similar results are obtained (i.e. the results are 'repeatable'), there is greater confidence in the validity of the results. There is even greater confidence when replication under different conditions achieves very similar results (i.e. the results are 'reproducible'). Alternatively, if replication of a study fails to produce the same basic findings, researchers have less confidence in the findings reported for the original research.

The research hypothesis for an experiment may refer to the operationalised variables (but this is not essential). For example, consider a possible hypothesis for the experiment on exercise and depression:

> *People with depression who exercise regularly will have fewer symptoms of depression than people who do not exercise.*

Note in this research hypothesis that:

- the IV is stated, including both its values i.e. regular exercise and no exercise
- the DV is clear i.e. number of symptoms of depression
- the expected effect of the IV on the DV is also stated, specifically, the direction of the predicted effect — that is, the way in which the two groups (exercise and no exercise) are predicted to differ (not simply that there would be a difference).

UNCORRECTED PAGE PROOFS

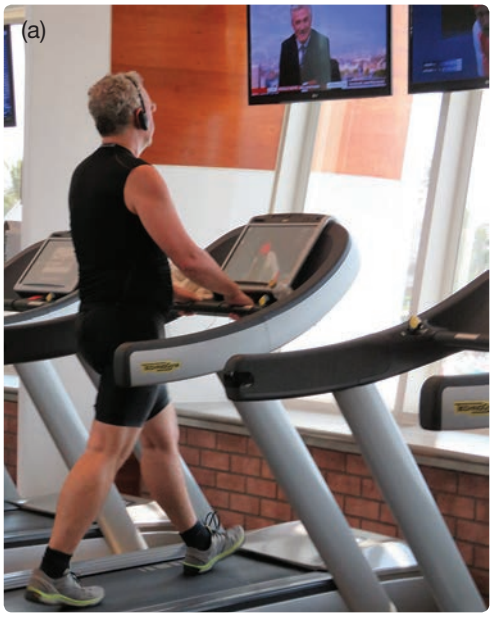


Figure 1.13 (a) Exercise might be defined as walking at 7 km/h for 30 minutes on an automated treadmill. (b) Is this cat attractive? The answer depends on how you operationalise attractive.

BOX 1.2 Operational definitions

In psychology, as in any other science, definitions of the concepts under investigation (and any other descriptions) must be clear and precise. Consider the statement 'the research is about crime'. 'Crime' could refer to riding a bicycle without an approved helmet, parking in front of a fire hydrant, shoplifting, assaulting someone, committing armed robbery, smuggling protected Australian birds out of the country, and so on. Like many of the words that we use in everyday conversation, the term 'crime' covers a broad range of behaviours and is therefore too inexact to use for research purposes.

Similarly, a term such as 'generous', while appropriate to use in everyday conversation, is too imprecise for research purposes because generosity can be demonstrated in many different ways, such as donating money to a charity, volunteering to coach a junior sports team, or spending time with a friend who is unwell or facing difficult times.

Researchers overcome this problem by defining their subject matter in terms of the way they observe or measure it — they define *what* they are measuring by describing precisely *how* they are measuring it.

The resulting definitions are called operational definitions. An *operational definition* is a definition of a variable, condition or some other observable event. It defines an observable event in terms of the specific procedures (or 'operations') used to measure that variable. Operationalisation of the IVs and DVs in an experiment essentially involves stating their operational definitions.

Consider, for example, a researcher who is interested in the conditions under which a rat turns left rather than right in a maze. It may seem relatively simple to determine the direction a rat turns in a maze. But what exactly is a *turn*? What will the researcher *observe*? How will the researcher *measure* the turn of a rat in a maze? Will the rat sticking its nose around the corner be considered a turn? What if it gets most of its body around the corner and then scoots back? Does the rat's tail have to make it all the way around? As basic as it may seem, the researcher would have to operationally define a *turn in a maze* by specifying exactly how it will be measured. For the purposes of this study, the operational definition of a *turn* might be 'when a rat's tail makes it all the way around a corner'. And, for the examples of *crime* and *generous* referred to previously, *crime* might be operationally defined as 'any act listed as a felony by Australian law', and *generous* might be operationally defined as 'donating more than 5% of one's annual salary to charity'.

Figure 1.14 A researcher will operationally define a turn in a maze by specifying exactly how it will be measured.

UNCORRECTED PAGE PROOFS

studyon

LEARNING ACTIVITY 1.5

Review questions

1. What does operationalisation of an experiment's independent and dependent variables involve?
2. List three potential benefits of operationalising variables.
3. Suggest how the IV and DV could be operationalised for an experiment on each of the following research hypotheses.
   (a) Anxiety causes forgetting.
   (b) Crowding increases aggression.
   (c) Relaxation minimises stress.
   (d) Practice assists learning.
   (e) Girls talk more than boys.
4. Suggest an operationalised IV and DV for each of the following research questions.
   (a) Does offering an incentive result in greater motivation to succeed?
   (b) What is the effect of rote learning of information on a person's ability to recall the information when needed?
   (c) Does being permitted to take a bottle of water into an exam improve performance in the exam?
   (d) Does parental attention increase the incidence of tantrum behaviour by toddlers?
   (e) Does sleep deprivation time cause an increase in reaction time when riding a bike?

Experimental groups and control groups

In a relatively simple experiment, the participants selected are allocated to one of two groups. One group of participants, called the **experimental group**, is exposed to the IV (i.e. the 'treatment') under investigation. This group may also be described as being in the **experimental condition**. A second group of participants, called the **control group**, is not exposed to the IV. This group may also be described as being in the **control condition**. This experiment is shown below.

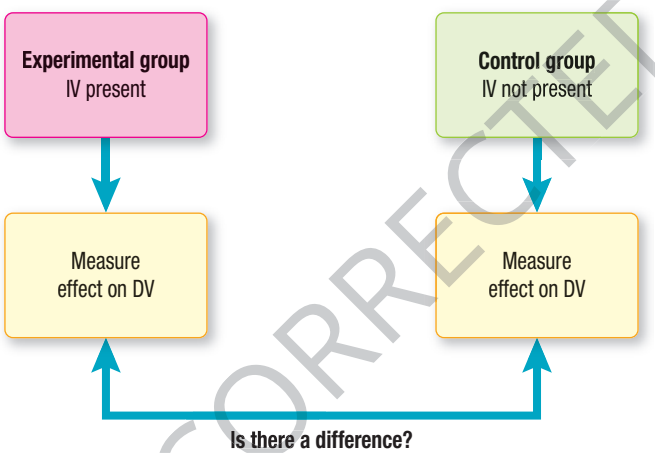


Figure 1.15 A simple experiment

For example, consider an experiment to test whether displaying posters of rock stars wearing a particular brand of jeans increases sales of that brand. The experimental group is exposed to a condition whereby posters showing rock stars wearing the jeans are displayed prominently in a jeans store, while the control group is not exposed to the posters. The DV might be the number of pairs of that brand of jeans sold (this experiment is shown on the next page).

The control group provides a standard or 'baseline' against which the performance of the experimental group can be compared to determine whether the IV has caused some change in, or affected in some way, the behaviour or event being measured (the DV). Without a control condition, it would not be possible to assess the influence of an IV; for example, whether the posters of rock stars wearing the jeans affected the number of pairs of jeans purchased.

If a significantly greater number of these jeans are purchased by participants in the experimental group, the experimenter may assume that the difference between the two groups was caused by the exposure of the experimental group to the posters of rock stars (IV). However, in order to make this assumption, the experimenter must be confident that no variable other than the IV being tested had an excessive influence on the purchase of jeans.

It is important that the experimental group and the control group are as similar as possible in personal characteristics that might cause a change in the DV. For example, one group should not have significantly more participants who have access to more spending money so that this doesn't become a possible reason for the difference in jeans purchased that may be recorded. It is also necessary to treat the two groups the same, except for exposure of the experimental group to the IV. For example, one group should not receive more or better quality customer service than the other. Both of these conditions are necessary so that if a large enough 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 most likely caused the change and not some other variable.

Some experiments do not have an experimental group and a control group with different participants. Instead, they have one group of participants who are exposed to *both* the control condition and the experimental condition. For example, to study the influence of rock music on people's concentration while driving, a group of participants could have their driving abilities tested in a simulator while no rock music was playing (control condition). The same group would later be tested again in the simulator

UNCORRECTED PAGE PROOFS

while there was rock music playing (experimental condition). The test results of the same group under the two different conditions would then be compared.

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 is often referred to as the 'treatment condition' because the IV is the 'treatment' to which the participants are exposed.

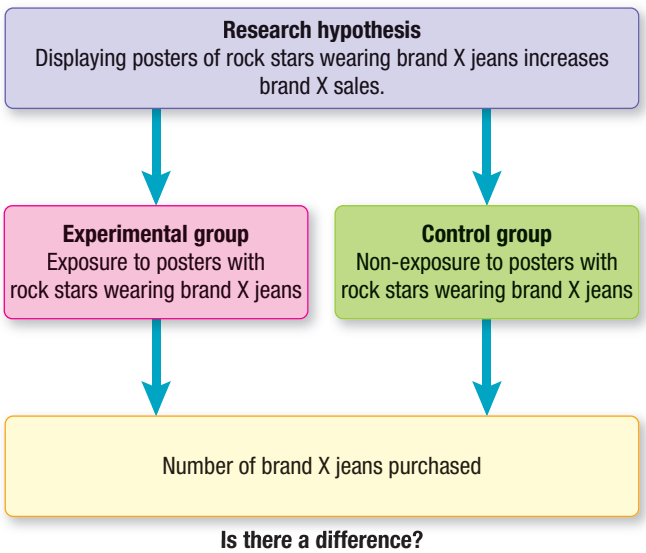


Figure 1.16 Experimental design to test whether displaying posters of rock stars wearing a particular brand of jeans increases sales of that brand

LEARNING ACTIVITY 1.6

Review questions

1. Distinguish between an experimental group and a control group in relation to the IV.
2. What is the purpose of using a control group in an experiment?
3. Why is it important for the experimental and control groups to be as similar as possible in personal characteristics that may affect the DV?
4. (a) In what other way(s) must the experimental and control groups (or conditions) be alike?
   (b) Why?
5. A researcher will conduct an experiment to find out whether people get a better score on a video game when cheered or jeered ('booed') by an audience of peers about the same age rather than when they play the game by themselves.
   (a) Suggest how the IVs and DVs could be operationalised for the experiment.
   (b) What are the experimental and control groups and how are their conditions different?
   (c) Suggest three variables other than the IV that have the potential to influence the DV and would need to be controlled so that the effects of the IV can be isolated.

Extraneous variables

In an experiment, the researcher predicts that manipulating the IV will cause a change in the DV. If a change in the DV is found, the researcher would like to conclude that the change is due *solely* to the influence of the IV. In order to draw such a conclusion, it must be shown that all other variables that could have influenced what is being measured (the DV) have been controlled, minimised or eliminated. Other, or 'extra', variables in an experiment are called extraneous variables.

An **extraneous variable** is any variable other than the IV that can cause a change in the DV and therefore affect the validity ('accuracy') of the results of the experiment in an unwanted way. Extraneous variables are 'unwanted' because they can make it difficult for the researcher 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.

Although extraneous variables are not intentionally studied (as are the IV and DV), the researcher tries to monitor and control or keep constant the influence of extraneous variables by using procedures that will minimise their influence to an acceptable level. Consequently, the researcher attempts to identify all relevant extraneous variables when designing their experiment. A 'relevant' extraneous variable is one that is believed to have the potential to cause a change in the DV.

For example, suppose a researcher will conduct an experiment to investigate age differences in navigating between different spatial locations. Participants in different age groups ranging from very young to very old will be required to find their way through a series of increasingly complex mazes. The maze navigation tasks will be presented online using a computer so that the time taken to successfully complete each maze can be electronically measured and recorded. The IV is age and the DV is navigational ability. In this study, extraneous variables that could impact on the DV and would therefore need to be controlled or eliminated include:

- biological sex of participants e.g. one sex may have inherently better ability for this kind of spatial task
- prior experience with computers and online task completion e.g. a large number of very young or very old participants may lack computer and online experience which may interfere with maze navigation
- prior experience with maze navigation e.g. to ensure no age group is advantaged
- motivation e.g. to ensure one age group is no more or less motivated to complete the maze tasks than the others

UNCORRECTED PAGE PROOFS

- instructions e.g. to ensure all clearly understand what needs to be done and that no age group is unduly advantaged or disadvantaged by the administration or comprehension of instructions
- test conditions e.g. to ensure they are the same for all participants.

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 can 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 (or at least monitoring) those variables that are likely to have a significant effect on the DV. For example, in an experiment on colour perception or discrimination between different shades of white, it would be very important to control light intensity. However, in an experiment to test the effect of flexibility on performance of some manually performed physical task, light intensity 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 can affect participants' responses (i.e. the DV), such as all those variables collectively referred to as extraneous variables.

Extraneous variables are inevitable and do not necessarily pose a problem if controlled in an appropriate way. By strictly controlling their unwanted effects 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, extraneous variable is commonly 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 predicted change in the DV. A confounding variable is not manipulated or controlled by the researcher (and therefore not intentionally studied). Nor is it considered to be some kind of 'random error'. It systematically varies together with the IV (or DV) so its effects are 'confounded', confused or mixed up with those of the IV. Because its effects are entangled with and cannot be separated from those of the IV, the researcher cannot conclude with any confidence that the predicted change in the DV was caused by the IV alone.

The presence of one or more confounding variables does not necessarily mean that the IV did *not* cause the change 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. For example, if there is a difference in the results for the experimental and control groups, it could be caused by the IV, the unwanted confounding variable or the combined effects of both. The more alternative explanations there are for the results, the less confident the researcher will be that the IV alone was responsible for the specific outcome.

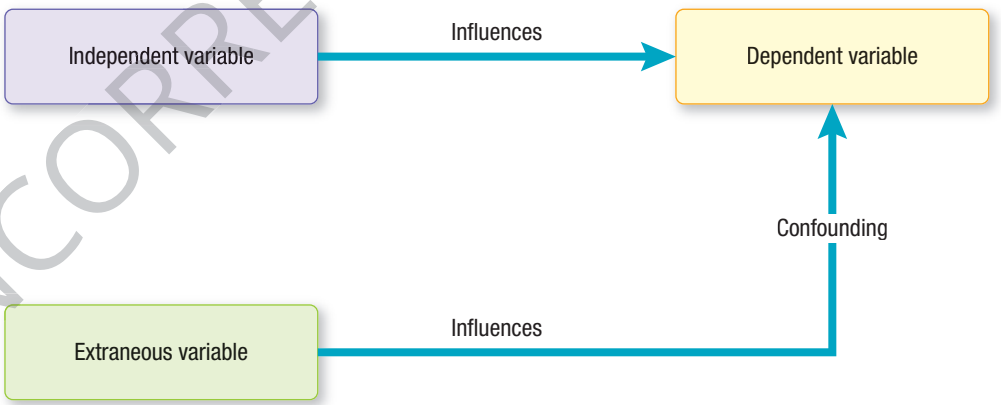


Figure 1.17 Confounding occurs when the effects of the IV on the DV cannot be separated from those of an extraneous variable. The change in the DV is consistent with what was predicted in the hypothesis. When this occurs, the variable will usually be referred to as a confounding variable or simply as a 'confound'. The presence of a confounding variable provides an alternative explanation of the experimental results and thereby also interferes with their validity.

UNCORRECTED PAGE PROOFS

A well-known example of the importance of controlling all variables in an experiment involves a taste test conducted by the Pepsi-Cola Company. Coca-Cola® drinkers were asked to taste two unidentified cola drinks and indicate which of the two they preferred. The drinks were Coca-Cola and Pepsi. The brand of cola was the IV, and the participants' taste preference was the DV. To prevent the participants from knowing which cola they were tasting, they were given Pepsi in a cup labelled 'M' and Coca-Cola in a cup labelled 'Q'. The results showed that most of the participants preferred Pepsi.

The Pepsi-Cola Company proudly advertised this as evidence that even Coca-Cola drinkers preferred Pepsi. But, to test the findings, the Coca-Cola Company replicated the experiment, this time filling both cups with Coca-Cola. The results showed that most of the participants still preferred the cola in the cup labelled 'M'. It seems that the Pepsi taste test had not demonstrated that Coca-Cola drinkers preferred Pepsi. It had demonstrated that Coca-Cola drinkers preferred the letter M to Q. The letters were an uncontrolled variable that had an unwanted effect on the DV (taste preference). Consequently, it remained unclear as to whether the IV (the kind of cola) or the unwanted variable (in this case the confounding variable of the letters) had affected the DV (taste preference).

A confounding variable is often described as a type of extraneous variable. A confounding variable may have its origin as an extraneous variable but there is an important distinction. 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 variables have in common is that they create problems for the researcher in isolating the real effect of the IV, moreso a confounding variable.

An experiment with one or more confounding variables compromises the interpretation of the results and the validity of the experiment, specifically its *internal validity* (which is examined on page 108). 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 think, feel or behave in any given situation, good experimental design involves anticipating potential extraneous and confounding variables and developing strategies to control and minimise their influence and ensure that extraneous variables do not become confounding variables.

UNCORRECTED PAGE PROOFS

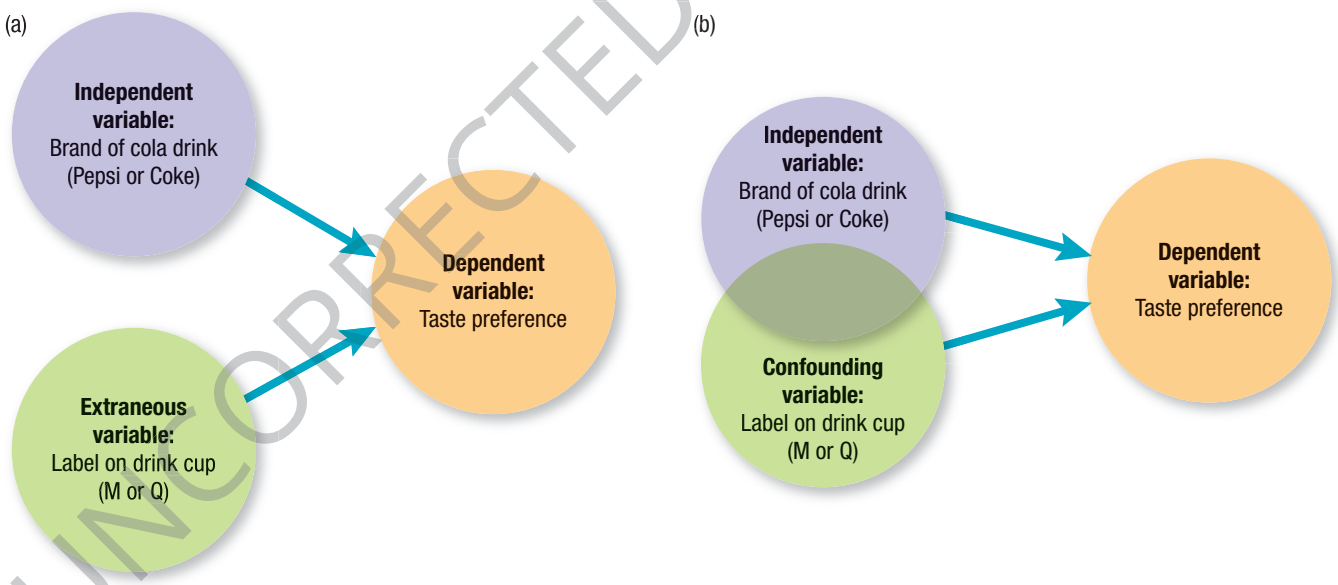


Figure 1.18 (a) In the taste test experiment, the label on the drink cup was an uncontrolled extraneous variable. (b) The effect of the label on the DV (taste preference) could not be isolated from that of the IV (brand of cola drink). The label on the drink cup had therefore become a confounding variable that provided an alternative explanation for the experimental results, making it impossible to be certain that the IV (rather than the confounding variable) caused the DV result.

Figure 1.19 (a) This participant has a high error rate in a driving simulator. (b) Is his poor performance due to his driving ability or because he went out with his friends and stayed up late the night before, slept poorly and felt excessively tired on awakening?

studyon

BOX 1.3 Other types of variables that apply to VCE Psychology

Type of variable	Definitions
Categorical	Categorical variables are qualitative variables that describe a quality or characteristic typically addressing 'what type?' or 'which category?' They are generally represented by non-numeric values and may be further classified as ordinal or nominal. **Ordinal variables** can take values that can be logically ordered or ranked, for example, birth order (1st, 2nd, 3rd), level of stress (low, medium, high) and attitudes (strongly agree, agree, disagree, strongly disagree). **Nominal variables** can take values that cannot be organised in a logical sequence, for example, gender, colour, taste (sweet, sour, bitter, salt, savoury) and type of sleep (REM, NREM). Bar charts and pie graphs are used to graph categorical data.
Numerical	Numerical variables are quantitative variables that describe a measurable quantity as a number, typically addressing 'how many?' or 'how much?' They are further classified as continuous or discrete. **Continuous variables** can take any value between a certain set of real numbers, for example, distance, height (2.85 metres), length of time (12.5 seconds) or temperature (25.4 °C). **Discrete variables** can take a value based on a count from a set of distinct whole values and cannot take the value of a fraction between one value and the next closest value, for example, number of neurons in a brain or number of facts recalled from a list. Scatter plots and line graphs are used to graph numerical data.

Source: VCE Advice for teachers: Psychology (2017). [VCAA digital document]. Retrieved May 2, 2018, from http://www.vcaa.vic.edu.au/Pages/vce/adviceforteachers/psychology/defining_variables.aspx

BOX 1.4 Experimental settings

Experiments can be conducted in a *laboratory setting* (called a *laboratory experiment*) or outside the laboratory in a *field setting* (called a *field experiment*).

A laboratory setting usually enables stricter control of variables but is sometimes criticised because of its artificiality, depending on what is studied. In a field setting, the conditions of the experiment are usually less strictly controlled, but it has the advantage of being able to make observations of participants' behaviour in a real world environment where their behaviour is likely to occur more naturally.

Some experiments conducted in field settings are called natural experiments. A *natural experiment* takes advantage of a naturally occurring event – the independent variable is naturally occurring and is not manipulated by the experimenter, often because it is impossible or unethical to do so. For example, a researcher may study the effect of a catastrophic event such as bushfire or flooding on stress-related ill health. In this case, the IV is a bushfire or flood, a naturally occurring event. However, the IV is not controlled by

UNCORRECTED PAGE PROOFS

the researcher and there may be many uncontrolled extraneous or confounding variables. For example, in a bushfire study, it may be difficult to determine whether stress-related ill health (the DV) is caused by fire, smoke or stress due to loss of one's house.

One of the best-known examples of a natural experiment was conducted by the Nobel Prize-winning American neuropsychologist Roger Sperry (1968) to study the functions of the brain's cerebral hemispheres when surgically separated. His participants were 11 patients who had already undergone 'split-brain surgery' to sever their corpus callosum in order to treat their debilitating epilepsy.

In a true experiment, the researcher can manipulate the IV, so a natural experiment is often described as a *quasi-experiment* because it looks like an experimental design ('quasi' means 'resemble') but the researcher does not actually manipulate anything and it lacks random allocation.

Figure 1.20 (a) A laboratory setting where social behaviour within a group of friends is observed in a controlled situation established by the researcher. (b) A field setting where the social behaviour of interest is observed in the real-world situation of an outdoor café area, but where less control of conditions is possible.

UNCORRECTED PAGE

LEARNING ACTIVITY 1.7

Review questions

1. (a) Define the meaning of the terms extraneous variable and confounding variable.
(b) When is the best time to identify these variables?

2. In what way are extraneous and confounding variables similar and different?

3. (a) Give two reasons to explain why it is essential to control extraneous and potential confounding variables in research.
(b) Sometimes researchers refer to the 'presence of a confound' in an experiment. What do you think this means?
(c) Similarly, a researcher may refer to a variable having 'confounded' the results or their interpretation. What do you think this means?

4. For each of the following experiments, identify the IV, DV and three potential extraneous or confounding variables.
(a) The reaction time of 20 people who have just awoken from a night's sleep is compared with that of a group who have just run a kilometre.
(b) The goal shooting accuracy of one group during a 10-minute period is measured when alone and compared with that of another group who goal shoot in the presence of others.
(c) Participants read a description of a person. All read the same description but half are told the person is of the same cultural background as themselves, whereas the other half are told the person has a different cultural background. All participants are then required to select characteristics they believe best describes the person; for example, 'good vs bad', 'warm vs cold' and 'friendly vs unfriendly'.

5. A researcher is planning an experiment to investigate the rate of forgetting (how much time it takes) and amount of forgetting (how much information) that occurs when new information (a list of nonsense 'words' such as qab and jir) is learned.
(a) Identify the operationalised independent and dependent variables.
(b) Suggest two extraneous or potential confounding variables that could affect the DV (in addition to the IV) and therefore need to be controlled.
(c) Suggest a way that each variable referred to in part (b) could be controlled.

6. An experiment was conducted to test whether people make fewer errors in detecting spelling errors in an interesting text than in a boring one. Two groups of randomly selected and allocated participants were used. Group 1 looked for errors in a physics text on string theory (a boring task) and Group 2 looked for errors in the script of a recently released blockbuster movie (an interesting task). The results showed that Group 1 detected significantly fewer spelling errors than did Group 2.
(a) Identify the operationalised independent and dependent variables.
(b) Suggest a potential confounding variable in the experiment. Explain your choice of confounding variable.

Identifying extraneous and potential confounding variables

Researchers have described different types of variables that can be extraneous or potential 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 on a DV.

Individual participant differences

The unique combination of personal characteristics, abilities and backgrounds each participant brings to an experiment (or any other research study) are commonly referred to as **individual participant differences**. These *participant variables*, as they are sometimes called, make one individual different from another, are expected by the researcher and may be biological, psychological or social in nature. They include age, gender, athletic ability, intelligence, personality, memory, educational background, family environment, social relationships, work experience, ethnicity, cultural background, religious beliefs, motivation, emotional state, mood, problem-solving ability, self-esteem, social skills, physical health, mental health, strength, eye–hand coordination, prior experience with materials or tasks to which they will be exposed in the experiment, and so on.

Each of these variables, and many other specific participant characteristics (including abilities) and prior experiences, can affect how participants respond in an experiment. For example, mood may affect participants' responses and make them more or less reactive to the experimental procedures. Some participants may be more or less competitive than others, and some may pay more or less attention to instructions or tasks required of them. Thus, the researcher tries to take into account those participant-related variables that have the potential to impact on the DV (in addition to the IV), and therefore possibly distort the results.

For example, a researcher conducting an experiment on sex differences in aggressive behaviour after playing a violent video game will recognise that participant characteristics such as age, personality, mood, prior experience with violent video games, cultural background, and so on, can also influence aggressive behaviour. Consequently, the researcher will try to ensure that the influence of these other participant variables is controlled or minimised and will do so before the experiment is conducted. How this can be achieved is described in the next section.

UNCORRECTED PAGE PROOFS

Figure 1.21 Even if they have the same cultural background, individual participants will differ in a diverse range of other personal characteristics and abililites. Participant variables arising from such differences are potential extraneous or confounding variables that need to be identfied and controlled, or at least monitored.

BOX 1.5 Participants' roles in experiments

American psychologists Stephen Weber and Thomas Cook (1972) conducted one of the earliest investigations on how being a participant in research can influence attitudes and behaviour. Their investigation led them to identify four distinct roles that some participants may adopt in a laboratory experiment and thereby influence the results in an unwanted way:

- The *'faithful' participant* tries to react to the experimental conditions as naturally as possible, either deliberately or out of disinterest.
- The *cooperative participant* tries to work out the hypothesis being tested in order to help support it.
- The *negativistic participant* tries to work out the hypothesis and what the researcher hopes they will do so that they can do the opposite or give obviously 'silly' replies.
- The *apprehensive participant* believes that the researcher is out to discover some hidden truth about them, and makes every effort to avoid a negative evaluation of themself.

BOX 1.6 The Hawthorne effect

In a series of well-known experiments over a period of five years, American researchers Fritz Roethlisberger and William Dickson (1939) tested different ways of increasing productivity among employees at the Western Electric Company's Hawthorne plant in Chicago.

At various times the employees were subjected to different work conditions such as shorter working periods, longer working periods, long rest breaks, short rest breaks, better lighting conditions, poorer lighting conditions, and work incentives such as bonus payments. Under most of these conditions, it appeared that productivity increased. This observation led to a conclusion called the *Hawthorne effect* — that if participants are aware that they are members of an experimental group, performance may improve simply because of that fact (rather than because of the IV — or experimental treatment — to which they are exposed).

Some psychologists believe that the Hawthorne effect is best described as the *Hawthorne defect*. This is because follow-up research suggests that many of the original Hawthorne experiments did not actually produce the increased productivity reported by the researchers. Reports of the Hawthorne study concentrated on only one of the experiments in which there was a big improvement in productivity. Nonetheless, it is possible that participants' knowledge of being in an experiment may affect the results in an unwanted way, and researchers make every effort to control this unwanted influence.

Figure 1.22 These workers took part in one of the Hawthorne experiments.

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.8

Review questions

1. In what way can individual participant differences be a source of extraneous or confounding variables? Explain with reference to an example not used in the text.
2. A researcher is planning to conduct an experiment to test the influence on exam performance of the amount of time spent studying.
   (a) Formulate a research hypothesis that could be tested.
   (b) How will the IV and DV be operationalised?
   (c) Identify three participant variables that are potential confounding variables in the experiment and explain your choice of each variable.

eGuide*plus*

Practical activities on individual differences

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. This can be a particular concern when research is carried out by multiple research assistants or when large sample sizes require many researchers to complete the project. For example, suppose that the researcher is interested in studying factors influencing the reaction time of helicopter pilots when flying over a hostile war zone at night. The researcher sets up an experiment in which participants perform a task in which they have to detect the blink of a faint red light in a dark room as quickly as possible.

Imagine how the results could be affected if 20 participants received different *instructions* on what the experiment is about, what they are supposed to do, whether they can sit or stand, how much time they have to respond, and so on. What if some participants complete the task early in the morning and others late at night (and may therefore be more or less alert)? Or what if some participants complete the task in a room with 'darker' conditions?

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 uniform, or 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 variables, and this feature is a source of potential extraneous and confounding variables that can influence the DV and therefore the results.

PAGE PROOFS

Figure 1.23 A researcher studying the reaction time of helicopter pilots when flying over a war zone at night may set up an experiment to measure how quickly participants detect the blink of a red light in a dark room.

Order effects

In some experiments, participants are exposed to more than one treatment condition (IV) and they may be required to perform the same type of task twice or even many times under different treatment conditions. For example, in an experiment to determine the effects of alcohol on driving performance, the *same* group of participants may be exposed to one treatment condition (or IV) for which they do not drink any alcohol before a driving test in a simulator (a control condition). After a short break, the participants may then be exposed to another treatment condition (or IV) for which they are given an alcoholic drink before completing the test. The order or sequence in which these conditions are administered can be a problem in an experiment with this type of design (called *repeated measures* and described in detail on pages 56–7).

An **order effect** occurs when performance, as measured by the DV, is influenced by the specific order in which the experimental tasks, treatments or conditions are presented rather than the IV. This essentially means that 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 illustrate 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 and/or prior experience doing 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 fatigue or tiredness (sometimes called a *fatigue effect*). Similarly, their performance may be influenced by boredom due to repeating the same task, especially if the task takes a long time, is simple and does not change. Boredom is quite common in experiments in which participants are required to complete multiple trials or tests, especially when task requirements are not particularly interesting.

Carry-over effects are the influences that a particular task has on performance in a task that follows it. They arise simply from experiencing a task. The effect of merely experiencing a task has the potential to 'carry-over' to the next task (or condition), regardless of whether the task is the same or simple. 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.

As with practice effects, a carry over effect can assist or hinder performance. For example, 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', improving or lowering performance the next time the task is completed (driving in the simulator again) depending on the participant's perception or feeling the first time. Either way, this is an unwanted effect as it is a potential confound.

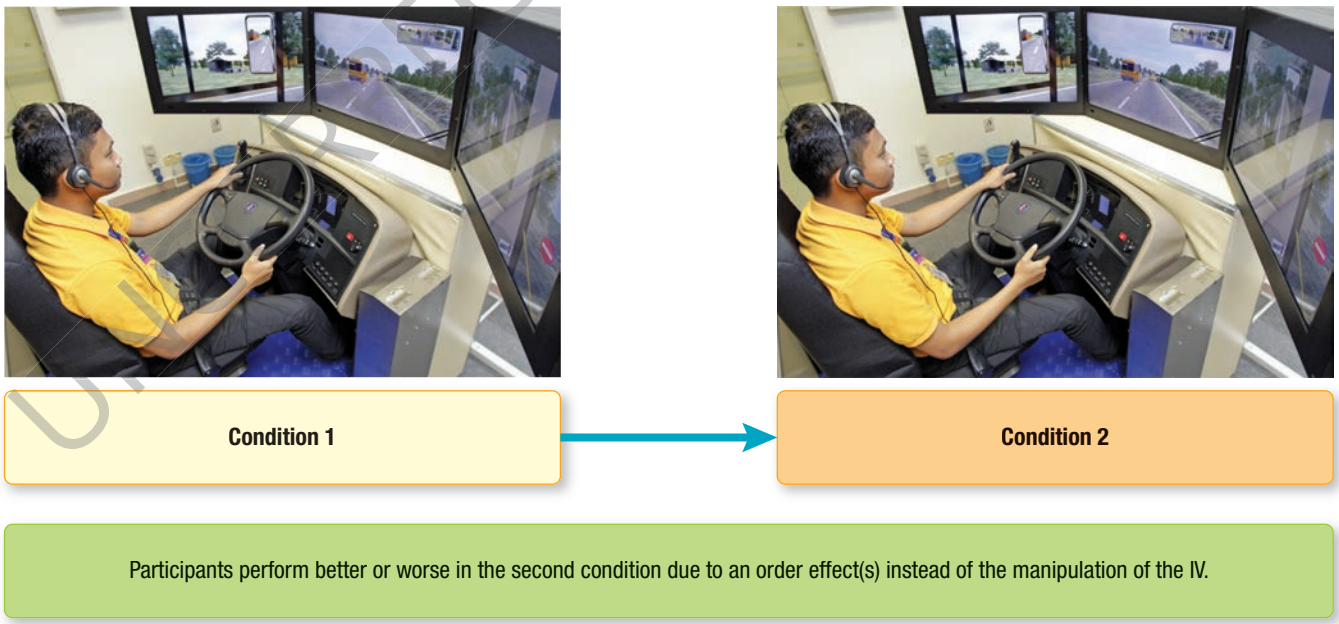


Figure 1.24 Order effect

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.9

Review questions

1. (a) Explain the meaning of the term non-standardised in relation to research procedures.
(b) Give an example of an instruction to participants and one or more experimental procedures that would be considered to be non-standardised.
(c) Explain why non-standardised instructions and procedures are a source of potential confounding variables.

2. (a) Explain the meaning of order effect.
(b) Distinguish between practice effects and carry-over effects with reference to how these types of order effects can lead to higher or lower scores on a measure of the DV.
(c) Explain why order effects are potential confounding variables.

3. Suggest a possible order effect for each of the following experiments and explain whether it is likely to help or hinder performance. In all examples, there is a minimal rest break between the different experimental conditions.
(a) An experiment is conducted on the effect of the rate of presentation of new information on memory. Participants are exposed to two treatment conditions in which they are presented with a list of words and asked to recall as many words as they can. In condition A, the words are presented at a rate of one word per second. In condition B, a list of similar words ('parallel version') are presented at a rate of two words per second. The number of words correctly recalled is measured for both conditions.
(b) An experiment is conducted to test the effectiveness of a new keyboard design which has a different arrangements of keys. In condition A, participants use a keyboard with a non-standard QWERTY arrangement to type a page of text. In condition B, the same participants use a keyboard with a QWERTY arrangement to type a parallel version of the text. Speed and accuracy is measured for the two different treatment conditions.
(c) An experiment is conducted to test whether listening to music when essay-writing will help students write better. In condition A, participants write a short essay in response to a prompt. In condition B, the same participants listen to music playing through headphones and write another short essay in response to a parallel version of the prompt. The essays are assessed using a standardised set of criteria.

eGuide*plus*

Practical activity
Order effect

Experimenter effect

Personal characteristics of the experimenter (or any other researcher) and their behaviour during an investigation are also sources of extraneous and confounding variables. The **experimenter effect** is an unwanted influence(s) on participant performance and therefore the results that is produced intentionally or unintentionally 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, sometimes called an *expectancy effect*, 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, not simply because they have correctly anticipated a response, but rather because they have helped to shape the response through their expectations (Rosenthal & Rubin, 1978). The results may therefore be attributable to behaviour associated with their expectations rather than the IV. Actions that can promote a self-fulfilling prophecy include:

- facial expressions, such as smiling at participants in the experimental or control 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.

American psychologists Robert Rosenthal and Lenore Jacobson (1966, 1968) demonstrated in a well-known experiment involving teachers and schoolchildren how experimenter expectancy can promote a self-fulfilling prophecy. They found that primary school teachers' expectations of the performance of their students affected how well the children actually performed. Students whose teachers were led to believe that they were 'bloomers' who were showing 'unusual potential for intellectual growth' and were therefore expected to develop rapidly performed better than students whose teachers were led to believe that they were not 'bloomers'. Yet the students hardly differed in their initial intellectual abilities, assessed at the outset of the study using an intelligence ('IQ') test. Teachers were found to have unintentionally influenced the performance of their students, depending on what they had been told by a researcher.

UNCORRECTED PAGE PROOFS

The experimenter effect involves not only the expectations and cues or actions of the researcher that influence participant responses in research settings, but also unintentional biases in the collection and/or treatment of data. This kind of experimenter effect is commonly referred to as *experimenter bias*. Studies have found that when the person measuring the dependent variable is aware of the purpose or hypothesis of the experiment, it is possible on some occasions for them to do such things as misread data, misperceive a rat's reaction in a maze, misinterpret a participant's verbal response, or give unintentional assistance to participants. This is more likely when the researcher wants a particular pattern of data, a particular verbal response from a participant, or a rat to take a particular turn in a maze. These are not examples of intentional dishonesty. They are examples of unintentional or unconscious errors that can be made in collecting and analysing data because of a researcher's close involvement with their study (Gerow, 1995).

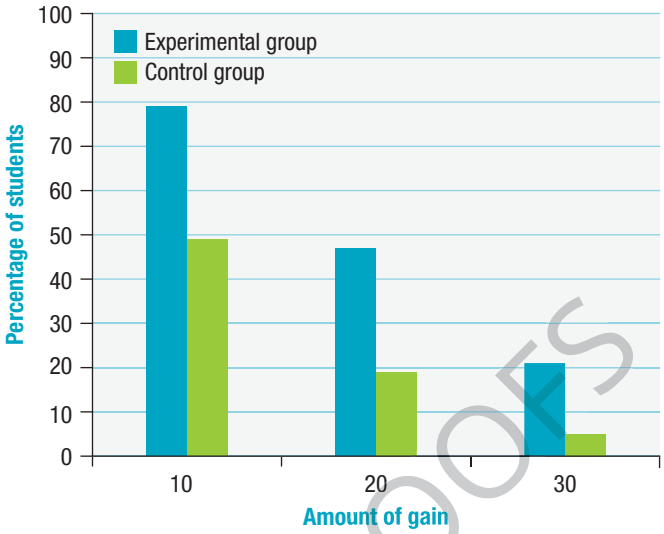


Figure 1.25 IQ score gains by Grades 1 and 2 students in the Rosenthal and Jacobson experiment.

Source: Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinates of pupils' IQ gains. *Psychological Reports*, *19,* 115–118.

Figure 1.26 A teacher's expectations of the performance of their students can affect how well the students actually perform.

UNCORRECTED PAGE PROOFS

BOX 1.7 Demand characteristics

Suppose that you are a participant in an experiment being conducted mid-afternoon in a laboratory. The 'laboratory' is a classroom in the psychology department at a university that looks as if most of the furniture has been rearranged or removed. A researcher wearing a lab coat and holding a clipboard and plate of biscuits walks into the room, places the plate on the table in front of you and says, 'Normal people crave biscuits at this time of day, so eat if you want to'. There's a good chance that you'll eat one. Alternatively, suppose that the researcher says, 'Only people who have no self-control eat at this time, but eat if you want to'. There's a good chance that you won't. In any situation, the social and physical surroundings provide cues, or signals, that essentially 'demand' that we behave in a certain way. In research, these cues are called demand characteristics.

A *demand characteristic* is a cue expressed by the researcher or present in some aspect of the research study that communicates the kind of response that is expected from participants and leads them to believe that the research study requires, or 'demands', that they respond in a particular way. In an experiment, demand characteristics guide or bias a participant's behaviour in some way within that setting. Participants rely on demand characteristics to answer such questions as 'What's really going on here?', 'What is this person trying to find out from me?' and 'What am I supposed to do?' Participants don't necessarily respond to demand characteristics intentionally or even consciously. However, demand characteristics typically result in reactions that are not valid or reliable responses to the variables under investigation (Heiman, 2002).

American psychologist Martin Orne (1962) was one of the first researchers to scientifically study demand characteristics and describe the concept as an extraneous variable that can influence the results of a research study, particularly in experimental research. Orne and his research assistants developed a set of tasks which he believed most people would refuse to do, or would do only for a short period of time. According to Orne, 'the tasks were intended to be psychologically noxious, meaningless, or boring, rather than painful or fatiguing'. For example, one task involved additions of rows of random numbers on a sheet of paper. In order to complete just one sheet, participants were required to perform 224 different additions. A stack of about 2000 sheets was presented to each participant, which was clearly an impossible task to complete.

After the instructions were given, the participants' watches were taken away and they were told to continue working until they were told to stop. 'Five and one-half hours later, the experimenter gave up!' (p. 777), amazed that the participants had conscientiously stuck to the task without question. In a variation of the task, which was designed to be 'a more frustrating situation', participants were asked to do the additions on each sheet, then tear up their answer sheet into at least 32 pieces before continuing with the next one, and so on. Again, participants tended to continue for several hours at this completely meaningless task. All persisted without question, complaint or hostility, simply because they were 'in an experiment'.

Orne argued that one reason that participants willingly do things in an experiment/research study which they would not normally do is because of *social desirability*; that is, their desire to provide socially acceptable responses. Essentially, participants 'adjust' their responses so that they aren't embarrassed or so they won't appear strange or abnormal. Similarly, participants tend to do their best to be 'good participants' and not 'upset' the experimenter. But this is not always the case.

Consider a well-known research study that highlights the influence of demand characteristics. German psychologists Fritz Strack, Leonard Martin and Sabine Stepper (1988) tested a hypothesis that the facial muscles used for smiling provide feedback to the brain that actually improves mood. In other words, being happy makes you smile, but then smiling makes you even happier. To test the hypothesis, they measured the dependent variable of mood after participants had held a pen in their mouths, either using their teeth (which mimics smiling) or using their puckered lips (which does not mimic smiling). Imagine that you are in this experiment, sitting in a room with a pen sticking out of your mouth, with a 'psychologist' watching you. What would you be thinking and feeling, and how would that influence what you said and did? Would you really react naturally? If the answer is 'No', then it's because of the experiment's demand characteristics. This variable is one of a number that may account for the inability of many researchers to repeat or reproduce the experiment's results (Heiman, 2002; Reber, 2016; Wagenmakers, 2016).

Figure 1.27 Under 'stage hypnosis', these audience members perform various silly behaviours that they would not ordinarily do in public. Can their behaviour be explained by demand characteristics?

UNCORRECTED PAGE PROOFS

Placebo effect

In medicine, the placebo effect refers to an improvement in health or wellbeing due to an individual's belief that the treatment given to them will be effective. The placebo effect is evident when a patient recovers from an illness or pain after they have been given a substance or a treatment that has no actual medicinal or therapeutic value, such as a 'sugar pill' or fake injection. This inactive substance or fake treatment, which substitutes for the real substance or treatment, is called a **placebo**. The mere suggestion to the patient that they have received, or will receive, some kind of treatment is often sufficient to minimise or eliminate the symptoms. For example, some people begin to feel better if they are put on a waiting list for treatment compared with how they might feel if they were not on a waiting list.

In an experiment, the **placebo effect** occurs when there is a change in the behaviour or responses of participants 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 due to their belief that they have received some treatment. There is a change in behaviour in the absence of an experimental manipulation.

For example, consider an experiment on the effects of alcohol on driving performance that uses an experimental and control group, each with different participants. The experimental group would be given an alcoholic drink before a driving test in a simulator. Participants in the control group would not drink any alcohol before completing the test. Suppose that the experimental group makes many more driving errors than the control group. Although the researcher would like to conclude that the difference was due to alcohol consumption impairing performance of the experimental group, this conclusion would not be valid.

The problem is that alcohol consumption may not have been the only variable that adversely affected performance of the experimental group. The act of being given an alcoholic drink by a researcher might have promoted expectations in participants about how they should behave. Experimental group participants might have perceived that they were given alcohol because they were expected to act drunkenly, so they did. Consequently, they may have driven as if they were drunk and made more driving errors. Furthermore, 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 — the researcher cannot be certain whether it was the effects of alcohol or the placebo effect that caused the performance difference.

UNCORRECTED PAGE PROOFS

Figure 1.28 The placebo effect involves a change in behaviour in the absence of an experimental manipulation. It can be triggered by the belief that a treatment is real, even though it isn't.

BOX 1.8 Experiment on the role of cognition in emotion using a placebo

In the early 1960s American psychologists Stanley Schachter and Jerome Singer (1962) conducted a controversial experiment to investigate the role of psychological factors in human emotion, especially cognitive processes. Their experiment provided evidence for their theory of human emotion in which they proposed that the experience of any emotion arises from a combination of two factors. One factor is our state of physiological arousal which accompanies an emotion and the second factor is our interpretation of the situation in which the arousal is experienced.

The participants were 184 volunteer male students enrolled in an introductory psychology course at the University of Minnesota who consented to an injection of a new medication called 'Supoxrin', which was actually made up of vitamins. They were informed that the purpose of the experiment was to determine the effects of the drug on vision. All received two extra points on their final exam for every hour they served as experimental participants.

There were four conditions in which participants were actually given either an injection of epinephrine (adrenaline) or a placebo. Epinephrine is a hormone that has the effect of arousing the body and producing physiological reactions such as increases in blood pressure and heart rate, rapid breathing, palpitations (trembling), flushing and sweating palms. These changes are also associated with strong emotional states such as excitement, elation and anger. The epinephrine effects usually begin within 3–5 minutes and may last from 10 minutes to an hour or so (depending on dosage and the individual). For most people, the effects wear off within 15–20 minutes after injection.

Although all participants in an experimental group were given the same drug, they were told different information about the effects of the drug that might be expected. Some participants were told a quite accurate account of the possible effects of the drug (the 'informed' group), some participants were told that the drug caused numbness and itching, with no suggestion that there might be excitement or agitation (the 'misinformed' group), and other participants were told nothing of what effects to expect (the 'uninformed' group). There was also a control group of participants who received a placebo injection that had no effect and were given no instructions of what to expect.

After receiving their injections, the participants were asked to wait in a room until the experiment began. In this room was another person who they thought was also a participant. However, the other person (called a 'confederate') was actually a research assistant. All instructions to participants were standardised as much as possible.

Schachter and Singer wanted to observe how the increased arousal level would be interpreted in different situations. They predicted that the 'informed' participants, who knew that the injection caused particular physiological changes, would not experience any strong emotion as they would attribute their arousal to the drug. They also predicted that the 'misinformed' and 'uninformed' participants would attribute their arousal to an emotion that could be associated with the situation they were in.

Two different situations were created using the confederate, who acted in either of two ways. In one situation, the confederate played the clown and acted in a very happy and silly manner. For example, he played with a hula hoop and tossed scrunched-up paper into a bin. In the second situation, the confederate acted as if angry and irritated. The confederate never knew which condition any participant was in.

Emotional responses were measured by researchers, who observed the participants through a one-way mirror. In addition, participants completed a self-report measure — a questionnaire requiring them to rate their feelings in a given situation.

The results showed that participants who waited with the happy confederate tended to interpret their increased arousal level as happiness. Those who waited with the angry confederate tended to interpret their increased arousal as anger. This effect, however, was achieved only with participants who were misinformed or uninformed about the effects of the injection. The informed participants tended to attribute their increased arousal level to the drug.

When there was nothing specific to help explain their physiological arousal (e.g. an injection), participants responded in ways consistent with their interpretation of what was happening in the situation they were in.

Figure 1.29 All participants were given either an injection of epinephrine or a placebo.

UNCORRECTED PAGE PROOFS

Not only did the uninformed participants respond differently in different situations, but they also reported different emotions. Their labelling of how they were feeling (i.e. their emotional state) depended on cognitive factors.

What does this experiment demonstrate? Generally, that physiological arousal affects a person differently, depending on their cognitive interpretation of the situation they are in. When a person cannot explain their physiological arousal, they seek explanations from their environment. In the experiment, the confederate provided cues that helped explain a participant's emotional experience. However, when participants thought that their physiological arousal was caused by an injection, they generally did not attribute their physiological changes to any particular emotion (i.e. happiness or anger). The experiment also demonstrates that physiological arousal is important in emotion because the group who received a placebo would have behaved in the same way as participants in the misinformed groups.

In sum, the Schachter and Singer experiment suggests that physiological arousal and cognition interact to produce emotions. However, their relationship in the experience of emotion is likely to be more complex than proposed by their two-factor theory.

Other psychologists have attempted to replicate this complex experiment in order to test the findings. However, replications of the experiment found differing results. Some have confirmed the findings while others have not. Furthermore, when found, differences in the results between the groups of participants were generally small and not very significant when statistically tested.

LEARNING ACTIVITY 1.10

Review questions

1. (a) Define the meaning of the term experimenter effect.
(b) What is experimenter expectancy and how can it produce
(i) a self-fulfilling prophecy?
(ii) experimenter bias?
(c) Explain why experimenter effects are potential confounding variables.

2. (a) What is a placebo in relation to experimental research?
(b) Define the term placebo effect in relation to experimental research with reference to an example not used in the text.
(c) Explain why the placebo effect is a potential confounding variable.
(d) (i) Sam has the sticker shown at the right on her school diary. Sam's friends notice that her behaviour has changed since the appearance of the sticker. Explain Sam's behaviour change in terms of a placebo effect.
(ii) Give three other everyday life examples of how the placebo effect may influence thinking, feeling or behaviour.

3. Suggest a way of controlling or minimising the influence of
(a) an experimenter effect
(b) a placebo effect.

eBook*plus*

Weblinks
- TED-Ed Lesson: The power of the placebo effect 4m 38s
- Theories on how placebos work

LEARNING ACTIVITY 1.11

eBook*plus*
Word copy of table

Summarising potential extraneous and confounding variables

Complete the following table.

Variable	Description	Example	Why a potential extraneous or confounding variable?
individual participant differences			
non-standardised instructions and procedures			
order effect			
experimenter effect			
placebo effect			

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.12

Evaluation of research

Consider the research by Schachter and Singer (1962) on the role of cognition in emotion summarised in Box 1.8 on pages 36–7 and answer the following questions.

1. Formulate a research hypothesis that could have been tested by the procedures used for the experiment.
2. Identify the operationalised independent and dependent variables in the experiment.
3. Identify the characteristics of participants and how they were selected.
4. Explain whether the sample was representative.
5. Suggest an appropriate procedure for assigning participants to different groups and explain your choice.
6. Identify the experimental and control groups used in the experiment and briefly outline the conditions for each group.
7. Describe two strengths and two limitations of the experiment.
8. Explain two significant ethical issues that are relevant to this particular experiment.

LEARNING ACTIVITY 1.13

Reflection

Comment on how accurately the cartoon below depicts a placebo effect.

"If this doesn't help you don't worry, it's a placebo."

eGuide*plus*

Weblinks

Media items on the nocebo effect

Ways of minimising extraneous and confounding variables

When conducting an experiment, a researcher will approach the task in a systematic way, controlling the variables under investigation in order to observe and measure what happens, and to prevent or minimise the influence of variables other than those being tested that might influence the results. The extent to which extraneous and confounding variables are anticipated and controlled determines the quality of an experiment and the validity and reliability of the procedures and 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 and allocation

The way participants are selected and how they are allocated to different groups or conditions are very important features of experimental research. Selection and allocation procedures provide the most commonly used means of minimising the influence of extraneous and confounding variables associated with participants.

For example, suppose a psychology lecturer at a university wanted to find out which of two teaching methods is more effective. The lecturer teaches two first-year psychology classes, one that starts at 8 am and one that starts at 4 pm. The lecturer uses one teaching method for the morning class and a different method for the afternoon class. At the end of the semester, the lecturer finds that the final examination scores are higher for the morning class. The researcher concludes that from now on all lecturers will use that particular teaching method for all classes. Is this a valid or legitimate conclusion to draw on the basis of the results obtained from the research?

The problem is that the two groups of participants may not be sufficiently alike in personal characteristics of relevance to the study and which may therefore have influenced the results. For example, people who enrol for lectures that start at 8 am may differ in some ways from those who enrol for a 4 pm lecture. Some people prefer to get up early, while others like to sleep late. Perhaps some students had commitments, such as casual work or other activities scheduled late in the afternoon, that prevented them from enrolling in the 4 pm class.

The likely cause of this problem is that the participant sampling and allocation procedures were inappropriate. Therefore, it cannot be concluded with confidence that the differences in the two groups' examination scores were caused solely by the difference in teaching methods (Carlson & Buskist, 1997).

UNCORRECTED PAGE PROOFS

Given the importance of participant sampling and allocation, we consider different types of procedures that can be used to minimise extraneous and confounding variables relevant to this aspect of the research design.

Participant selection

The process of selecting participants from a population of research interest (the 'target population') is called **sampling**. It is usually undertaken with the goal of being able to use the participants in the sample to draw conclusions about the larger group who form the population. This is not unlike the goal of a medical researcher who analyses a sample of someone's blood to draw one or more conclusions about all of that person's blood.

A sample should 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. When sampling, it is important to ensure that the sample lacks bias and is like its population in as many ways as possible so the results can be generalised ('applied') to that population. 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. For example, for an experiment on the soft drink and fast food preferences of adolescents, personal characteristics of participants such as their age, sex, income, access to retail outlets and cultural background are among the variables that could be assumed to be important. Variables such as height, hair colour, intelligence and spatial abilities may be assumed to not be important.

When a researcher selects a sample that mirrors or is approximately the same as its population, the sample is called a representative sample. A **representative sample** is a sample that closely matches the population from which it is drawn in every important participant variable. When a sample does not adequately represent the key characteristics of its population it is referred to as a **biased sample**.

Larger samples also tend to minimise the likelihood of an unexpected '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. For example, 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.

Researchers can use many different procedures to select a sample. Commonly used procedures are called random sampling, stratified sampling and convenience sampling. Convenience sampling is the simplest method but is less likely to achieve a representative sample than the other procedures.

UNCORRECTED PAGE PROOFS

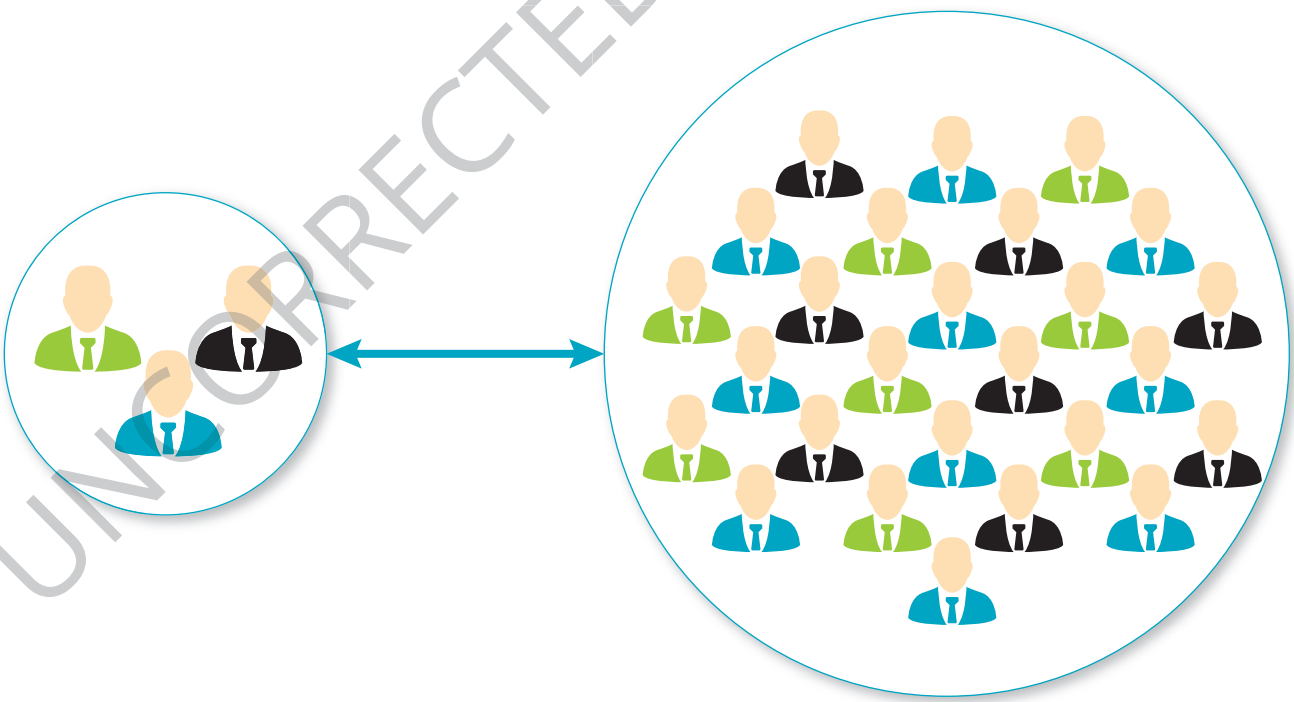


Figure 1.30 A representative sample closely matches the population of research interest in the distribution of key participant characteristics. Generally, the larger the number selected, the more likely it is to be a representative sample rather than a biased sample.

Random sampling

When used in relation to sampling and samples, 'random' does not mean 'haphazard'. In fact, it is anything but haphazard. Nor does random mean selecting participants (or allocating them to experimental or control groups) according to the whims of the researcher. Using a random sampling procedure actually involves a very careful methodical approach or plan.

Random sampling is a sample selection procedure that ensures every member of the population of research interest has a genuinely 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 target population (with appropriate contact information). This source material or 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. Other sampling frames can include a list of employees in an organisation of research interest, a list of members of an AFL club, patient files in a hospital, and so on. Ideally, the sampling frame will be in an electronic format so that it can be used on a computer or tablet.

After the sampling frame is accessed, the researcher could obtain a random sample using a simple lottery procedure to select the required number of participants. The lottery procedure could involve putting the names or ID numbers of all members of the target population on equal-sized slips of paper, placing the slips in a container and mixing them thoroughly, then drawing out the required number of slips 'blindly', one at a time. As a result of this simple but methodical 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. For example, if this procedure was used to select a sample of five students from a small box with the names of all 20 students in a psychology class, any group of five names is equally likely to be selected as any other group of five names.

A commonly used method when a large number of participants is required is to assign a number to each member of the target population, then use a digitally generated list of random numbers to select sample members.

Suppose you are interested in studying some aspect of student behaviour at your school and you want a random sample of 20 students. You would begin with a list of all students currently enrolled at your school. Then, each student is assigned a number. If, for example, there are 1000 students, the first student in the list is assigned number 1 and the last student assigned 1000. A random number generator (available online, in an app or scientific calculator) could then be used to produce 20 numbers that fall between 1 and 1000. The students whose numbers are selected become the sample. For example, if the first random number is 47, then the 47th person in the list is included in the sample; if the second random number is 10, then the 10th person in the list is selected, and so on until the 20th participant has been selected.

This sampling procedure ensures that every student in the school (the relevant population) has an equal chance of being chosen to be part of the sample. Therefore, 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.

If everyone in a target population does not have an equal chance of being selected as a participant, then *sampling bias* is said to occur. Sampling bias increases the likelihood of a biased sample being obtained. For example, a researcher might conduct a study on stress management strategies used by Victoria Police. A random sample could be obtained by allocating a number to all Victorian police officers and then selecting participants' names using a lottery method. However, sampling police officers 'at random' in a 'hit-or-miss' way in the street or at a nearby police station would achieve a biased sample rather than a truly random sample, because not all Victorian police officers (the target population) will have an equal chance of being selected into the sample at these sampling locations or when the sampling is done.

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.

An advantage of random sampling is that the sample obtained is not biased in any way by a researcher's behaviour or preferences. In addition, every potential participant has an equal chance of being selected, which helps ensure a highly representative sample, particularly when everyone who has been selected can be contacted and agrees to participate. This allows generalisations that are more likely to be considered to have external validity (as described on page 109). The larger the sample, the more likely it is that the sample will be representative.

However, there is no guarantee all random samples are representative. For example, not all those who have been selected may be contactable, available or agree to participate, which can be a problem when the sample size is small. Others may agree to participate then refuse to do so or withdraw (dropout) after the study commences.

UNCORRECTED PAGE PROOFS

A limitation of random sampling is its need for a complete and up-to-date sampling frame that includes all members of the target population. There may be no single list available with relevant details of all the population. As a result, it may be difficult and time consuming to bring together numerous sub-lists to create a final list from which to select the sample. If available, it may be difficult to gain access. For example, the list may be protected by privacy policies or require a lengthy process for permission to access. If accessed, the process of random selection may be time-consuming.

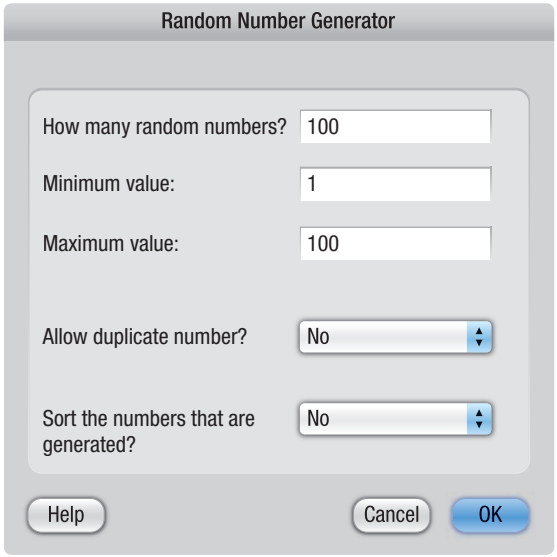


Figure 1.31 One of many random number generators that are freely available online

eBook*plus*

Weblinks
Online random number generators

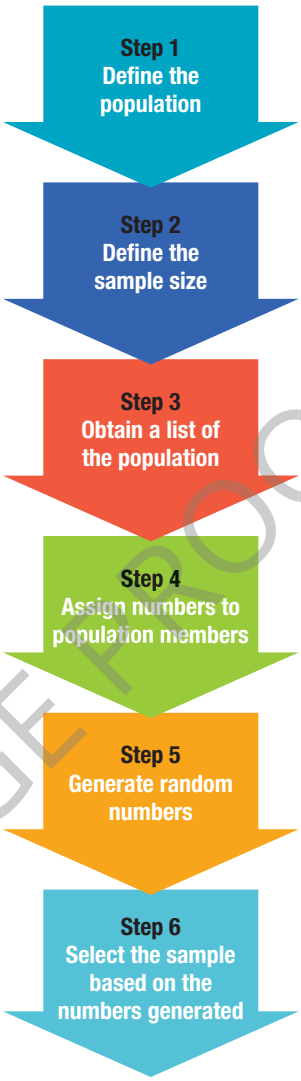


Figure 1.32 A random sampling procedure using a randomly generated number sequence

TABLE 1.2 Example of 100 randomly-generated numbers (unsorted) to select a sample of 20 from a school's student population of 1000. The first participant selected for the sample would be the 47th student in the school list, the second selected would be the 10th in the list, and so on.

47	113	958	780	970	553	464	936	767	23
10	220	410	818	167	792	578	197	935	188
963	389	990	846	10	673	537	790	300	577
323	362	597	32	518	232	665	802	298	103
404	860	252	631	401	191	414	624	770	26
559	193	861	383	917	650	972	997	358	878
120	459	448	472	489	823	703	871	400	671
821	617	883	21	62	130	169	274	746	84
284	981	605	372	393	656	16	516	809	610
451	141	799	687	490	628	90	155	533	912

UNCORRECTED PAGE PROOFS

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, if a psychologist wanted to determine the attitudes of Australian voters to asylum seekers, they could reasonably expect that people's attitudes would differ depending on their age, sex, religion and cultural or ethnic background. Consequently, the psychologist would want to ensure that each of these groups was represented in the final sample in the same proportions that they were known to exist in the voting population. This can be achieved by using a stratified sampling procedure.

Stratified sampling involves dividing the population to be sampled into distinct subgroups (called *strata*), then selecting a separate sample from each subgroup (*stratum*), usually in the same proportions as they occur in the target population. Age, sex, religion, cultural background, residential area, educational qualifications, IQ score, income level and income type (e.g. wages or pension) are examples of characteristics that may be used to divide a population into strata.

The stratified sampling procedure is commonly used to study psychological characteristics or attitudes that vary greatly among different subgroups of a population. For example, suppose you were going to undertake an investigation on behavioural responses of students at your school towards their teachers' use of rewards and punishments in the classroom. You expect that responses may differ among students in different year levels so you want to ensure each year level (stratum) is represented in your sample, more so in about the same proportions as they occur in the school population.

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 school population. Using this *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. It would also be referred to a type of *quota sampling* because the sample includes a predetermined percentage of each subcategory of the target population.

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 a phobia and will be exposed to a new type of relaxation therapy to help manage their anxiety. In addition, when there is random sampling from appropriately sized proportions of the strata, this helps ensure a high degree of representativeness of all the strata, which means that there can be greater precision in the study and its findings when compared to the standard 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 relatively complete lists of the target populations (strata) are available and accessible. If accessed, a representative sample cannot be obtained unless stratified random sampling is used (assuming there is limited missing data). Either way, stratified sampling can be a very time-consuming procedure, more so than standard random sampling.

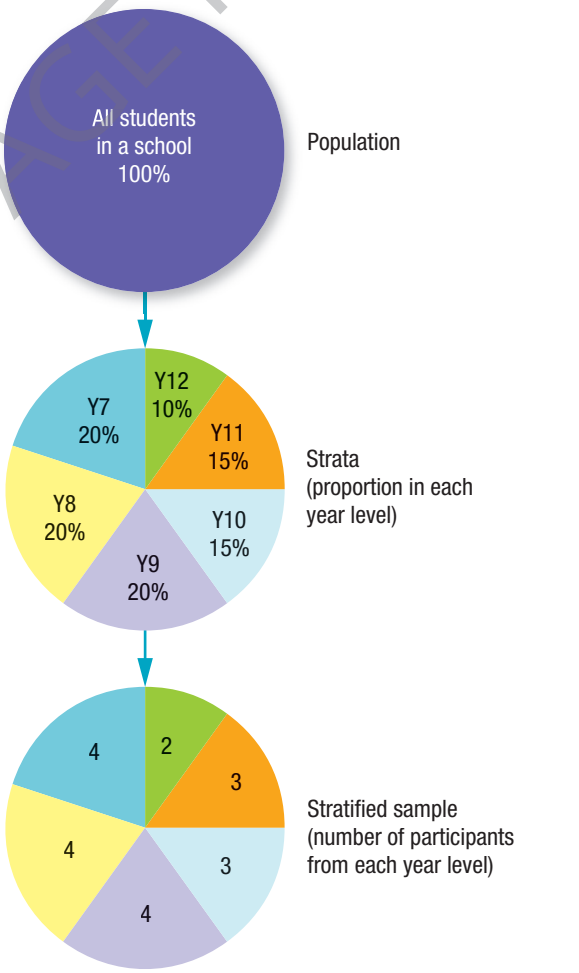


Figure 1.33 An example of stratified sampling for a school-based investigation. Notice how the proportions in the sample match the percentages in the population.

UNCORRECTED PAGE PROOFS

BOX 1.9 Stratified random sampling

Stratified random sampling involves identifying all of the people within each stratum of research interest, then randomly selecting samples of a set size from within each stratum.

Suppose, for example, you were going to undertake an investigation involving volunteer participants from the VCE population at your school or college. You expect that volunteer rates may differ among students doing maths and science courses, performing arts courses and so on.

You could first make separate lists of the students in each of these courses and then randomly sample from each list. If, for example, 20% of the VCE population is enrolled in a maths and science course and 10% is enrolled in a performing arts course, then the sample should ideally also consist of 20% maths and science students and 10% performing arts students. Students from these courses would be represented in the same proportions in the sample as they are in the population.

An outcome of this sampling procedure is that the sample can be assumed to be truly representative and unbiased. However, obtaining a stratified random sample using set proportions is usually very time-consuming and difficult to achieve so the procedure is not often used.

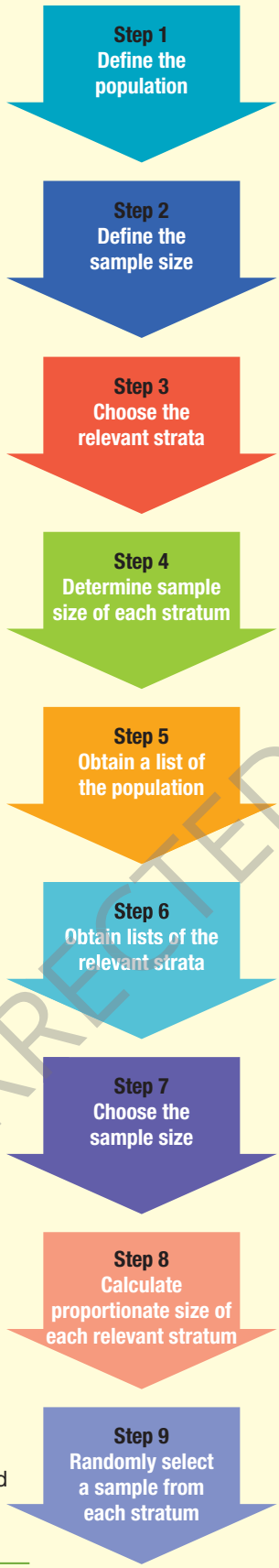


Figure 1.34 Steps in a stratified random sampling procedure using set proportions.

Convenience sampling

For some research studies it is not convenient, suitable or possible to obtain a representative sample. In such cases, a convenience sample may be used and the researcher may use anyone who is available or present at the time of the study.

Convenience sampling, also called *opportunity sampling* and *accidental sampling*, involves selecting participants who are readily or most easily available. There is no attempt to make the sample representative of a population. For example, a representative sample of illegal drug users or homeless teenagers is not often easily available. Consequently, the researcher may go to locations known to be frequented by the required participants and simply select the first individual they come across who is in the target population and who is willing and available to participate. They may also ask this participant to identify others who would possibly be suitable for the study and agreeable to inclusion (which is called *snowball sampling*).

Similarly, a researcher seeking to conduct a study on drivers who do not obey red traffic lights at a particular intersection at a particular time could use convenience sampling. Participants will be the drivers observed to disobey a red traffic light, and they are simply entered into the study until the desired sample size is reached.

Psychology students often use convenience sampling; for example, when selecting participants they have the opportunity to study such as other students sitting in the library at their school, people walking by in the street, children who happen to be playing in the school yard at a local primary school, friends, parents or relatives.

In most cases, convenience sampling produces a biased sample because only those people present and 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 appear 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 convenience sampling involves sampling bias and the resulting 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 generalised to the entire population.

Despite these limitations, convenience sampling is widely used in psychology. Participants are readily available, so it tends to be quick (time-efficient), simple and inexpensive compared to other sampling procedures. Sometimes it is the only way to access participants to conduct a study. Convenience sampling can also be of considerable value to identify possible trends or patterns in results or when conducting

UNCORRECTED PAGE PROOFS

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 1.35 Convenience sampling involves selecting participants who are readily available. For example, a researcher may be interested in strategies of buskers that are more or less likely to promote audience donations, such as eye contact, verbal interaction, and involvement of audience members as active participants. Different buskers at known locations may simply be entered into the sample and observed.

SWINBURNE UNIVERSITY OF TECHNOLOGY

Research on hair pulling

Do you pull out your hair to the point of causing unintended hair loss?

Do you find your hair pulling difficult to control?

Does your hair pulling interfere with your life or cause you emotional distress?

Psychology researchers at Swinburne University (Melbourne, Australia) are currently developing a measure of thoughts and beliefs associated with **Trichotillomania** (aka **Hair Pulling Disorder**). It is hoped that the final measure can be used across Australia and internationally to improve our understanding of this disorder so we can develop more comprehensive, effective, and targeted treatments.

If you think you might have this problem and are aged 18 or older, we need your help!

In order to help us develop this measure you will be required to:

1. participate in a 30-minute interview about your hair pulling with a student investigator. Interviews can be conducted via Skype for interstate and international participants.
2. complete a series of online questionnaires about hair pulling, related thoughts and feelings, and other psychological symptoms.

If you think you can participate in this research or would like more information, please do not hesitate to contact either:

Dr M.N., Principal Investigator, ph. (03) 1234 5678, email: mn@email.edu.au

A.R., Student Investigator, ph. (03) 1234 6789, email: ar@email.edu.au

eBook*plus*

Weblinks
Advertisements for research participants

Figure 1.36 A commonly used means of accessing participants in a population of research interest is to advertise. These participants are volunteers who 'self-select' to be in the sample but they are not necessarily included. They may be in a predefined group but may not meet the researcher's criteria to be in the sample. Nonetheless, using a *volunteer sample* or *self-selected sample* as it is also called, despite the likelihood of bias, is a convenient way to find willing participants.

LEARNING ACTIVITY 1.14

eBook*plus*

Word copy of table

Review questions

1. What does sampling involve?
2. Explain whether the sample is considered to be all those selected for participation in research or the group who actually participate.
3. Explain the difference between a biased and representative sample with reference to an example of how each sample type is achieved.
4. What are two potential limitations of a small sample size?
5. Why is convenience sampling often described as 'opportunity sampling'?
6. Explain how the type of sampling procedure used can either reduce or increase the likelihood of an extraneous or confounding variable influencing research results in an unwanted way.
7. Complete the table below to summarise key features of three sampling procedures.
8. Suppose that you are required to determine the typical amount of nightly sleep of students at your school or college.
   (a) Briefly describe a convenience sampling, random sampling and stratified sampling procedure for selecting research participants.
   (b) Explain which of the three sampling procedures would result in the most highly representative sample.
9. You want to compare the lifestyles of VCE students in Melbourne and Mildura.
   (a) Define your population.
   (b) How could you obtain a random sample from each of these populations?
10. You want to test short-term memory capacity in preschool children, teenagers and people aged over 65 years.
   (a) Define your population.
   (b) How could you obtain a random sample from each of these populations?

Sampling procedure	Description	Example	Advantages	Limitations
random sampling				
stratified sampling				
convenience sampling				

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.15

Evaluation of samples and sampling procedures

1. (a) A researcher investigating variables that influence consumer decision-making will conduct a brief, 3-item survey outside a Myer store in a shopping mall on a Friday evening. The researcher will toss a 20 cent coin when ready to interview someone. If the coin shows heads, the person exiting the store will be interviewed. If the coin shows tails, the person will not be interviewed.

Explain whether the researcher will obtain a random sample.

(b) Another researcher will conduct the same survey at the same location but will vary the sampling technique. Instead of tossing a coin, the researcher will interview every 20th person who exits the store.

Explain whether the researcher will obtain a random sample.

2. Explain why each of the following research studies is likely to have sample bias.

(a) a survey on binge-drinking behaviour in a popular teenager's magazine

(b) a television or radio call-in survey

(c) a telephone survey at 6 pm on weeknights using landlines to people's homes

(d) a survey based on the number of 'likes' for a Facebook post

(e) a psychologist working at a rehabilitation centre for people with a brain injury accesses some of their relatives for a study on their coping strategies

(f) a researcher interested in the age and sex of gamblers who play the pokies interviews people entering a local gaming venue during a 4-hour period on a weekday afternoon

3. To which type of sampling procedure is the cartoon below most likely referring? Explain your choice.

"How would you like me to answer that question? As a member of my ethnic group, educational class, income group, or religious category?"

4. The rating of a TV show is determined by the size of the audience who watches it. The rating is based on data from an electronic recording device attached to a TV set in a viewer's home. The device automatically records which TV show is being watched and for how long. Viewers also indicate who in the household is watching. Suppose you are responsible for determining the TV ratings of the viewing audience in Melbourne (or a regional town) for a one-week period.

(a) Briefly describe the sampling procedure you could use to select the participants who will receive a recording device.

(b) Explain how your procedure would ensure a representative sample of the viewing population.

UNCORRECTED PAGE PROOFS

5. Consider the newspaper advertisement below, then answer the following questions.

How does long-term cannabis use affect your brain and memory?

The University of Melbourne is conducting a study examining how heavy, long-term cannabis use (daily or almost daily use for 10+ years) affects the brain. There are two parts to the study: a memory testing session and a brain scanning session. Each session takes approx. two hours and participants receive $50 in Coles Myer vouchers for each session. Participants should be betw. 18–35 years old, not using other drugs or alcohol regularly, and NOT have a diagnosed mental illness.

(a) Identify the target population for the sample.
(b) Is this sampling procedure better described as convenience sampling or random sampling? Explain your choice.
(c) A sample obtained through an advertisement is sometimes called a 'self-selected' sample. In what way is this type of sample self-selecting?
(d) Will advertising for research participants and using a gift voucher or incentive payment result in sample bias? Explain your answer.
(e) Will exclusion of some respondents from the sample result in sample bias? Explain your answer.
(f) How representative is the sample obtained using the advertisement likely to be?
(g) Will the researcher be able to generalise their results from the study described in the advertisement? Explain your answer.

eGuide*plus*

Practical activity
Testing different sampling techniques

UNCORRECTED PAGE PROOFS

Participant allocation

After participants have been selected, they have to be assigned to the different groups required for the experiment. This must be completed as systematically as the selection process so that personal characteristics of participants are distributed relatively evenly across the different groups (and therefore within the conditions of the experiment). In this way, participant variables that may become confounding variables are controlled. This type of control through random allocation is an *essential* characteristic of experimental research and also distinguishes it from other research methods.

Random allocation

It is to be expected that individual participants will have different abilities and other personal characteristics or backgrounds that may affect the outcome of an experiment. One way of minimising differences in the composition of the control and experimental groups is to randomly allocate, or assign, participants to the groups.

Random allocation, also called *random assignment*, is a procedure used to place participants in groups or conditions 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 to be used.

As with random selection, random allocation can be achieved using a lottery method in which chance alone will determine the group to which each participant is assigned. For example, drawing 'names out of a hat' 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 an 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 people who prefer to sit at the front or back of the classroom.

Random allocation does not guarantee that participants in the different groups or conditions are entirely equivalent before the experiment begins. However, it does reduce the likelihood of differences being present due to 'chance' factors, which helps rule out possible alternative explanations of the change measured in the DV. The likelihood of chance differences in participant variables between or even within groups tends to be reduced further as the sample size of each group increases.

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

Random sampling and random allocation can also be distinguished in relation to the validity of the experiment (or any other research study). Since random sampling involves selection of participants from a target population, it is crucial for the generalisability of the results and therefore *external validity*. After all, random sampling is used so that the participants will better represent the population from which they are drawn — the goal being to generalise the results for the sample to its population.

UNCORRECTED PAGE PROOFS

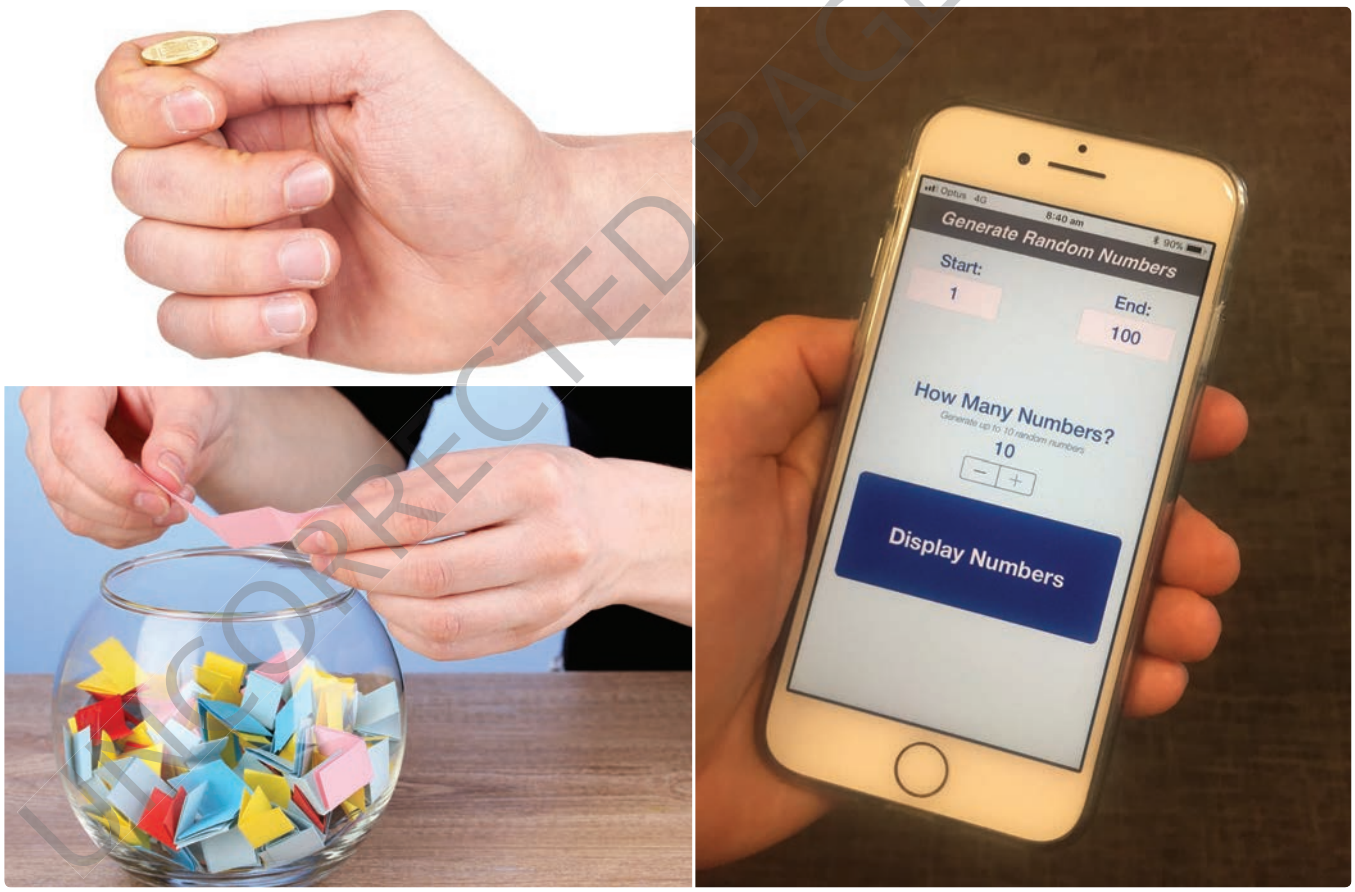


Figure 1.37 Coin tossing, drawing names from a container and a digital random number generator can be used for random allocation to groups.

eGuide*plus*

Practical activities on random allocation

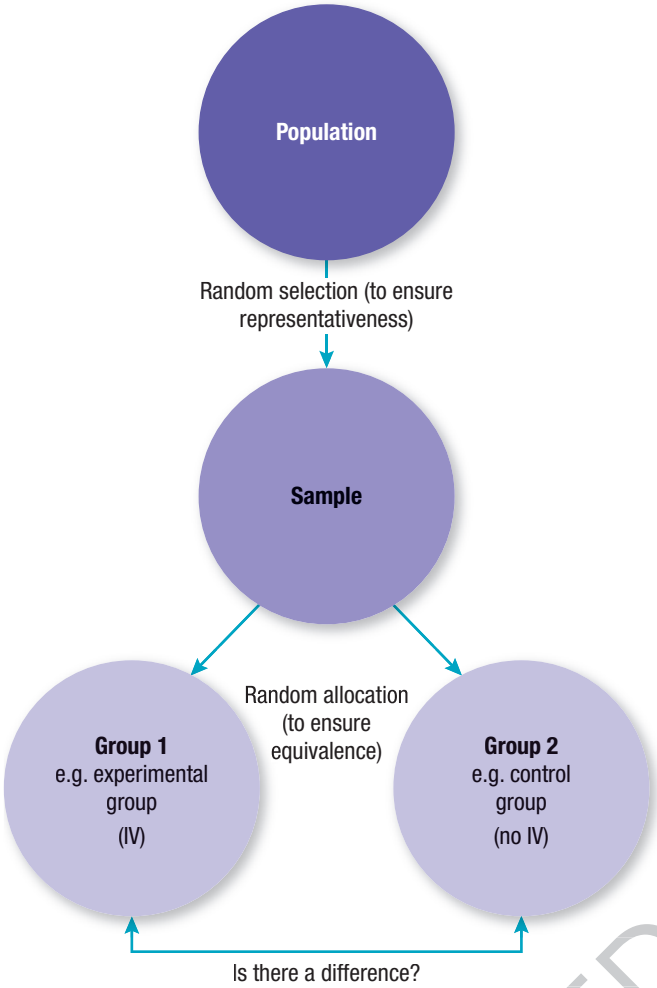


Figure 1.38 A simple experimental design using random sampling to select participants and random allocation to assign them to either condition

In contrast, random allocation is most related to the research design. It is used with the goal of being able to draw conclusions about the causal effect of the IV on the DV. After all, participants are randomly allocated in order to help ensure that the experimental and control groups are as similar to each other as possible (i.e. equivalent) prior to manipulation of the IV. Non-random allocation usually leads to non-equivalent groups, meaning that any change in the DV might be due to the groups being different in participant variables rather than the IV alone. Therefore, random allocation is most related to *internal validity*.

The consequences of random selection and random allocation are clearly very different, and a good research design will use both whenever possible to help ensure both internal and external validity (see pages 108–9).

LEARNING ACTIVITY 1.16

Review questions

1. What is the main difference between participant selection and participant allocation?
2. What is random allocation?
3. What does random allocation achieve?
4. Why is random allocation considered to be a crucial feature of a psychological experiment?
5. How are random sampling and random allocation different?
6. How do random sampling and random allocation relate to the validity of an experiment?

LEARNING ACTIVITY 1.17

Evaluation of research

A VCE Psychologist wanted to investigate what effect, if any, a reward of money would have on doing a simple motor task. The student therefore devised the task of having participants put together small nuts and bolts. The 'score' consisted of the number of nuts and bolts put together in a three-minute period. The student non-randomly allocated a large number of conveniently-sampled, volunteer individuals into two groups, ensuring the groups were equivalent in size. The sample is shown in Table 1 below.

TABLE 1

Group	Males	Females	Total
A	22	18	40
B	15	25	40

The student researcher explained to the participants what to do. Then the student had Group A assemble the nuts and bolts, which were placed in separate containers next to each participant. At the end of the three-minute period, the student counted the number of units assembled by each participant. Next, the student repeated the same instructions to Group B, except Group B members were offered 50 cents for each completed unit. The results are shown in Table 2 below.

TABLE 2

Group	Mean number of units assembled
A	37
B	52

1. Identify the operationalised independent and dependent variables.
2. What conclusion do the results suggest about the effect of the IV?
3. Suggest two potential confounding variables and therefore alternative explanations of the results. Ensure that you clearly explain your choice of confounds.

UNCORRECTED PAGE PROOFS

Counterbalancing

A counterbalancing procedure is commonly used to control or 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 confounding variable and cannot be controlled or eliminated through other means. Experiments with a repeated measures design are most vulnerable to order effects (see page 56).

There are different types of counterbalancing procedures that vary in complexity. The simplest and most commonly used is called between-participants counterbalancing.

The *between-participants counterbalancing* procedure involves alternating the order in which the experimental and control groups are exposed to each condition of the experiment. Each group is exposed to each condition in a different order.

For example, suppose a researcher will conduct an experiment in which all participants first learn a list of words when rap music is playing and then learn a list of similar words when there is no music. It is possible that the participants may demonstrate better learning in the no music condition because of a practice effect. To address this order effect, the researcher could split the sample into two groups — A and B. Group A could learn words in the rap music condition first, then learn words in the no music condition. Group B would learn words in the no music condition, followed by the rap music condition. Participants would also be randomly allocated to each group to experience either condition first or second. The procedure is shown in Figure 1.39 below.

The results for all participants are then combined across the entire experiment to achieve counterbalancing. In this way, whatever order effects impact on learning the words are controlled. Although an order effect may have occurred for each participant, because they occurred equally in both groups, they have balanced each other out in the results.

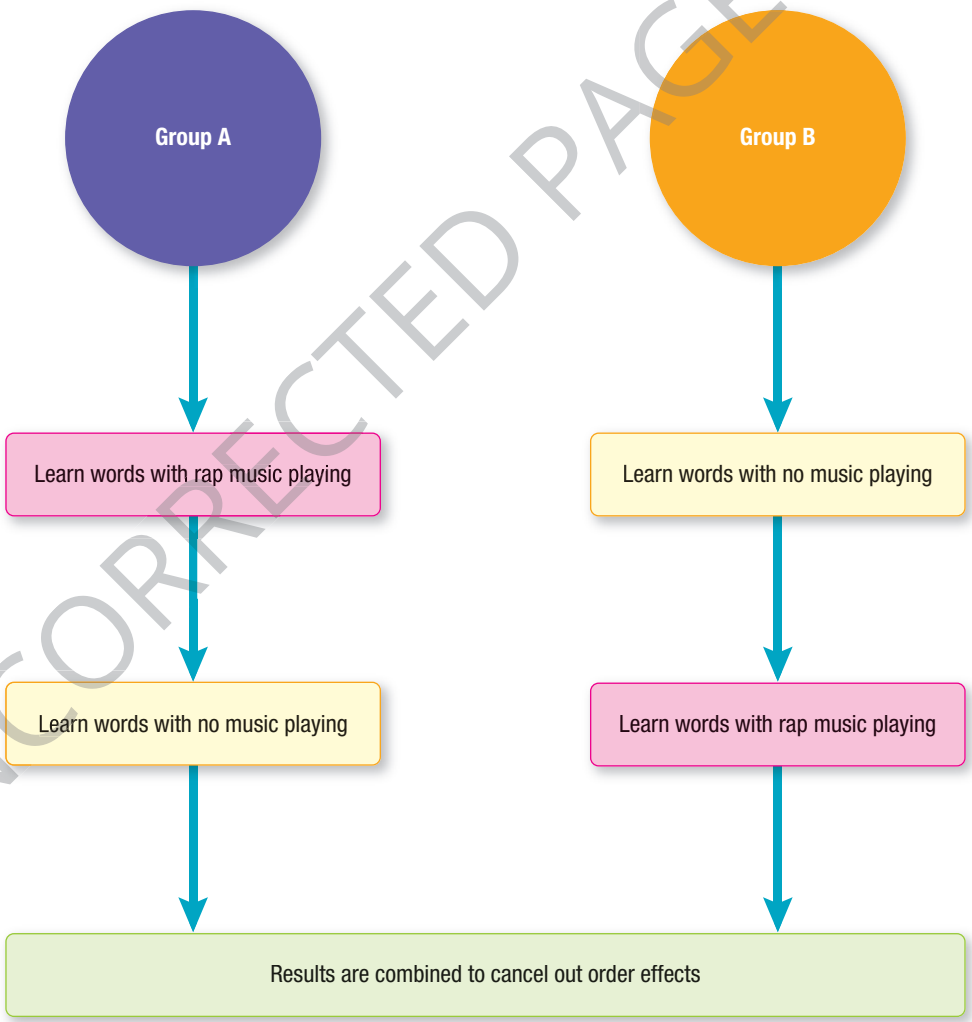


Figure 1.39 A simple counterbalancing procedure for the learning and rap music experiment. Table 1.3 on the next page and Table 1.4 on page 57 show other examples of counterbalancing.

UNCORRECTED PAGE PROOFS

BOX 1.10 Within-participants counterbalancing

The *within-participants counterbalancing* procedure requires each participant to be exposed to the same combination of conditions. That is, each participant is exposed to the treatment conditions in one order and then again in the reverse order.

The within-participants counterbalancing procedure called *ABBA counterbalancing* can be used when there are only two experimental conditions. In this form of counterbalancing, participants receive condition A first and then condition B, followed by condition B and then condition A. If more trials are desired, then this sequence is repeated.

For example, in the Pepsi and Coke taste-testing study described earlier on page 25, participants would rate one cola first followed by the second cola, then the second cola again followed by the first cola. In this way, any order effects are held constant so that their impact is balanced out over the entire experiment.

TABLE 1.3 An example of a within-participants counterbalancing procedure for a taste-testing experiment

Condition A (Pepsi)	Condition B (Coke)	Condition B (Coke)	Condition A (Pepsi)
P_1	P_1	P_1	P_1
P_2	P_2	P_2	P_2
P_3	P_3	P_3	P_3
P_4	P_4	P_4	P_4
P_5	P_5	P_5	P_5
P_6	P_6	P_6	P_6
P_7	P_7	P_7	P_7
P_8	P_8	P_8	P_8
P_9	P_9	P_9	P_9
P_{10}	P_{10}	P_{10}	P_{10}

Note: P = participant

Single and double blind procedures

Participant beliefs and expectations (including the placebo effect) can influence the results of any investigation, 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). Note, however, that single blind is not always possible. For example, in an experiment assessing which type of psychotherapy is most effective, it would be impossible to keep participants unaware of whether or not they actually received some type of therapy.

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 conditions. 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 the researcher(s) 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 affect the behaviour of the researcher as well as the participants; for example, when testing the effects of a drug. In drug testing studies, called 'clinical trials', use of double blind is a standard procedure. However, as with single blind, double blind is not always possible.

In some experiments the participants may know which condition they are in and one or more researchers do not. For example, a researcher conducting a vital data collection, recording or assessment procedure may intentionally be kept unaware of the condition to which participants are allocated in order to avoid bias or some other 'experimenter effect' on participant performance or results. In such cases, the researcher is said to have been 'blinded' or 'blinded to' the condition, so it is considered that single blind has been used.

Figure 1.40 This 10-year-old boy is a participant in a double blind, oral food trial to test for a food allergy. The boy is required to eat a variety of foods one by one, while being observed for symptoms of an allergic reaction by a doctor. The foods include suspected allergens and placebos. Neither the patient nor the doctor knows which foods are which.

UNCORRECTED PAGE PROOFS

Placebos

In an experiment, participants in the experimental group are exposed to the treatment (the IV) and participants in the control group are not. Because only the experimental group receives the treatment, only the participants in this group may be influenced by their expectations or beliefs 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 treatment or because of their expectations of how they should behave.

For example, suppose an experimental group is given an alcoholic drink so that its effects on performance of a task can be observed, whereas the control group receives nothing. Impaired performance observed in the experimental group may be due to the alcohol, or it may have arisen because the act of giving the participants alcohol suggested that they were expected to act drunkenly, so they did.

In order to control this potential confounding variable, control groups can be given a placebo, or fake treatment, so that they form the same expectations and beliefs as the experimental group. Thus, the control group would be given a drink that smells and tastes like alcohol but is not alcohol. The control group would not be informed that their drink is not alcoholic and they would have no way of distinguishing it from a real alcoholic drink.

Using this procedure, both groups will form the same expectations for acting drunkenly, so any differences in performance can be assumed to be due to the real alcohol given to the experimental group. Similarly, when testing other drugs, researchers give placebo pills, substances or even 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).

A placebo can be any type of inert (inactive) or fake treatment. It may be a drug or any other type of pill or substance, an edible product such as a food, a special diet, a psychological therapy, a physical therapy, exercise or even surgery (such as an incision and a procedure that is faked so that the participant doesn't know they actually had nothing done).

When a placebo is given to a control group, the group is sometimes referred to as the *placebo control group* or the *placebo condition*. The use of placebo treatments, including ethical considerations, is discussed further in Chapter 11.

UNDER THEIR PAGE PROOFS

Figure 1.41 The inability of many astronauts to sleep well when on space missions led to research designed to test whether they could be helped by taking melatonin, a hormone known to have a role in sleep onset. In one experiment, half the astronauts aboard a space shuttle took a pill containing melatonin, and the other half took a placebo pill that looked the same but did not contain any active ingredient. All astronauts were blind as to which experimental condition they were in — they did not know whether they had taken the melatonin or placebo. This photo shows space shuttle astronauts in their sleeping bags. Their arms are floating free but their bodies are restrained and kept in place.

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. More specifically, the goal is to minimise any differences among participants that might occur within the experiment itself.

This is achieved by standardisation ('uniformity') 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** means that the directions and explanations given to all participants in each condition are predetermined and identical in terms of what they state and how they are administered. They should be clear and avoid jargon, and there should be no ambiguities or variations for individual participants. Generally, the researcher should describe the sequence of events, identify the stimuli participants should attend to and explain how to respond. Questions by participants should be anticipated and the specific answers or type of response to be given by the researcher should be predetermined.

The use of standardised instructions starts with participant contact and includes the consent form. To reliably present instructions during an experiment, 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 during the experiment, beginning with greeting the participants and ending with the debriefing.

A key purpose of using standardised instructions is to have all participants perform the intended task in the same way to avoid introducing potential extraneous variables that make the task inconsistent or different for different participants. For example, if the researcher must stop during testing to further explain a task or correct a behaviour, the researcher's actions can become an inconsistency.

It is also essential that all participants are exposed to the same environment and procedures, with the only exception being exposure to the independent variable. Therefore, it is also necessary to use **standardised procedures** — the techniques used for making observations and measuring and recording responses should be identical for all individual participants. All participants should be treated in the same way, as appropriate to the experimental condition to which they have been assigned

CORRECTED PAGE PROOFS

Figure 1.42 A researcher conducting a study on food cravings during pregnancy. The participants are shown sampling foods from a selection on a table, and notes are being made by a researcher (upper left) of what they choose to eat. The researcher's use of standardised instructions and procedures minimises the influence of unwanted participant variables and experimenter effects.

(e.g. experimental or control group). For example, using standardised procedures often means that:

- all participants will be organised and brought together in the same way
- all participants would interact with the same researcher in the same environment
- the experiment would be run at the same time of day for all participants
- all participants would have the same amount of time, learn the same amount of information and complete the same activities (except for variations required for IV exposure)
- all researchers will treat each participant in the same way when conducting the study (except for IV manipulation).

How the researcher presents a stimulus to initiate a response of interest and record the resulting physical reaction or a score during DV measurement can often be controlled through automation by using electronic or mechanical devices. Electronic timers, data projectors, video and audio recorders, computers, tablets and other digital technologies can help ensure controlled and consistent stimulus presentations (such as an image on a monitor to measure reaction time). Automating data collection ensures that the response recording and scoring systems are consistently and accurately applied. It can also provide for sensitive measurement of responses (such as the keystroke or screen tap used to measure reaction time to a stimulus).

The use of standardised instructions and procedures can help control unwanted participant variables and the placebo effect, because all participants will have the same experience. It can also help control any experimenter effect, 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.

LEARNING ACTIVITY 1.18

Review questions

1. Explain what the counterbalancing procedure is and which potential problem(s) it attempts to control.

2. (a) A researcher believes that the biological sex of participants is a potential confounding variable. Explain how counterbalancing could be used to control this variable.

(b) The researcher will use a number of research assistants to conduct the study and also believes that their sex is a potential confounding variable. Explain how counterbalancing could be used to control this variable.

3. Suggest a randomisation or random allocation procedure that could be used to counterbalance order effects in an experiment.

4. (a) In what way is the single blind procedure similar and different to the double blind procedure?

(b) Which of the two procedures gives more control? Explain your answer.

(c) Give an example of when it would be essential for participants *and* a researcher to be blinded to the condition to which participants are allocated.

5. Explain how a placebo can be used to control or minimise the influence of expectations or beliefs associated with the specific treatment received by participants in experimental and control groups.

6. (a) What are standardised instructions and procedures? Explain with reference to relevant examples.

(b) Explain how standardised instructions and procedures can be used to control or minimise the influence of participant variables and experimenter effects.

LEARNING ACTIVITY 1.19

Evaluation of research

Mardi conducts an experiment to find out if colour preference can be influenced by associating a colour with a pleasant experience such as eating. She delivers a supply of red, orange, yellow, green and blue feeding bottles to some mothers of newborn infants and the regular transparent feeding bottles to the mothers of other newborn infants in the sample. The mothers have consented to let their infants be participants in Mardi's experiment.

1. What is the IV in Mardi's experiment?

2. How many experimental groups does Mardi have in her experiment?

3. Which participants make up the control group?

4. What is the DV?

5. How could Mardi randomly allocate the participants to different groups?

6. Identify two extraneous or potential confounding variables that should be controlled and explain why.

Source: Adapted from Sugar, J., & Offir, C. (1990). *Student resource manual* (3rd ed.). New York: Harper & Row.

LEARNING ACTIVITY 1.20

Reflection

To what extent are non evidence-based 'alternative medicines' (such as faith healing and herbal products) reliant on the placebo effect? Comment with reference to an example.

UNCORRECTED PAGE PROOFS

Use of an appropriate experimental research design

Various experimental research designs can be used to minimise the effects of potential extraneous and confounding variables, particularly variables associated with individual differences of participants. Three of these designs are the independent groups, repeated measures and matched participants designs. The designs primarily differ in how participants are allocated to different groups (or conditions).

Independent groups

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

The simplest independent groups design uses two groups — most often one group as the experimental group and the other as the control group. The experimental group is exposed to the IV (some kind of change) and the control group is not. For example, suppose a researcher is interested in investigating the effects of loud music (an IV) on problem solving (a DV). The experimental group could be given a problem-solving task to complete while loud music is playing and the control group would be given the task to complete without any music playing. Performance of participants under each condition would then be compared with reference to scores achieved on the problem-solving task. For example, a mean score could be calculated for each group to enable a quick comparison.

Random allocation is an essential feature of the independent groups design in order to control the influence of individual participant differences. This procedure is used after participants have been selected for the experiment, but before the experiment begins. Depending on the number of participants and groups, the random allocation procedure may involve tossing a coin to decide which participant goes into each group. Alternatively, a simple lottery method can be used. For example, participants may be asked to draw a number from a container holding as many different numbers as there are participants. Those who draw odd numbers are then allocated to one group, and those who draw even numbers are allocated to another group.

Using an appropriate random allocation procedure is likely to result in two groups that are well matched on participant characteristics and therefore fairly equivalent. However, this may not necessarily be the case if there are a very small number of participants. 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. A significant practical advantage is that the experiment can usually be completed on one occasion, which also helps ensure participant attrition (loss) is negligible. Unlike the repeated measures design, there is not often a need to spread out the time period between the different conditions. Another advantage of the independent groups design is there are no order effects between conditions to control.

However, there is often a need for a larger number of participants to help ensure the spread of participant variables within the sample will match the distribution within the population. In addition, there is less control over participant variables than in the repeated measures and matched participants designs, especially when a small sample is used.

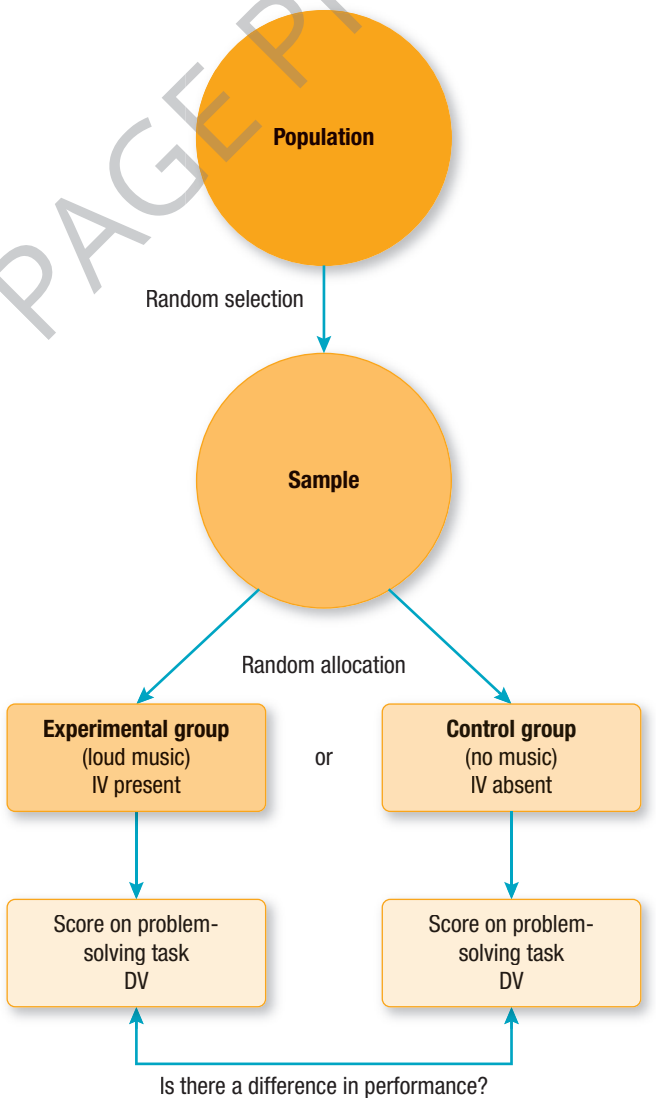


Figure 1.43 In an experiment with an independent groups design, participants are randomly allocated to the different groups (conditions) to ensure participant variables that can cause a change in the DV are uniformly distributed across the groups.

UNCORRECTED PAGE PROOFS

Repeated measures

Another way of controlling the influence of individual differences of participants is to design an experiment that uses the *same* participants in the experimental and control conditions. This is what happens in a repeated measures experiment, which is also called *within participants* and *within groups*.

In an experiment with a **repeated measures** design, each participant is in both the experimental and control groups (and therefore all conditions). The groups are identical so individual participant differences that may impact on the DV are controlled.

For example, consider the researcher interested in loud music and problem solving. Using the repeated measures design, a group of participants would be given a problem-solving task to complete while loud music was playing, and the same group would then be tested on a similar, equally difficult, problem-solving task but without any music playing. This means each participant would experience both the loud music and no music conditions while solving similar problems. How well all participants solved problems would be measured twice, once after each condition of loud music and no music respectively (hence the term 'repeated measures').

Using this design would give the researcher strict control over all relevant participant variables that can influence the results, such as individual differences in problem-solving abilities and levels of motivation. Any participant differences that may not have been identified by the researcher as potential confounding variables have also been controlled as the participants in both conditions are identical in every respect.

When planning a repeated measures experiment, the researcher has to consider order effects that may arise from the design; for example, whether a problem-solving task is performed first or second. Performance on a problem-solving task that is completed second may be better because of the experience gained in completing the first task. Participants may perform better in the second condition because they have practised the task or have gained other useful knowledge about the task or the experiment. Alternatively, participants' performance may be impaired by effects such as fatigue or boredom, and they may not perform as well on the second occasion. In either case, the order effect is a potential confounding variable because the researcher cannot be confident about whether the IV or order effect caused the change that was measured 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 (in this case, completion of the problem-solving tasks). For example, participants might be in the experimental group one day, then return a week later for the control condition. When this procedure is inappropriate, inconvenient or impractical, the researcher should ensure that the order in which the problem-solving tasks are performed is counterbalanced across the participants. For example, half the participants would follow one order (solving a problem in the experimental condition first, then solving a problem in the control condition). The other half would follow the reverse order (solving the problem in the control condition first, then solving the problem in the experimental condition).

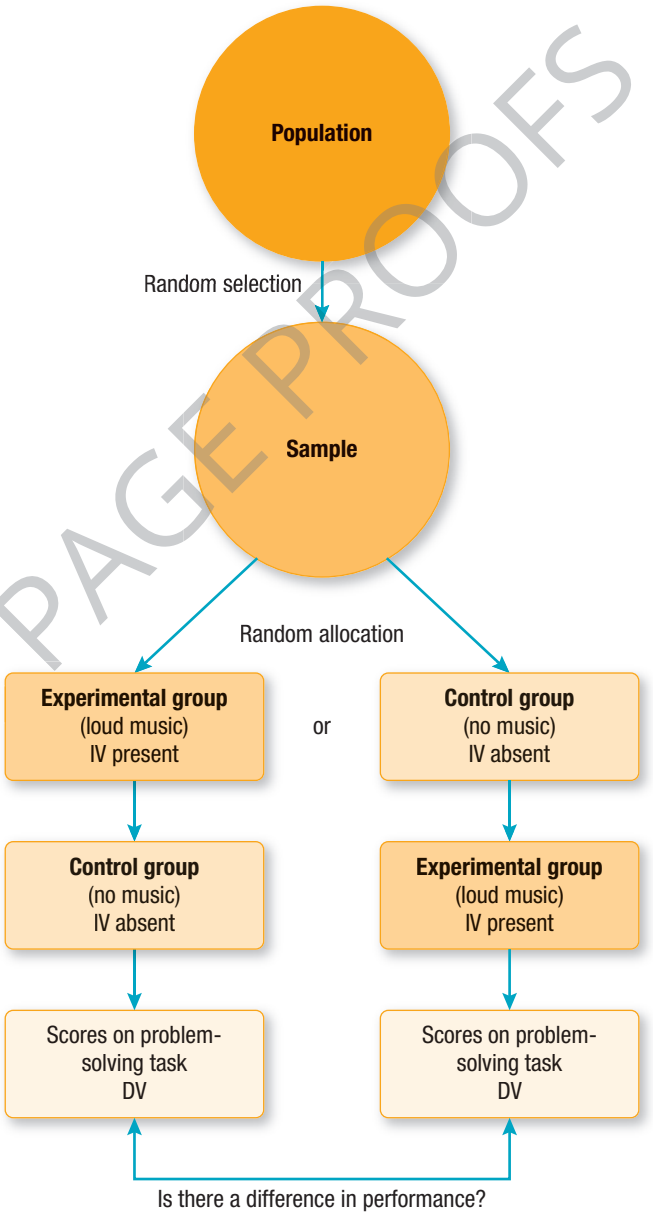


Figure 1.44 In an experiment with a repeated measures design, the same participants are in both the experimental and the control groups and are therefore exposed to all conditions.

Even stricter counterbalancing control of order effects may be achieved by using a random allocation procedure to determine the order in which each participant will be exposed to each condition. These procedures usually ensure that order effects are balanced out over the entire experiment, but there

UNCORRECTED PAGE PROOFS

is no guarantee that this will occur. Table 1.4 below shows a counterbalancing arrangement whereby individual participants are randomly allocated to the different conditions (also see Figure 1.39 on page 50).

TABLE 1.4 Counterbalancing for a repeated measures experiment

Participant	Experimental condition	Control condition
1	first	second
2	second	first
3	first	second
4	second	first
5	second	first
6	first	second

The main advantage of the repeated measures design is that it can effectively control potential confounding variables arising from individual participant differences. The researcher can assume that any difference in performance on the DV in each condition of the experiment should not be due to differences produced by any extraneous variable associated with individual participant differences because each participant, with the same personal characteristics (including abilities), is in every condition. Another advantage is that this design also tends to require a relatively smaller number of participants when compared with other experimental designs because the same participants are in all conditions.

However, the repeated measures design has several limitations. Although this design keeps individual participant differences constant, it does not necessarily control all participant variables that can influence the results. For example, some participants may guess what the experiment is about as they compare the two conditions, creating expectations and beliefs that lead to unnatural responses.

The repeated measures design can also result in unwanted participant loss before the experiment is completed. It is most common when the repeated measurement of the DV requires a considerable amount of time per participant, so that, to reduce fatigue or overload, the researcher spreads out the time between the different conditions over several days. Then, participants show up for the first session but do not return for the second one. It is also possible for some participants to find the first condition boring and not attend the second simply because they don't want to.

Another limitation is that order effects are more likely with the repeated measures design and can become a confounding variable if uncontrolled. However, the researcher can use counterbalancing to control any order effect.

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 to the experiment. This type of experiment usually involves selection of pairs of participants who are very similar in one or more personal characteristics that can influence the DV, then allocating each member of the pair to an experimental or control group. When pairs of participants are matched, the design is often called *matched pairs*. However, the matched participant design is not limited to pairs of participants. Several well-known experiments have used matched 'trios' (or 'triplets'). One of the best known is a series of experiments by Albert Bandura to investigate observational learning of aggression by young children (see pages 318–19). An experiment may also use matched quads and so on, if required.

The use of a matching process in terms of one or more relevant participant variables (such as intelligence, creativity, sex, age) is reflected in the name of the experimental design. Randomly allocating one member of each matched group to a different condition ensures that each experimental and control group is fairly equivalent in terms of the spread of participant characteristics that can cause a change in the DV.

Sets of identical twins are best for matching participants when pairs are used. They have the same genetic make-up, are identical in age and sex and tend to be very similar (but not identical) in socio-cultural background, mental abilities, temperament and various other personal characteristics (but not all). In a matched participants experiment, one member of each pair of twins would be allocated to the experimental group and the other would be allocated to the control group. In this way, both groups would be considered fairly equivalent, thereby minimising extraneous and confounding variables. However, identical twins are often not available.

In the loud music and problem-solving experiment, the intellectual ability of each participant could be reasonably assumed as being likely to affect their problem-solving ability. Suppose that the experimental group (loud music condition) performed poorly on the problem-solving task, as compared to the control group (no music condition). The researcher would want to be in a position to conclude that this difference in performance was due solely to the IV (loud music). If the experimental group had all the participants who were least intellectually able and the control group had all the participants who were most intellectually able, the experimenter would not know whether it was the loud music or the problem-solving ability of participants that caused the poor performance.

In order to avoid this problem, the researcher could administer an intelligence test to each participant after they had been selected for the experiment,

UNCORRECTED PAGE PROOFS

but before the experiment began. This pre-testing would provide an IQ score for each participant. Each participant would then be paired with someone else with a similar IQ score until all the participants had been matched on intelligence. The participants would then be allocated to a group (experimental or control condition) on the basis of their IQ scores.

For example, in allocating participants to groups in the loud music and problem-solving experiment, the two participants with the highest IQ scores would be randomly allocated to the loud music and no music groups respectively. Then the two participants with the next highest IQ scores would be randomly allocated to the two groups, and so on. In this way, the two groups in the experiment would be *matched* in terms of intellectual ability, thereby controlling the influence of this potential confounding variable.

The loud music and problem-solving experiment could also have three groups — a control group who will not be exposed to any music, an experimental group who will be exposed to loud music and an experimental group who will be exposed to soft music. For this specific design, participants would be organised into groups of three based on IQ score matching, then each member would be allocated to one of the three experimental conditions.

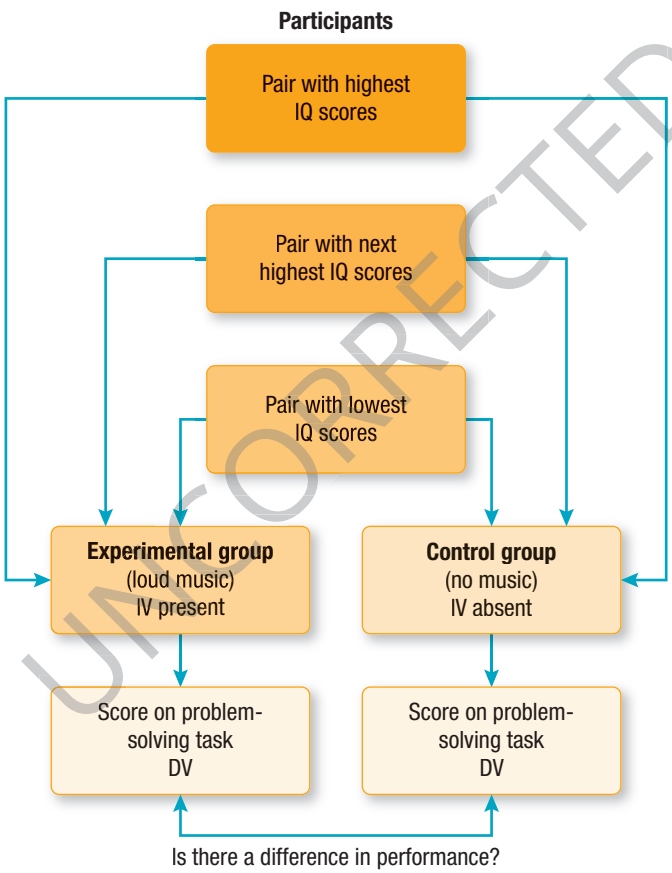


Figure 1.45 Matched participants design for the experiment on loud music and problem-solving

The main 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 constant across the conditions, thereby eliminating them as potential confounding variables. In addition, like the independent groups design, the experiment can usually be completed on one occasion. Furthermore, unlike the repeated measures design, there are no order effects between conditions to control. Participant loss is also 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 the difficulty of knowing which specific participant variables should be matched. Even if the researcher can identify the variable(s) that is likely to most influence performance on the DV, it is often difficult and time-consuming to actually recruit participants who are sufficiently alike in the variable. 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. And the more matching variables that are specified, the more difficult and time-consuming it is to match.

As well as being time-consuming, pre-testing can create order effects. And the loss of one participant through attrition means the loss of a whole pair, trio and so on. Pre-testing is usually not required for the repeated measures design and is only used in an independent groups design 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 commonly used.

TABLE 1.5 Three experimental designs

Experimental design	Feature
independent groups	Each participant is randomly allocated to either the experimental or control group and is in one group (condition) only.
repeated measures	Each participant is involved in both the experimental and control groups i.e. all conditions.
matched participants	Each participant in one group (or condition) 'matches' a participant in the other group(s) on one or more participant variables relevant to the experiment.

UNCORRECTED PAGE PROOFS

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 or confounding variables to concentrate entirely on the effect the IV has on the DV. Elimination of all unwanted and irrelevant variables is not always possible, but their control is usually greater than in other research methods, especially if the experiment is conducted in a laboratory setting. Therefore, the ability to control variables, often in a very precise way, is an important advantage of an experiment when compared to other research methods.

Another important advantage of the experiment is that the IV can be manipulated in order to observe and measure the effect on the DV, therefore making it possible to test if there is a cause and effect relationship between the IV and DV. In this way, experimental research has the principal advantage of enabling conclusions about cause-and-effect relationships between variables. Furthermore, because controlled conditions are known conditions, the experimenter can set up the experiment a second time and repeat it to test the results.

Alternatively, the experimenter can report the conditions of an experiment in such a precise way that others can replicate it 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 research and its results.

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, this type of setting 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 (see Box 1.11 on the next page).

Furthermore, some things cannot be measured in a laboratory. The researcher cannot break up families, for example, to measure the effects of family separation. It is not possible because of practical realities and ethical concerns. 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.

COPYRIGHTED PAGE PROOFS

Figure 1.46 Would practical realities and/or ethical concerns limit an experimental research study of responses to an emergency situation by an audience in a real-life setting? If so, how could these limitations be overcome?

BOX 1.11 Artificiality and ecological validity of laboratory experiments in psychological research

Psychologists often conduct experiments in laboratory settings, usually located at a university, so that the environment and procedures can be controlled and the participants' responses to the IV can be carefully observed. Volunteer participants of any age may be brought into a research laboratory to study virtually any aspect of human thinking, feeling and behaviour.

Some laboratories are quite elaborate. For example, a sleep laboratory has a diverse range of technical equipment that enables researchers to monitor a participant's eye movements and brain wave patterns, record the exact time they fall asleep and get dream reports the moment they awaken. Similarly, a laboratory for studying infant–parent social interactions may be set up as a special playroom equipped with hand-picked toys, two-way mirrors, and a hidden camera and microphone so that researchers can record every word that is spoken and analyse every little detail of each participant's behaviour.

Thus, in order to make certain observations, psychologists often find it necessary to simulate situations or events in a laboratory. However, the laboratory is an artificial setting. It is a humanmade, 'non-natural' setting used for research purposes (but can be basic as a classroom or lecture theatre).

This type of environment can produce responses that are distorted or do not adequately resemble how people would 'naturally' think, feel or behave when in the 'field' or real-life settings. For example, can someone sleep normally in a strange bed with metal electrodes pasted to their scalp? Will a parent and child interact in the playroom the way they do at home (Kassin, 1995)?

Laboratory-based research in psychology is often criticised because of its *artificiality* — its lack of realism and differences to real-life settings. The artificiality of the environment in which a study takes place can produce cues called *demand characteristics* that cause participants to react unnaturally. For example, participants sometimes use such things as random noises, changes in lighting or a broken pencil point as cues to work out what is being studied and how they should respond. More often than not, the conclusions drawn by participants are misleading or wrong (see Box 1.7 on page 34). Furthermore, artificiality can limit the generalisability of the results from the laboratory setting to real-life contexts. This means that a study conducted in a laboratory setting may be lacking in ecological validity and therefore external validity. *Ecological validity* is the extent to which an experiment's results (or any research findings) can be generalised to everyday, common real-life behaviours and natural situations.

However, too much emphasis may be placed on the setting in which a study is conducted. Many variables are impossible to study in the field or a real-world context for ethical and practical reasons. For example, Stanley Milgram's controversial studies in the 1960s on obedience to authority were criticised on the basis that they were conducted in a laboratory-based environment at a university. Participants knew they were in an experiment and, due to the artificial nature of the environment, may have responded differently to how they would have acted in similar circumstances in real life. For example, they may not have administered electrical shocks at deadly levels, if at all. Milgram recognised the limitation of his university laboratory setting by conducting further studies in a more realistic setting (a shopfront location in the community outside the university). He subsequently obtained a similar overall pattern of results to those observed in the 'artificial' laboratory setting. Such a study would be virtually impossible to conduct in a truly real-life context, and although the study was ethically questionable, it was better to examine the type of obedience involved in a controlled environment.

Figure 1.47 Stanley Milgram's controversial studies on obedience to authority were criticised on the basis of their artificiality as they were conducted in a laboratory-based environment. The photo shows an obedient research participant administering a shock he believes is potentially lethal.

UNCORRECTED PAGE PROOFS

BOX 1.12 Correlational research

In an experiment, the researcher manipulates one or more IV(s) in order to establish whether a change in the IV(s) brings about a change in the DV(s). If strict experimental control is maintained, a cause–effect relationship between an IV(s) and DV(s) can usually be established. Sometimes, however, using the experimental method is impractical or inappropriate.

Suppose a researcher wanted to find out how a severe emotional 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 the harmful experimental condition of the particular severe emotional trauma. In such cases, the researcher may choose to use existing information in order to assess the relationship, or correlation, between the variables of interest.

Correlational research is used to identify and describe the 'co-relationship' between two (or more) variables of interest. No attempt is made to manipulate any variable, as in experimental research. Nor is there any random allocation to conditions. The researcher merely assesses the type and strength of relationship between the variables.

The term *correlation* is a statistical one — a measure that indicates how variables are 'co-related'; for example, the relationship between stress and cancer, between level of anxiety and incidence of bedwetting, or between personality test scores and birth order. Correlation does not indicate that one variable (such as stress) causes another (such as cancer). Rather, it indicates whether a relationship exists, the strength of the relationship and the direction of the relationship (e.g. as one variable increases, is the other likely to increase or decrease?).

Direction of correlation

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

A *positive correlation* means that the variables vary (change) in the same direction — as the value of one variable increases, the value of the other variable also tends to increase. For example, as job satisfaction increases, work productivity tends to increase (and vice versa). Similarly, as one variable decreases, so does the other.

A *negative correlation* means that the variables vary in opposite directions — as the value of one variable increases, the value of the other variable tends to decrease (and vice versa). The variables change like a seesaw, rather than in tandem. For example, as the amount of alcohol in the blood increases, reaction time tends to decrease (and vice versa). And, as the number of cigarettes smoked daily increases, the number of years of life decreases.

It is said that there is a *zero correlation* or little or no relationship between two variables if the way that they vary is totally independent of each other. For example, there is no relationship between height and VCE grades or the amount of coffee drunk and IQ score.

A correlation is usually described by a number known as a correlation coefficient. A *correlation coefficient* 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 preceded by 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, if there is a high positive correlation (e.g. +0.75) between the rate of pupil dilation and problem-solving ability, then people with rapid pupil dilation should tend to be good problem-solvers (e.g. they would solve many problems in a 20-minute period). People with slow dilation would tend to be poor problem solvers (e.g. they would solve 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 (e.g. –0.75) between the rate of pupil dilation and problem-solving ability, then people with rapid pupil dilation would tend to be poor problem solvers and those with slow dilation 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 the variables; that is, whether the relationship is strong, moderate or weak. A correlation coefficient that is close to +1.00 indicates a high positive correlation (i.e. very strong relationship) between the variables. A correlation coefficient that is close to –1.00 indicates a high negative correlation (i.e. very strong relationship) between the variables. Correlation coefficients of +1.00 and –1.00 indicate perfect correlations, but these very rarely occur.

A correlation coefficient that is close to zero indicates little or no relationship between the variable. For example, 0.12 and –0.12 would be considered a weak positive and negative correlation respectively.

Correlation and causation

Correlations show the existence and extent of relationships between variables but they do not necessarily indicate a cause–effect relationship (that one variable causes the other); for example, as the world rotates on its axis, people get older.

(continued)

UNCORRECTED PAGE PROOFS

(continued from previous page)

Figure 1.48 Sleeping with shoes on and waking up with a headache have a strong positive correlation. Does this mean that sleeping with shoes on causes a headache? Of course not. Both variables also correlate with and are affected by a third variable — going to bed after drinking too much alcohol.

UNCORRECTED PAGE PROOFS

There is an extremely strong correlation between these two variables but it would be faulty to assume that the Earth's rotation causes people to age or that people's ageing causes the Earth to rotate. Although scientists know that the two variables correlate, they have not been inclined to discover a way of stopping the Earth's rotation so that we stop getting older!

There are also many instances when high correlations suggest a logical cause–effect relationship, and sometimes correlations really do represent causal relationships. The amount of ink left in a pen is closely related to the length of time spent using the pen. But a significant 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 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.

Similarly, there is a very high correlation between the number of years spent in schooling and income as an adult. Both of these variables, however, have also been found to correlate not *only* with each other but *also* with a third variable — parents' income. When two variables are correlated, this is not accepted by psychologists as proof of causation in the absence of other research evidence. In such cases, researchers will test the possible cause–effect relationship by conducting a carefully controlled experiment.

Scatter plots

Correlational data are often displayed in a scatter plot (also called a *scattergram* or *scatter diagram*). A *scatter plot* is a graph of scores (or other values) on two different variables. The values of one variable are shown on the vertical axis (Y axis) and the values of the other variable on the horizontal axis (X axis). Each pair of scores is plotted as a single point (a dot) in the scatter plot.

The spread of the dots on a scatter plot gives an idea of the *strength* of the correlation — the extent to which the two variables are related. Widely spread dots in the scatter plot in Figure 1.49 on the next page suggest that the two variables, attractiveness and reaction time, are not related. This would be represented by a correlation close to 0 (zero). In a zero correlation, individuals with high scores on one variable may have high, middle or low scores on the other variable. Figure 1.49 shows that participants with high scores on attractiveness have high, medium and low reaction times.

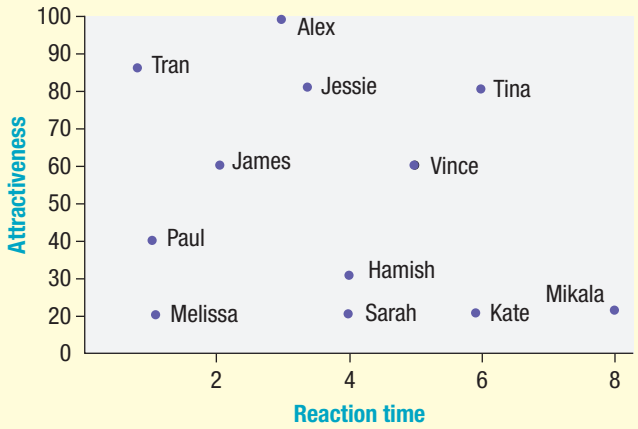


Figure 1.49 Scatter plot showing data for 12 participants obtained from a correlational study that investigated the relationship between physical attractiveness and reaction time

Figures 1.50 and 1.51 below both show a strong correlation as the dots cluster close together in a cigar-shaped pattern. Figure 1.50 shows a strong positive correlation and Figure 1.51 shows a strong negative correlation.

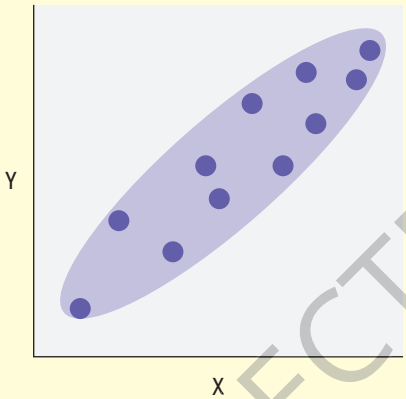


Figure 1.50 Strong positive correlation

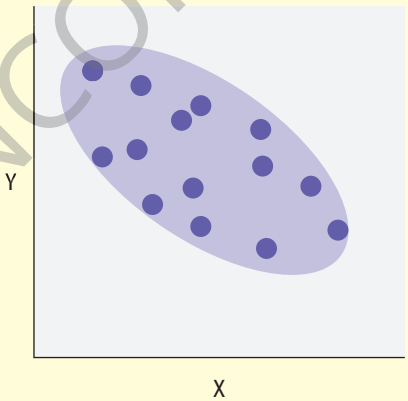


Figure 1.51 Strong negative correlation

The *direction* of the correlation — whether the correlation is positive or negative — is indicated by the slope or 'lean' of the dots, that is, whether they slope upwards or downwards (or neither). For example, in figures 1.52 and 1.53, a line has been drawn through the middle of the dots to help identify the slope. In Figure 1.52, the upward sloping line indicates a positive correlation, whereas the downward sloping line in Figure 1.53 indicates a negative correlation. Note that in both figures 1.52 and 1.53, the dots are closely clustered around each line, indicating a strong positive correlation in Figure 1.52 and a strong negative correlation in 1.53.

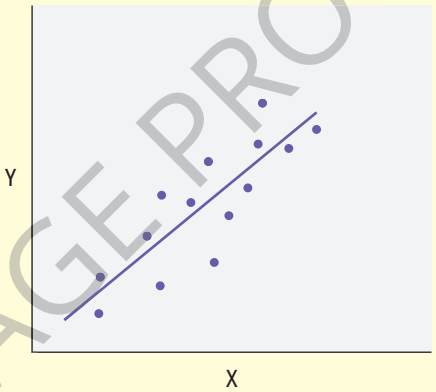


Figure 1.52 Strong positive correlation

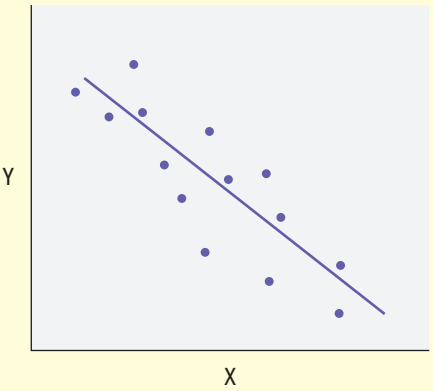


Figure 1.53 Strong negative correlation

eBook*plus*

Weblink
TEDx presentation: The danger of mixing up causality and correlation 5m 57s

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.21

Review questions

1. Make a copy of Table 1.5 on page 58. Add two columns, one headed 'Advantages' and the other 'Limitations'. Complete the table including at least two advantages and two limitations for each experimental design.
2. Explain the meaning of experimental design.
3. Which of the three experimental designs uses each participant as his or her own control?
4. Which of the three experimental designs targets comparison of a participant's behaviour both before and after exposure to a treatment?
5. Which of the three experimental designs exposes all participants to the same IV(s)?
6. Sometimes participants in an experiment may guess what the study is about, thereby producing unnatural responses. Which procedure would be used to control this extraneous variable?
7. Sometimes participants in an independent groups experiment may be naturally better at performing an experimental task. For example, in a study on attention, some people might happen to be better at concentrating than others. Which procedure would be used to control this extraneous variable?
8. For each of the following extraneous or potential confounding variables, rank the three experimental designs from 1 to 3 to indicate the extent to which the design controls the variable, as compared to the other designs. A ranking of 1 indicates greatest control and a ranking of 3 indicates least control. If appropriate, more than one design may be given the same ranking. Explain your choice of rankings.
   (a) individual participant differences
   (b) placebo effect
   (c) order effects
   (d) experimenter effect
9. Suggest two research topics or questions that you believe could not be investigated through experimental research because of ethical or practical concerns and explain each choice.

LEARNING ACTIVITY 1.22

Identifying and analysing experimental designs

1. Read the following summaries of experiments and indicate whether an independent groups, repeated measures or matched participants is used.

Experiment 1
A researcher is interested in the effectiveness of a particular treatment for insomnia. Fifty adult insomnia sufferers are contacted from a newspaper advertisement, and each is given a pill with instructions to take it before going to sleep that night. The pill actually contains milk powder (a placebo). The participants are randomly allocated to receive one of two instructions about the pill: half are told that the pill will make them feel 'sleepy' and the other half are told that the pill will make them feel 'awake and alert'. The next day, all the participants meet with the researcher and are asked how long it took them to fall asleep after taking the pill. The participants who were told the pill would make them feel sleepy reported having fallen asleep significantly faster than the participants who were told the pill would make them feel awake and alert.

Experiment 2
A researcher wants to examine the effects of massed practice versus distributed practice on the learning of nonsense words such as qoh, nal and fub. The researcher randomly allocates first-year university students studying psychology into one of three groups.

- Group 1 is required to learn a list of 20 nonsense words in one 90-minute session on one day.
- Group 2 learns the same list for 45 minutes per day for two successive days.
- Group 3 practises the same list for 30 minutes per day for three successive days.

The researcher assesses each group's performance with a test of free recall of the nonsense words after each group completes the designated number of sessions. The mean recall of the 20 words for Group 1 is 6.2; for Group 2, 11.1; and for Group 3, 14.9. These mean scores are found to be significantly different from one another, and the researcher concludes that distributed practice is more effective than massed practice.

Experiment 3
A researcher conducted an experiment to test a new medication for treating phobias. 347 volunteers ranging in age from 21 to 75 years responded to an advertisement. 24 males and 24 females who best met the selection criteria were included in the sample. The participants were grouped into pairs on the basis of two variables requiring control — sex and age. For example, Pair 1 was two females, both in the age group 21–25 years, Pair 2 was 2 males, both aged 21–25, and so on. Then, within each pair, participants were randomly allocated to either of two conditions — a treatment condition in which they used the medication, or a placebo control condition.

Experiment 4
A researcher studied how having previously seen an image of an object may influence the ability to name it again when it reappears later. Participants are first shown pictures of common objects such as a purse, a wristwatch and keys on a computer monitor. The participants then leave and return one week later. At this time, they are shown some of the original pictures they had seen in the first session, some similar but not identical pictures, and some entirely new ones. They are then asked to name the objects as quickly as possible. The researcher found that the original objects were named significantly faster than the new objects, but that the similar objects were named more slowly than the new ones.

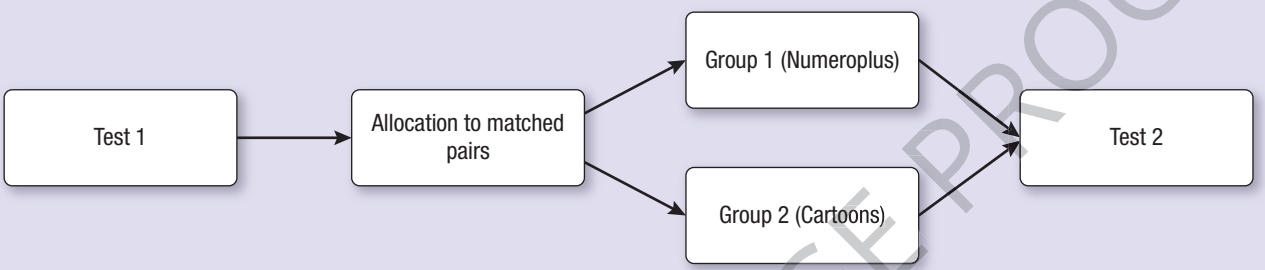
UNCORRECTED PAGE PROOFS

2. Read the following summary of an experiment and answer the questions about the design.

A researcher conducted an experiment to test the effectiveness of Numeroplus – a new teaching program for developing numeracy skills in young children. There were 24 Year 2 participants from a local primary school whose parents had given written consent. All completed a numeracy test (Test 1). The results were used to organise participants into 12 matched pairs based on the similarity of their scores. A member of each pair was then allocated to either Group 1 or 2. Each participant had an equal chance of being in Group 1 or 2. Group 1 undertook the Numeroplus program every school day over a two-week period. This was presented through a 20-minute video tutorial on a TV at the front of the classroom. Group 2 watched cartoons on a TV for 20 minutes, every school day for 2 weeks. All participants then completed a second numeracy test (Test 2).

The flow chart shown below summarises the procedure.

(a) Explain why the researcher placed the participants into matched pairs.
(b) What was the purpose of Group 2?
(c) Identify the operationalised independent and dependent variables.
(d) Describe two features of the procedure that identify this study as an experimental design.
(e) What is an advantage of using an experimental design for this investigation?



Source: Adapted from Stangor, C. (1998). *Research methods for the behavioural sciences*. Boston, Massachusetts: Houghton Mifflin.

LEARNING ACTIVITY 1.23

Evaluation of research

An experiment was conducted to compare the effectiveness of learning online and learning from a textbook. The participants were 70 volunteer first-year psychology students. After arriving at the experimental room, each participant was given written instructions describing the task and learning materials. These instructions were identical except for reference to either a computer or text for learning.

Participants were then randomly allocated to either of two groups. Group 1 went to an adjacent computer lab where they were each assigned to a computer in separate learning compartments and worked individually to learn new psychology content that was presented online in 20 screens over a 40-minute period. Group 2 members went to an adjacent classroom, remained together and learnt the same psychology content for the same amount of time. However, the computer screen content was in a hard-copy format, presented as a printed booklet.

Both groups then returned to the experimental room for a test on the content. This was completed using a standard pen and paper format. The results showed that Group 1 performed significantly better than Group 2. The researcher concluded that online learning is more effective than textbook learning.

1. Formulate a research hypothesis that would be supported by the results obtained for the experiment.
2. Give an advantage and a limitation of the sampling technique used for this particular experiment.
3. Identify the experimental design.
4. Identify the operationalised independent and dependent variables in the experiment.
5. Briefly describe the conclusion that was made on the basis of the results that were obtained.
6. Identify a potential confounding variable and explain why it may provide an alternative explanation of the results.

LEARNING ACTIVITY 1.24

Designing an experiment

Choose a specific topic of interest from Unit 3 or 4 and outline an experiment with an independent groups design, an experiment with a repeated measures design and an experiment with a matched participants design that could be conducted on your chosen topic.

Present the experiments using flow charts (see figures 1.43–1.45) so that the three designs can be compared.

UNCORRECTED PAGE PROOFS

Cross-sectional studies

A **cross-sectional study** selects and compares different groups of participants on one or more variables of interest at a single point in time. In this way, it provides a 'snapshot' of mental processes or behaviour in relation to the variables included in the study. A cross-sectional study is often used to study age-related differences. For example, a study might investigate how performance on a memory task can vary according to age. Groups of people selected at 10- or 20-year intervals from 5 to 85 year olds (i.e. a 'cross-section' of ages) could be tested and the results compared.

A cross-sectional study may also be used to study differences between groups in one or more other variables at a particular point in time. For example, a study may compare the recovery of a group of stroke patients who participated in a specific rehabilitation program with a group who did not, or it may be used to study the behaviour of students in three year seven English classes, each of which is taught by a teacher with a different approach to classroom management and discipline. The choice of samples and variables for study is virtually endless. Samples of participants may be selected on the basis of their school type, year level, course, occupation, attitudes, mental health status, mental abilities, sleep habits, exercise habits, parental style, personality, cultural background, social media use, use of a specific medication, and so on. However, the data are collected all at the same time (or within a short time frame).

A cross-sectional study may use an independent groups design and is sometimes called a *quasi-experiment* because of this resemblance to an experiment. However, it is not a true experiment because participants cannot be randomly allocated to experimental and control groups. Instead, a cross-sectional study uses existing, naturally formed or occurring groups. For example, 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 measures characteristics or events that already exist or occur naturally in a sample (or population) without directly manipulating any variable(s).

In sum, in a cross-sectional study, the independent variable of interest occurs naturally and the researcher takes advantage of pre-existing variables or conditions such as a participant's age, or an event that the researcher has no control over.

A cross-sectional study may be repeated periodically to study a trend or pattern. It may therefore also involve 'repeated measures', but this does not mean it is the same as a repeated measures experiment.

Advantages and limitations of cross-sectional studies

One advantage of a cross-sectional study is that it can be used to efficiently study change over time. Compared to other research methods, it tends to be simpler to undertake, less 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.

Another advantage is that a cross-sectional study provides a means of conducting research on certain topics that are unethical and/or impractical to conduct through experimentation. For example, to study the effects of exposure to a natural disaster on mental health, the researcher could access one or more groups who have been exposed to a specific type of natural disaster and assess their mental health. Similarly, to study maternal risk factors for low birth weight infants, a research team could access mothers who have recently given birth and collect data to compare maternal environments and behaviours. Four groups of mothers could be used, based on World Health Organization definitions of 'birth weight'. For example, 'low birth weight' is defined as a birth weight of a liveborn

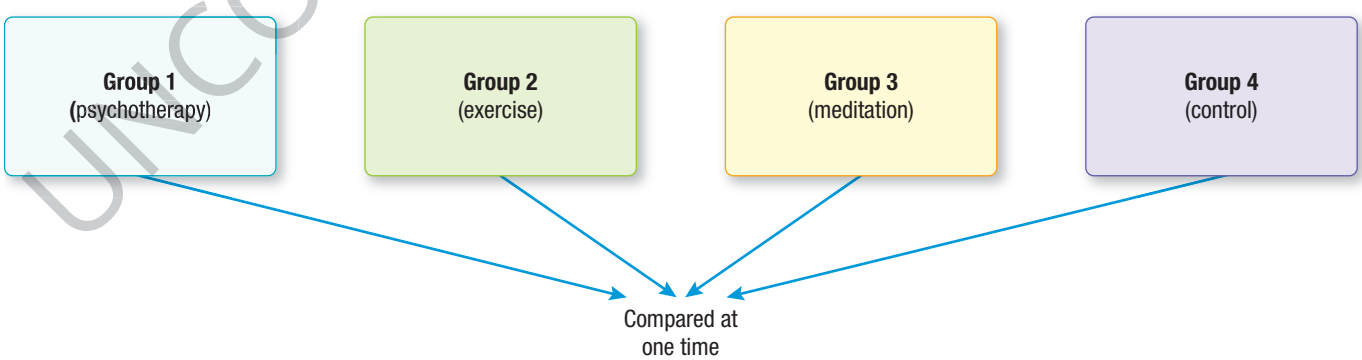


Figure 1.54 An example of a cross-sectional study that uses an independent groups design to enable comparison of one or more variables in existing groups at a single point in time. This study is comparing outcomes associated with three different stress management programs. There is a fourth group who do not engage in any stress management program.

UNCORRECTED PAGE PROOFS

infant of 2499 grams (5 pounds 8 ounces) or less, 'very low birth weight' is less than 1500 grams (3 pounds 5 ounces), and 'extremely low birth weight' is less than 1000 grams (2 pounds 3 ounces). A fourth group could be mothers who gave birth to an infant with a 'normal birth weight', defined as 2500–4200 grams (5 pounds 8 ounces–9 pounds 4 ounces).

Cross-sectional studies also tend to have a relatively low level of participant attrition (especially when compared to repeated measures experiments). Since data are collected all at once, it is less likely that participants will quit the study before all data have been collected. Furthermore, cross-sectional studies tend to have a low level of artificiality and therefore the results may better reflect real-life everyday thoughts, feelings and behaviour (especially when compared with laboratory experiments) (see Box 1.10 on page 51).

A major limitation of cross-sectional studies is that a cause-effect relationship between different variables cannot be tested or determined. For example, in the natural disaster and mental health study, if the results showed that exposure to a natural disaster adversely impacted on mental health, a researcher would not be able to conclude with confidence that exposure to the disaster was the sole cause. Numerous other participant variables can influence mental health and these would not have been subject to the control that occurs in an experiment. Note that this also highlights another limitation of cross-sectional studies — little control over extraneous variables (which may also become confounding variables). However, despite the inability to assess cause–effect, the researcher can still make inferences about possible relationships that may exist at a particular point in time, or collect data that may give direction to further research through true experimental designs.

In addition, when age differences are studied, participant 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 makeup, number of siblings, family environment and schooling may account for differences found in a cross-sectional study of language development in young children.

In particular, one variable that cannot be adequately controlled in some cross-sectional studies is called a cohort effect. A **cohort effect** occurs when the researcher measures characteristics in groups of people ('cohorts') born at significantly different times and members of each group share life experiences associated with the social and cultural experiences of the period and/or place in which they grew up. One or more of these experiences can impact on their development and how they think, feel and/or behave. These perceptions, characteristics and other changes are unique to the group in question.

For example, people who are currently in their nineties experienced childhood during the 1930s depression. They may think, feel and behave differently from 30-, 50- or 60-year olds, not only because they differ in age, but because their backgrounds differ in terms of access to nutritional diets and health care, educational backgrounds and other socio-economic circumstances. Similarly, some age groups experienced adolescence during World War II in the 1940s, without television in the 1950s or when disco was popular in the 1970s. People in these groups may have shared experiences that make them different from other age groups. Consider also the fact that the current adolescent 'cohort' has grown up during a period marked by widespread access to the internet, new digital technologies and social media, which collectively provides a different experience from that of their parents and grandparents when they were growing up.

The larger the difference in age between groups in a cross-sectional study investigating age-related differences, the greater the potential for a cohort effect that causes confounding — when age differences are entangled with differences in participants' life experiences that are actually associated with growing up at a particular time, as a member of a particular generation, in a particular historical context.

Figure 1.55 The larger the difference in age between groups in a cross-sectional study investigating age-related differences, the greater the potential for a cohort effect that may cause confounding.

UNCORRECTED PAGE PROOFS

BOX 1.13 Longitudinal studies and cohort sequential designs

A *longitudinal study* 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, thereby involving a non-experimental (i.e. quasi-experimental) repeated measures design.

A longitudinal study is a relatively useful way of examining consistencies and inconsistencies in mental processes or behaviour over time; for example, discovering if intelligence test scores change with age or remain stable, whether memory declines with age, whether regular physical or mental exercise inhibits onset of a dementia such as Alzheimer's disease, how identical twins reared together or apart may differ on a variety of variables, and how symptoms of a mental health disorder may progress over time.

Because longitudinal studies use the same group(s) of participants, they also allow researchers to study ways in which early experience may affect later experience. However, a longitudinal study can be expensive and take a long time to get results. Keeping in touch with the same group over a period of time can also be difficult — participants may lose interest in the study, move to another location where they are unable to be contacted, or even die.

A research study with a *cohort sequential design* combines the cross-sectional and longitudinal methods to investigate change over time. It has some of the advantages of each method and eliminates some of the disadvantages of both.

The cohort sequential method involves two or more groups of participants (i.e. cohorts) who overlap in age. For example, a study might begin with three groups of adolescents aged 14, 16 and 18. Every two years, this group is tested on risk-taking behaviour, the characteristic of research interest, until the 14-year-olds have turned 18. In addition, every two years a new group of 14-year-olds is added to the research study.

In this way the study is longitudinal in that it spans a four-year period. But the study is also cross-sectional in that it provides data from three different groups that can be compared directly. It also provides a comparison of adolescents who were the same age (14, 16 and 18 years) at three different times. This third comparison enables researchers to identify socio-economic and historical factors that may influence age-related differences. Furthermore, even though the study spans only a four-year period, it provides longitudinal data over an eight-year period.

Although studies with a cohort sequential design are not used often, primarily because of the time and cost involved, they enable a comparison of cross-sectional and longitudinal data for the same groups of participants.

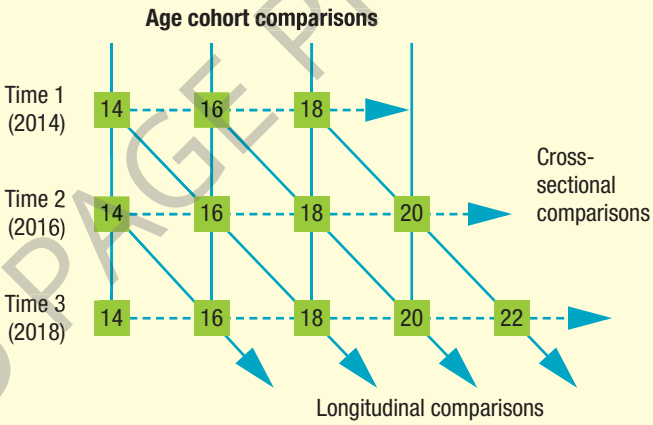


Figure 1.56 Example of a cross-sequential design

LEARNING ACTIVITY 1.25

Review questions

1. (a) Explain what a cross-sectional study is, with reference to an example.
(b) Is a cross-section taken of participants, variables or both?

2. Why may some cross-sectional studies be referred to as quasi-experiments?

3. Which type of experimental design is a cross-sectional study most like?

4. Give two reasons to explain why a cross-sectional study cannot be used to test a cause–effect relationship between the variables of interest.

5. (a) Explain the meaning of cohort effect with reference to an example.
(b) Give an example of when a cohort effect would be considered a confounding variable and explain why it would be considered a confounding variable in relation to your example.

6. (a) Suggest an example of cross-sectional research that could be conducted to investigate whether the duration of a typical night's sleep changes with age.
(b) If age-related differences in total sleep time are found, would your research design enable you to account for or explain the differences? Give a reason for your answer.

7. Suggest an example of cross-sectional research that could be conducted to investigate each of the following.
(a) mental and physical health problems associated with dementia
(b) behaviour management strategies used by student teachers during their training years
(c) problematic mobile phone use in adolescence
(d) quality of life of people who regularly experience panic attacks

8. Which research method is considered better than the cross-sectional study for measuring change over time?

9. What are the main advantages and limitations of the cross-sectional study?

UNCORRECTED PAGE PROOFS

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 researcher may be conducting 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. Usually, the 'case' that is the subject of 'study' is a person. It may involve the normal or abnormal behaviour or functioning of an individual, over a short period of time or a long period. Sometimes it may involve a specific group or event, such as the decision-making process used by an ethics committee when reviewing a research proposal or by a NASA team when deciding to proceed with or abort a space mission launch.

A case study may also involve an organisation; for example an investigation of a specific issue such as how staff respond to or cope with news of their workplace's relocation, permanent closure or some other significant change.

Sometimes a researcher may conduct multiple case studies on a research question and combine the results to look for trends or patterns that permit tentative conclusions. For example, data collected from case studies of different individuals diagnosed with a particular mental disorder may be combined to better understand one or more factors contributing to its onset, maintenance and/or relapse.

Case studies are most often used when large numbers of participants are not available for study; for example, to study individuals with a relatively rare or unusual disorder, problem, ability or characteristic. 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, 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. However, a case study is different from a single participant experiment because the method does not actually involve manipulation of any independent variable.

Clinical psychologists who treat people with mental health problems and disorders routinely conduct case studies involving their clients. However, this research is usually for diagnostic and treatment goals (rather than for scientific research purposes). Therefore, when used in a clinical setting for therapeutic purposes, a case study is often referred to as a *case history* or a *clinical observation*.

TABLE 1.6 Examples of case studies

Person	The study of one single individual, generally using several different research methods
Group	The study of a single distinctive set of people, such as a family or small group of friends
Organisation	The study of a single organisation or company and the way that people act within it
Event	The study of a particular social or cultural event and the interpretations of that event by those participating in it
Location	The study of a particular place and the way that it is used or regarded by people

Much of what is known about the role of the brain in behaviour and mental processes has come from case studies of people with a brain injury. Intensive study of individuals with brain damage makes it possible for researchers to gain detailed, valuable information about the roles of the brain in consciousness, speech, memory, perception and so on.

One of the earliest and best-known case studies of an individual with a brain injury is that of Phineas Gage, which was reported by his doctor John Harlow in 1848. Gage was a railway construction supervisor who accidentally exploded gunpowder that sent an iron rod through his skull, causing massive damage to his frontal lobes. No-one expected him to live. Although Gage survived, his temperament (mood), social behaviour and personality changed very noticeably after the accident. The last sentence of his doctor's report reads, 'His mind was radically changed, so decidedly that his friends and acquaintances said that he was "no longer Gage"'. The Gage case study provided one of the earliest detailed insights into the roles of the frontal lobes in mental processes and behaviour research.

Another well-known case is that of Henry Molaison, whose brain injury was intentionally caused through surgery. His case is examined in detail in your study of human memory (pages 388–90).

UNCORRECTED PAGE PROOFS

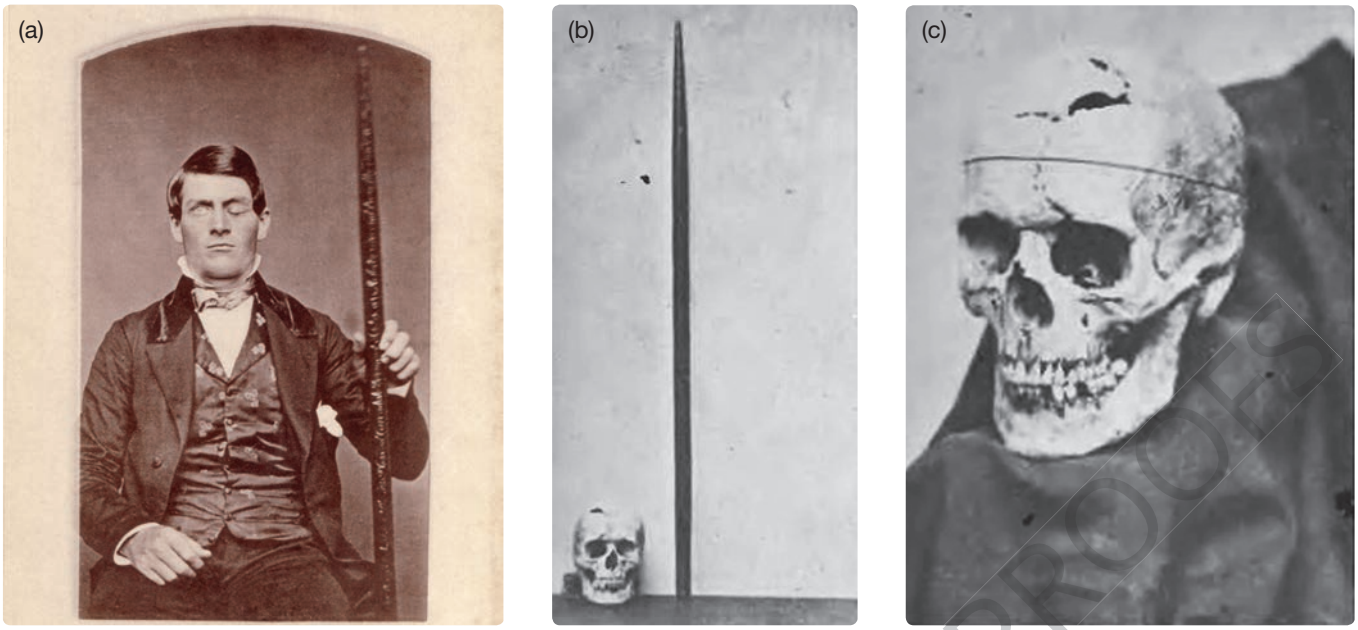


Figure 1.57 Some of the earliest information on the effects of damage to the brain's frontal lobes has come from a case study of the injuries of Phineas Gage in 1848. (a) Gage is shown holding the rod that caused the damage. (b) Gage's skull and the rod. (c) A close-up of Gage's skull.

UNCORRECTED PAGE PROOFS

BOX 1.14 Case study research on Anton's syndrome

Case study research in the late 19th century led to the identification of a rare disorder involving blindness without any awareness of being blind. A common pattern of symptoms emerged and the disorder was subsequently called Anton's syndrome, named after the Austrian neurologist and psychiatrist Gabriel Anton who first reported the disorder in 1893 (Förstl, Owen & David, 1993).

People with *Anton's syndrome* are cortically blind but do not realise that they are blind. Cortical blindness means that they are unable to see because of severe damage to their visual cortex, the part of the brain that initially receives and processes incoming visual sensory information. An unusual aspect of Anton's syndrome is that individuals do not have any damage to their eyes or visual pathways to the brain. However, they believe that they can still see and have an explanation for why they cannot. For example, someone with Anton's syndrome may claim that they can't see because there is insufficient light in the room where they are being examined (Andrewes, 2001).

Each case study published on Anton's syndrome since it was first reported has enhanced understanding of the disorder. For example, a more recent case involved an adult female patient called H. S. to maintain confidentiality.

H. S. was of particular interest to the researchers because her visual cortex was entirely destroyed. Despite all the evidence that H. S. was totally blind, she would deny her blindness and describe her sight as only 'unreliable'. She reported that sometimes things around her would appear very clearly, only to disappear a few minutes later. Sometimes she would reach out for an object, such as a cup, only to find that it was not where she expected it to be (Goldenberg, Mullbacher & Nowak, 1995).

The researchers believed that H. S. might have been mistaking her visual imagery of objects for sight, believing that what she was imagining was what she was actually seeing. They tested their imagery hypothesis by making sounds that related to various objects — for example, the sound of rattling keys or scissors opening and shutting — and then placing the object out of sight. At other times, they let H. S. touch the object and then placed the object out of sight.

Figure 1.58 Gabriel Anton (1858–1933)

Each time they did this, the researchers would ask H. S. whether she saw the object. When not allowed to hear or touch the object, H. S. would say that she couldn't see anything, but she would report seeing the object if the object was within her field of vision.

The following dialogue, in which R. is the researcher and H. S. is the patient, reveals the test. Although by this time H. S. had recovered some of her vision, she still only had a 5° visual window on the right side. Apart from this, she is functionally blind.

R. [Moves bunch of small keys, producing sound.] *I am holding an object. Do you have any idea what it might be?*
H.S. *Could that be a key?*
R. [Silently moves the keys beneath the table. The part of the conversation printed in italics takes place while the keys are hidden from view.] *What does it look like?*
H.S. *On top there is a big ring, and it has a dark key-bit.*
R. *Do you see the key well?*
H.S. *I am seeing the key.*

As can be seen from these notes, the case study of H. S. gives researchers much insight into different aspects of brain function, as well as providing evidence for a number of different aspects of brain function. For example, the visual cortex is shown to have a crucial role in vision, given that H. S.'s was entirely damaged and she consequently had no vision. Despite believing that she could still see objects, as indicated in her conversation above with the researcher, H. S. was blind — her description of the key was incorrect. However, she had excellent visual imagery, despite having no visual cortex. This suggests that visual imagery and visual perception do not necessarily depend on the same brain structures and processes, and that the relationship between visual imagery and visual perception is not as close as some psychologists have proposed. Furthermore, H. S. recovered some of her vision over time (and recovery may continue). This provides evidence for the plasticity of the brain; that is, the capacity of the brain (specifically its neurons) to take over part or all of a function of an area responsible for that function, but which has been damaged (Andrewes, 2001).

Advantages and limitations of case studies

Case studies provide a useful and effective way of obtaining highly detailed information (data and results) on behaviour and mental processes, particularly in relation to rare or unusual disorders or conditions. Their depth of analysis and the richness of the data are commonly described as their main advantage (or 'strength').

With case studies, there is usually no manipulation or control of variables, as with research conducted under strictly controlled experimental conditions (unless an experiment is used to collect some of the case study data). Consequently, case studies can avoid artificiality and provide a 'snapshot' of the actual or real-life experience of one or more individuals at a particular time in a particular situation. Case studies are, however, not only useful for a 'snapshot'. They can be conducted over a prolonged period, even many years where relevant and practical to do so, and may therefore also be useful for tracking and describing experiences and change over time.

Case studies can also provide insights into how others may think, feel or behave under similar circumstances, especially when information from different case studies on the same topic or research question is compiled and knit together to help identify a general pattern or trend in the results. Another advantage of case studies is that they can be a valuable source of hypotheses for further research or for data to support theory building or challenge a theory's assumptions.

A major limitation of case studies is that they cannot test a cause–effect relationship as does an experiment. For example, a case study cannot be considered to be a single participant experiment.

Their small sample size is another limitation. By their very nature, case studies usually focus on rare or unusual individuals, groups or situations. This means that the sample is often a convenience sample (rather than a random sample) and limited to a size of one. The results for such a sample can usually provide only very tentative and limited support for drawing conclusions.

Furthermore, generalising the results to others in a relevant population cannot be done with any certainty. Generalising is a bigger problem when the case study involves a rare or unusual disorder or ability. Because the mental experiences, processes or behaviours of such individuals (or groups) are 'extraordinary', they may not reflect typical ways of thinking, feeling or behaving. Therefore, the researcher can never be fully confident that the conclusions drawn from their study are representative of similar instances within the wider population or apply elsewhere over time. This means that a case study usually has poor external validity. In addition, a case study's use of a rare or unusual sample means that it often cannot be replicated to test the reliability of the results in the way that an experiment can.

Case studies have other limitations. Because of the very detailed and comprehensive data usually obtained, the process of analysing, summarising and reporting these data can be painstaking and time-consuming.

Case studies also have the limitation of being susceptible to biased information from the participants or the researcher. This can influence

UNCORRECTED PAGE PROOFS

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 and are vulnerable to experimenter effects. It is possible that the researcher may see or hear 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 a report on 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.

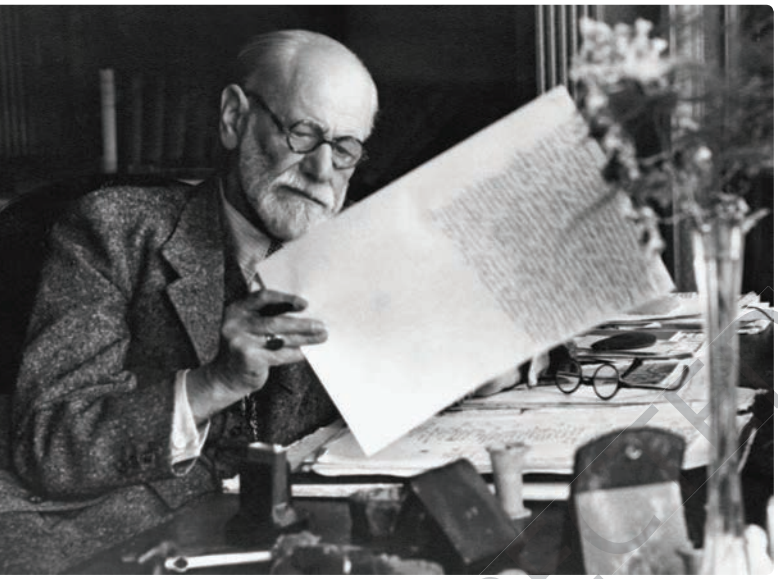


Figure 1.59 Sigmund Freud (1856–1939) developed his psychoanalytic theories mainly from case studies of patients who sought his help with mental health problems they were experiencing.

LEARNING ACTIVITY 1.26

Review questions

1. What is a case study?
2. Give two examples of possible case studies not referred to in the text, one involving an individual and another a small group or organisation.
3. Suggest an example of an experiment that could be conducted as part of a case study.
4. Describe three advantages and three limitations of case studies when used for research purposes.
5. What is a potential limitation of a case study conducted over a prolonged period of time, but other than a limitation referred to in Q4 above.

Observational studies

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

- *structured* — a pre-prepared system is used to guide and record observations e.g. a checklist of items to precisely guide what to look for and to record or exclude
- *unstructured* — observations are made without a predetermined format
- *semi-structured* — a part of the study involves using 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 on the next page. A structured study typically involves operationalising the behaviour of interest and variables that are involved. For example, a researcher observing aggression outside nightclubs in King Street, Melbourne, must define aggression precisely in terms of the variables to be measured and devise a list of the specific behaviours to be observed and recorded. In preparing their observation checklist, the researcher will determine whether, for example, aggression includes shouting or only physical contact and whether an accidental push or shove is to be recorded along with a deliberate push or shove.

One of the difficulties in observing and recording ongoing target behaviour is that there can be so much of it and/or it may be occurring within the context of lots of other behaviour. The researcher must therefore also decide in advance whether to record all occurrences of the behaviour and when observations are actually made during the study. For example, three behaviour sampling decisions for an observational study may involve:

1. *event sampling*: The researcher decides to focus on one or more specific types of behaviour ('events') and record all occurrences. All other types of behaviour are ignored. For example, in a study of bullying in the school grounds, only those acts involving intentional physical contact initiated by the bully. In some studies where relevant, the researcher may also record the antecedent (actions immediately prior to the event) and/or the consequences (what happens immediately after the event).
2. *time sampling*: The researcher decides that observation will take place only during specified time periods (e.g. 1 minute every 5 minutes per hour, or 10 minutes every hour, 1 hour per day) and records the occurrence of the specified behaviour during that period only.

UNCORRECTED PAGE PROOFS

INITIATION OF AGGRESSIVE ACT, SHOWN BY CHILD AND SEX

	#	1	M	#	2	M	#	3	M	#	4	M	#	5	M	#	6	F	#	7	F	#	8	F
Aggressive act against whom (shown by child's assigned number)*	5	6	8				6	7	2	1	3	5	1	6	8	2	2	4				6	3	5
Response:																								
Anger												X		X		X								
Surprise													X			X								X
Hurt, tears		X	X				X	X							X		X					X	X	X
Little or no emotion shown	X								X	X	X													
Help from other children			X				X	X					X									X		
Help from adult supervisor			X					X							X	X								X
Resolution favours:	1	1	8				8	7	3	?	?	5	5	5	8	6	2					8	3	5

*The observer assigned numbers to the children, keyed to clothing. The child in the red sweater was #1, the one in the blue shirt was #2, and so on.

Figure 1.60 An example of a checklist used for an observational study

Source: Wiseman, J.P., & Aron, M.S. (1970). *Field projects for sociology students*. Cambridge, Massachusetts: Schenkman. p.21.

3. *individual sampling*: Rather than trying to record the behaviour of all individuals at once, the researcher decides to focus on observing one individual (or group) at a time while ignoring the behaviour of others during the time period. One individual may be randomly selected to be the focus of an observational period and all others are ignored during that period. Over the entire period of the study, however, each individual may be observed. And, when dealing with a group's behaviour, the researcher might observe the entire group all at once, or only one group member for a certain time, then observe another, and so on.

Observations have become more accurate as new technologies permit increasingly precise measurements. For example, digital video cameras can be used to record then analyse rapidly changing behaviour. Even a single 'frame' within a long action sequence can be analysed. This technology can be used, for example, in studies of the way subtle changes in facial expressions of mothers and their babies become synchronised and similar over time. The miniaturisation and greater portability of recording devices has also enabled observational studies to become less obtrusive where desirable.

Sometimes an observational study might resemble an experiment. For example, to investigate creative problem-solving processes in a group, a researcher might present a friendship group with a problem requiring its members to come up with as many uses for a house brick as they possibly can in 10 minutes. The researcher could then observe and record who suggests answers and how often; who records answers and how the recorder is selected; how correct and incorrect answers are dealt with; whether judgments about answers are immediate or postponed; who gives or doesn't give feedback; who stays on task; who is time-conscious, and so on. The researcher might also observe problem solving in a group comprising strangers in order 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 using groups 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 creative problem-solving), but only a true experiment can establish a cause–effect relationship because there is an IV that is manipulated along with the use of random allocation. In an observational study, there may be naturally occurring variables of interest but the researcher does not interfere in the situation — there is no intervention that might disrupt what is occurring naturally.

UNCORRECTED PAGE PROOFS

Natural and contrived settings

Observations may be conducted within a participant's natural environment or in a contrived environment. In both settings, the 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 social behaviour of pre-schoolers, a researcher might observe children at play in a pre-school 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. From the observational records of each child's interactions with others, the researcher will make inferences about children's social behaviour. Similarly, a researcher studying how paramedics respond to traumatic events might observe paramedics in action by riding along with them on duty. In doing so, the researcher would be as unobtrusive as possible, trying to 'shadow' the paramedics as they respond to various types of trauma, communicating with them only when essential.

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 *controlled* or *laboratory* environment because of the degree of control the researcher has over it.

Generally, with this type of observational study, the researcher decides where the observations will take place, at what time, with which participants, in what circumstances, and uses standardised procedures. There may also be observations of different groups for which an IV may be present or absent.

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 away from the pre-school centre. Specific playthings may be made available and strategically located together with a table and chairs. Observations could then be made at pre-determined times from behind a one-way mirror so that multiple the children are not aware that multiple they are being observed. The children's behaviour might also be video recorded so that multiple researchers can also record observations to help ensure reliability of the data. However, despite the use of such control over the environment and observational procedures, the researchers must still wait for the behaviours of interest to occur naturally.

REJECTED PAGE PROOFS

Figure 1.61 (a) An observational study may be conducted in a natural, 'real-life' environment or (b) a formal contrived environment such as a formal laboratory setting.

Participant and non-participant observation

When researchers try to conceal their presence while making observations, it is often called *non-participant observation* (and sometimes *passive participation*). For example, when making naturalistic observations of the use of specific 'body language' in a real-life setting, a researcher might sit on a nearby seat pretending to be absorbed in a book in order to observe the non-verbal interactions of people being met at an airport. They will try to blend in with the crowd and observe from the sidelines, concealing as best as possible their identity and what they are actually doing so that the individual(s) being observed does not alter their actions, as may occur when people know that they are being observed.

Sometimes, the behaviour of interest involves private interactions between members of a group that cannot be reliably observed from the sidelines. In such cases, the researcher may engage in participant observation. In an observational study using *participant observation* the researcher is an active member of the group being observed (which is why it is sometimes called *active participation*). The researcher will actually join in the activities of the group and may deliberately try to be mistaken by the participants as being part of the group or situation being observed, but their identity as an observer is concealed.

Figure 1.62 We often watch people engage in normal behaviour in public settings. If you stop and watch for a while without intruding on whatever they are doing then you are engaging in non-participant observation.

In one well-known observational study that used participant observation, the researchers set up a situation where 8 pseudo-patients had themselves admitted to several different psychiatric hospitals by imitating symptoms of schizophrenia. After they had been admitted, the fake patients took part in ward activities and spent time observing and writing notes about ward staff and how patients were treated. Their record-keeping behaviour was regarded by the hospital staff as being a symptom of their mental disorder. In all, they remained in hospital for 7 to 52 days (average 19 days) and were eventually discharged with a diagnosis of 'schizophrenia in remission' (Rosenhan, 1973).

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 or unfamiliar 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. Controlled observations in

PREPRINT PAGE PROOFS

laboratory settings can also be more easily replicated by other researchers using the same procedures. This means that the reliability of results can be tested.

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 most aspects of true human crowd behaviour in an artificial laboratory setting. 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 doesn't 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. But this is often worthwhile as observational studies may enable generation of new hypotheses.

A practical advantage of naturalistic observation is that it does not require the cooperation 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 up 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.

A limitation of many observational studies, particularly when unstructured, 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 is not manipulating or controlling any variables and 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 the researcher is not immediately aware of. Consequently, an observational study may reveal a relationship between variables, but not a cause–effect relationship as does an experiment.

In addition, naturalistic observation studies such as the King Street example, often lack a representative sample. For instance, they may be biased in relation to participant variables such as age, sex, cultural and socio-economic background. This means that the results may not readily be generalised to a wider population (which impacts on external validity). Furthermore, because of the lack of control of variables in such a setting, the study is also more likely be low in reliability as it will be difficult for a researcher to replicate it.

A potential limitation of any observation procedure is *observer bias,* which is a type of *experimenter bias* (or experimenter effect). 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.

Figure 1.63 Using two or more observers to collect the same data in an observational study can help minimise the influence of observer bias on the results.

UNCORRECTED PAGE PROOFS

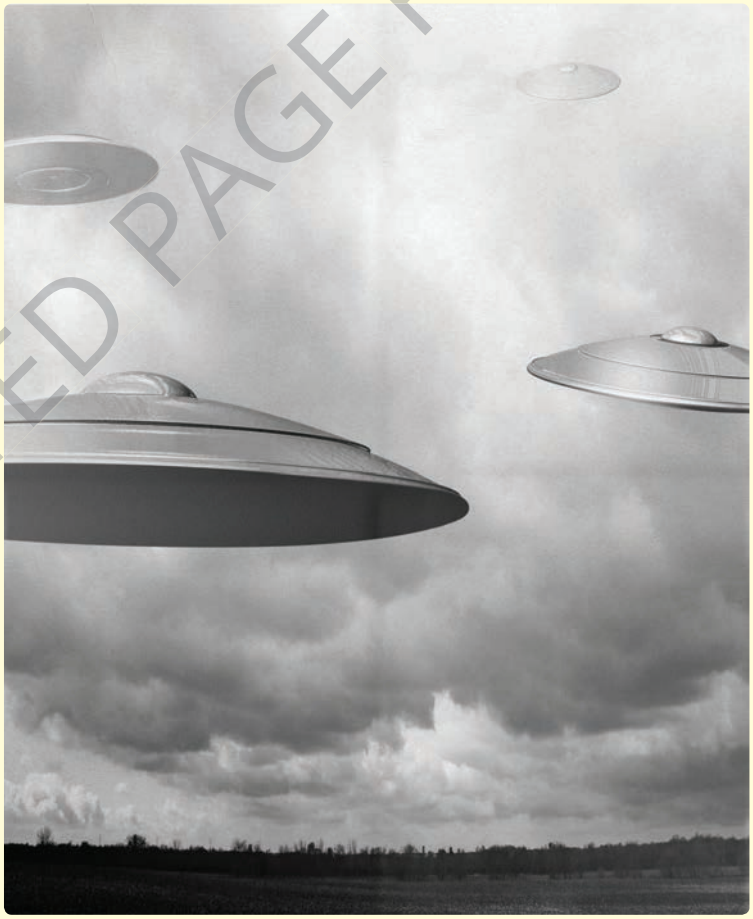
BOX 1.15 Studying a doomsday cult using participant observation

American psychologist Leon Festinger (1956) used participant observation in an observational study of a small UFO cult called the Seekers. The cult believed that a great flood would destroy the world on December 21, 1954 and that aliens would arrive in a flying saucer just before the hour of the apocalypse and take the true believers to safety on another planet. Members of the cult left their jobs and families and gave away money and possessions to prepare for the arrival of the superior beings from outer space.

The research involved participant observation because Festinger and three other researchers became members of the group that they were studying and interacted with them in various cult meetings over many weeks before and after the predicted apocalypse.

Festinger wanted to find out how the cult members would cope and react when they realised that the apocalypse did not occur. He observed that when neither the flying saucer nor the apocalypse arrived at the appointed time, the committed cult members did not abandon their doomsday beliefs despite the fact that they had been clearly discredited. Instead, after several uncomfortable hours on the appointed day, they accepted their leader's explanation that 'the little group, sitting all night long, had spread so much light that God had saved the world from destruction' and became even more fanatical about their beliefs.

Festinger proposed that the cult members' beliefs had strengthened because we all have an inner drive to reduce the psychological discomfort we can experience when we hold two conflicting beliefs simultaneously, such as when there is a mismatch between a firm belief (e.g. there will be an apocalypse at X time) and what may be true in reality (e.g. the predicted apocalypse did not occur).



UNCORRECTED PAGE PROOFS

Figure 1.64 Researcher Leon Festinger (1919–1989), shown above, used participant observation to study how cult members would react when the flying saucer they were expecting to carry them to safety did not arrive.

eBook*plus*

Weblink
Explanation of cult beliefs – Festinger's cognitive dissonance theory 3m 30s

BOX 1.16 Studying gorilla behaviour using both participant and non-participant observation

In a well-known observational study that spanned 18 years, American researcher Dian Fossey (1983) used both participant and non-participant observation. Fossey, whose work is featured in the 1988 movie *Gorillas in the mist*, lived among gorillas in their remote African highlands habitat.

After first using non-participant observation to learn about key aspects of gorilla behaviour, Fossey changed her method to use participant observation and started to interact with and behave like a gorilla. The more she learnt about the behaviour of gorillas, the more she was able to act like them. She imitated their feeding and grooming behaviours and even attempted to copy their vocalisations. By waiting for the gorillas to approach her, by avoiding actions that might threaten them, and by imitating their actions, Fossey gradually became accepted by them and was able to collect valuable data about their behaviour.

Fossey was not formally trained in scientific research but her contributions to the understanding of gorilla behaviour (and gorilla conservation) are widely recognised. In 1985, she was found hacked to death at her camp site in Rwanda. It is believed that she was murdered by gorilla poachers.

Figure 1.65 Researcher Dian Fossey (1932–1985) engaging in participant observation with gorillas

eBookplus

Weblink
Video on Fossey's observational studies 3m 03s

LEARNING ACTIVITY 1.27

eBookplus
Word copy of table

Review questions

1. What is an observational study?
2. (a) Suggest an example of an observational study with a non-experimental independent groups design (but not an example used in the text).
   (b) Explain why this study would be considered a quasi-experiment.
3. What are the key features of naturalistic observation?
4. Distinguish between each of the following with reference to examples not used in the text.
   (a) structured and unstructured observation
   (b) participant and non-participant observation
   (c) event and time sampling
5. (a) Explain the meaning of observer bias and whether or not it is a type of experiment effect.
   (b) What is a suitable means of controlling this variable?
6. Are naturalistic observations of people without obtaining their informed consent ethically acceptable? Explain your answer.
7. Describe three advantages and three limitations of observational studies.
8. To what extent do reality TV programs such as *Bachelorette, Real Housewives* and *Survivor* produce authentic human behaviour? Explain with reference to an example of a program and design features of observational studies that help ensure 'normal' behaviour is more likely to be recorded.

UNCORRECTED PAGE PROOFS

10. Complete the following table to analyse different observational studies.

Observational study	Structured vs Unstructured	Naturalistic vs Contrived setting	Participant vs Non-participant
(a) A teacher concerned about the unsafe behaviour of students at the school's bus stop at the end of the school day organises an observational study. Observations will be made from a nearby classroom with reference to a checklist.			
(b) Trainee counsellors will be assessed while they conduct consultations with each other, taking turns to be the counsellor then the client. All assessments will be conducted in a room at the university set up for that purpose. The course leader will video record each session and a criteria sheet will be used to guide feedback to trainees.			
(c) The captain of the school's senior hockey team will analyse the players' communication styles during an upcoming match.			
(d) A researcher will record the number of drivers who obey a give way sign at a roundabout.			
(e) Researchers will compare the behaviour of AFL football spectators who sit behind the goals with those who sit in a grandstand. Observations will target the number of comments directed at umpires and players (but not the content).			
(f) A VCE student is planning to conduct an observational study in a shopping mall to find out whether people look at their own reflections or avoid doing so when walking past a large department store. Sex differences will be recorded using a 3-point rating scale based on 'Yes', 'No' and 'Not sure'.			
(g) A VCE student is planning to conduct an observational study at school to find out whether junior school students are more disruptive with a replacement teacher. There will be two observation sessions conducted in the same period, in the same classroom, on the same day across two weeks. In week 1, the student will observe a Year 7 Maths class from a storeroom when taught by their usual teacher. Only six of the students will be observed — two seated mid-front row, two mid-middle row and two mid-back row. In week 2, the student will observe students in the same class, seated in the same positions, but when taught by a replacement teacher who has never previously been at the school. Disruptive behaviour has been operationalised to enable development of observation criteria.			

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.28

Reflection

The use of deception in an observational study allows the researcher to study a person's natural behaviour. Comment on whether deception contradicts the ethical principle of informed consent.

LEARNING ACTIVITY 1.29

Evaluation of research

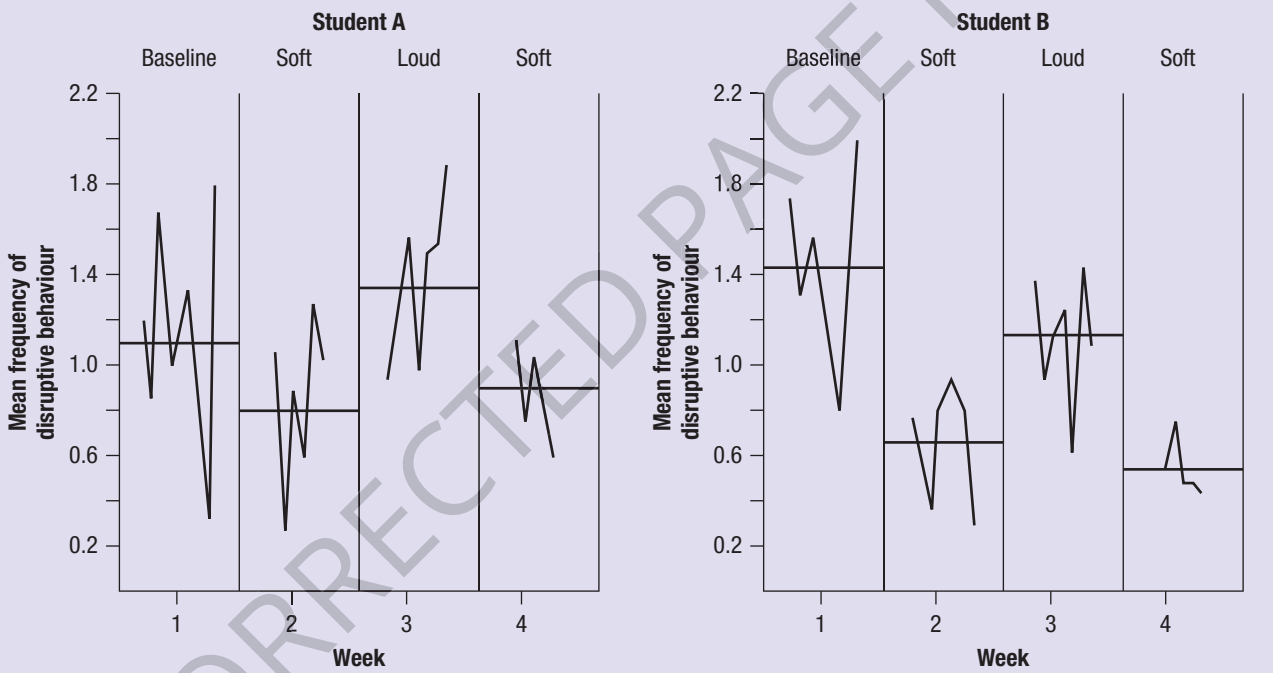
A researcher conducted an observational study to test the effectiveness of different types of reprimands when dealing with disruptive students. On the basis of observations from an earlier study, it was expected that giving soft reprimands to disruptive students would have a greater effect than loud reprimands.

Two disruptive students in a Year 3 class at a local primary school were observed for a 20-minute period during an arithmetic lesson each day for 4 weeks by two trained observers who sat at the back of the classroom and were as unobtrusive as possible. The observers followed a schedule of observing for 20 seconds then recording for 10 seconds using a checklist with eight types of disruptive behaviour, including whether the student was out of their chair, noisy, communicating with another student, interfering with another student's property, or being aggressive.

There were four phases of the study:

1. Baseline: Measurements of disruptive behaviour before any treatment intervention is introduced. The teacher used their normal loud or soft reprimands, however, it was found that the teacher predominantly used loud reprimands throughout this period.
2. Soft: The teacher used soft, 'private' reprimands so that only the student being reprimanded could hear; soft reprimands were used with all students, not just the target students being observed.
3. Loud: The teacher used loud, 'public' reprimands that could be heard by all students; loud reprimands were used with all students, not just the target students.
4. Soft: The teacher used once again used soft, 'private' reprimands.

The results are shown below.



1. Formulate a research hypothesis that would be supported by the results obtained for the study.
2. Identify the sampling procedure.
3. Give an advantage and a limitation of the sampling technique used for this particular study.
4. Explain whether the event sampling or time sampling procedure was used for this particular study.
5. What was the purpose of the baseline measurements?
6. Is the study best described as structured, unstructured or semi-structured?
7. Does the study use participant observation, non-participant observation, or both?
8. Explain whether the study can be described as a naturalistic observation study.
9. Identify the operationalised independent and dependent variables in the study.
10. Suggest a single, relevant title for the series of graphs showing the results.
11. What do the results show?
12. What conclusion can be drawn from the results?
13. To what extent can the results be generalised?

UNCORRECTED PAGE PROOFS

Self-reports

For some topics of research interest, it is best to ask people about their thoughts, feelings or behaviour. For example, a researcher may be interested in studying what people dream about, how often they have nightmares, thoughts and feelings that accompany a stressful experience, why people jump queues, why people with a spider phobia react as they do when they see a spider, or what people do to cope with the fear or anxiety triggered by a phobic event. To ask people how they think, feel or behave when conducting scientific research involves using a measure or technique that will prompt 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 take the form of answers about study habits before an important exam, to statements in a seven-point rating scale measuring anxiety, or a participant's diary records kept in response to a researcher's specific request (such as the sleep diary records described on page 452).

Assuming that 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 measures. 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 methods, 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.

Figure 1.66 Researchers conducting surveys often use a questionnaire to collect self-report data. Online administration is usually supported by efficient data processing and analysis.

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. How questions are asked and the categories of response are focused but not necessarily predetermined or fixed.

Interviews are most often conducted with individuals, in a face-to-face meeting or sometimes over the phone, or even through an app like Facetime or Skype. They usually require spoken answers to questions and are rarely used with very large samples as data collection would require a considerable amount of time. Answers may be recorded electronically or in hard copy writing, depending on how the interview is conducted. Unlike questionnaires, which are usually structured, interviews may be structured, unstructured or semi-structured (National Health and Medical Research Council [NHMRC], 2007).

In a *structured interview*, the participant (or 'interviewee') is asked specific, predetermined questions in a controlled manner. The most structured interview is when the interviewer simply asks a set of fixed-response questions and records the participant's answers; for example, questions such as 'Have you ever been awoken by a nightmare?' with predetermined answer options such as 'Yes', 'No', and 'Not sure'. The interviewer follows a script and the questions are read out from a list in a neutral manner with no comments or cues. This is done to ensure

UNCORRECTED PAGE PROOFS

that all participants are treated in the same way and thereby help maintain standardised procedures. A less structured interview may use free-response questions (such as 'How do you feel immediately after waking from a nightmare?'), but the researcher will follow a script and ask a set list of questions to ensure consistency across all participants.

In an *unstructured interview*, the researcher has an overall aim of what data should be collected, but the interview is driven by the participant and there is a spontaneous generation of questions in the natural flow of interaction with the 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. This means that questions asked in an unstructured interview can vary widely from participant to participant (NHMRC, 2007).

A goal of unstructured interviews is to allow people to describe their thoughts, feelings and behaviour in their own way using their own words and to give more or less emphasis to relevant issues. 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.

In a *semi-structured interview*, the researcher uses an interview guide listing a set of issues to be explored. The researcher aims to cover all issues but there are no set questions to be asked. As with the unstructured interview, there is a spontaneous generation of questions through interaction with the participant (NHMRC, 2007).

Figure 1.67 An interview is often a face-to-face discussion between a researcher and an individual for the purpose of obtaining detailed information.

UNCORRECTED PAGE PROOFS

BOX 1.17 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 participants to identify or describe their thoughts, feelings or behaviour 'freely' in their own words; for example:

'What are you thinking?'
'How do you feel when stressed?'
'How do you usually react when this happens?'

These kinds of questions enable participants to provide detailed responses without being restricted to giving answers that fit into predetermined categories (such as those of fixed-response questions). Furthermore, in an unstructured or semi-structured interview, 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.

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. Scoring or interpreting the responses also requires subjective interpretation by the researcher. Consequently, the scoring and interpreting of responses are susceptible to the biases and expectations of the researcher; that is, experimenter effects.

To avoid or overcome these limitations, researchers often ask fixed-response questions, which are more objective. *Fixed-response* (or *closed*) *questions* typically require participants to identify or describe their response by selecting from a number of 'fixed' alternative responses; for example:

'Do you dream in colour?' Yes, No, Not sure
'How often you remember your dreams on awakening?' Always, Often, Sometimes, Not often, Never
'How much time does it usually take to fall asleep when you go to bed at night?' 0–10 minutes, 11–20 minutes, 21–30 minutes, 31–60 minutes, 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 and describe the responses numerically (thereby obtaining quantitative data).

For example, a '0–10 minutes' response to the question about time taken to fall asleep can be assigned a score of 1, '11–20 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 1.68 Free- or fixed-response questions can be used in a questionnaire to collect self-report data from a large number of people in a relatively short period of time.

UNCORRECTED PAGE PROOFS

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 collect 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 fear or anxiety when making an oral presentation to a large group, how often they use social media just before going to sleep, their level of tiredness after completing homework on a weeknight, their level of confidence before sitting a test for which they have done the right amount of study, or the strength of their attitudes to racist comments by competitors in sports matches. The series of 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 that enable answers to be quantified (converted to numbers) for summary, analysis and interpretation. The rating scale is not unlike a multiple choice test, but the answer options represent levels or degrees of a particular characteristic. However, there is no correct answer for a rating scale item, other than what the participant decides to give. Box 1.18 on the next page shows examples of items in a rating scale for measuring anxious attachment in a romantic relationship.

The best known and most commonly used rating scale is the *Likert scale* (developed by American psychologist Rensis Likert in 1932). This consists of about 20 questions or statements to which the participant responds, typically using a five-point scale. It is most commonly used to measure attitudes. For example, in a study on attitudes to refugees and asylum seekers who reach Australia by boat rather than conventional means, a Likert scale statement could be 'Refugees and asylum seekers arriving by boat on Australian shores should be imprisoned until background checks can be completed'. 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. Alternatively, options may vary in terms of approval such as strongly approve, approve, undecided, disapprove and strongly disapprove. There are also other possible answer options, depending on what is measured (see Box 1.20 on the next page).

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 1.19.

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. In a true Likert scale, however, positive and negative statements are distributed in a random order. Box 1.18 below has an example of a 7-point scale with statements in a random order. For this scale, the higher the score the greater the level of anxiety experienced in a romantic relationship.

When a respondent has completed a Likert scale, all of the responses are scored and a total is calculated. The result is a score on a attitude. Generally, the higher the score, the more positive or favourable the attitude. Box 1.20 below describes how to construct a Likert scale.

BOX 1.18 Rating scale to measure anxious attachment in a romantic relationship

Please rate your agreement with each of these statements using a 1 to 7 scale, with 1 meaning 'disagree strongly' and 7 meaning 'agree strongly'.

1. I'm afraid that I will lose my partner's love.
2. I often worry that my partner will not want to stay with me.
3. I often worry that my partner doesn't really love me.
4. I worry that romantic partners won't care about me as much as I care about them.
5. I often wish that my partner's feelings for me were as strong as my feelings for them.
6. I worry a lot about my relationships.
7. When my partner is out of sight, I worry that they might become interested in someone else.
8. When I show my feelings for romantic partners, I'm afraid they will not feel the same about me.
9. My romantic partner makes me doubt myself.
10. I find that my partners don't want to get as close as I would like.

Source: Fraley, R.C., Waller, N.G., & Brennan, K.A. (2000). An item–response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78,* 350–365.

BOX 1.19 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

SA = Strongly agree A = Agree N = Neither agree nor disagree D = Disagree SD = Strongly disagree

BOX 1.20 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 or other topic of interest. Although your scale is likely to be a useful measure for your research questions, 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. The steps are written with reference to attitude measurement and a scale varying in strength of agreement or disagreement. However, the steps can be adapted to construct a Likert scale for any topic.

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 1.19 above 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.

UNCORRECTED PAGE PROOFS

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.

- 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 (i.e. 'neither agree nor disagree').
- Use simple, clear language that is suited to the experience, age, and educational and cultural background of the participants whose attitudes you are measuring.
- Write your items in such a way that 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' within items.

Note that Likert scale items may also have other answer options, depending on what is measured. For example:

Level of concern
- Not at all concerned
- Slightly concerned
- Somewhat concerned
- Moderately concerned
- Extremely concerned

Frequency
- Always
- Often
- Sometimes
- Rarely
- Never

Belief
- Almost always true
- Usually true
- Occasionally true
- Usualy not true
- Almost never true

Quality
- Excellent
- Very good
- Good
- Fair
- Poor

Level of awareness
- Extremely aware
- Moderately aware
- Somewhat aware
- Slightly aware
- Not at all aware

Affect
- Major affect
- Moderate affect
- Neutral
- Minor affect
- No affect

Knowledge of action
1. Never true
2. Rarely true
3. Sometimes but infrequently true
4. Neutral
5. Sometimes true
6. Usually true
7. Always true

Level of acceptability
1. Totally unacceptable
2. Unacceptable
3. Slightly unacceptable
4. Neutral
5. Slightly acceptable
6. Acceptable
7. Perfectly acceptable

Level of appropriateness
1. Absolutely inappropriate
2. Inappropriate
3. Slightly inappropriate
4. Neutral
5. Slightly appropriate
6. Appropriate
7. Absolutely appropriate

Level of importance
1. Extremely important
2. Very important
3. Moderately important
4. Neutral
5. Slightly important
6. Low importance
7. Not at all important

Level of difficulty
1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

Level of influence
1. Not at all influential
2. Slightly influential
3. Somewhat influential
4. Moderately influential
5. Extremely influential

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 one of the following:
  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.

Adapted from Grivas, J., & Lawrie, P. A. (1991). *Psychology: Experiments and activities*. Marrickville, NSW: Harcourt Brace Jovanovich, pp. 401–403.

UNCORRECTED PAGE PROOFS

Advantages and limitations of self-reports

Self-report measures 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. They also have the advantages of making it relatively easy to compare responses among participants and to replicate a study, especially when structured measures are used.

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 reliably recall 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 refugees, Aboriginal land rights, the death penalty 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 an 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 some 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-report techniques, 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.

Figure 1.69 Self-report methods of data collection provide useful information about how people think, feel and behave. However, they typically rely on participants having well-developed language skills and being able to accurately recall and state the information required of them.

UNCORRECTED PAGE PROOFS

BOX 1.21 Meta-analysis

A *meta-analysis* is an advanced statistical technique that combines and summarises the results of multiple research studies on a given topic (or research question). The results are integrated to identify patterns and draw conclusions about those studies. The technique takes account of the fact that the studies have used different samples, different research methods, different measures and have obtained different results of varying significance.

Following are examples of studies that have used meta-analysis. The Smith and Glass (1977) meta-analysis on the effectiveness of psychotherapy for mental health disorders summarised 375 original studies published in a diverse range of journals. There were an estimated 50 000 participants divided across experimental and control conditions. The average age of participants was 22 years and they each received an average of 17 hours of psychotherapy. Rosenthal and Rubin (1978) summarised 345 studies that had verified the presence or influence of the expectancy effect described on page 32.

Meta-analysis	Findings
Smith, M.L., & Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.	Psychotherapy is very effective (e.g. the typical psychotherapy client is better off than 75% of untreated individuals) and few important differences in effectiveness between behavioural therapies (e.g. systematic desensitisation) and non-behavioural therapies (e.g. psychodynamic)
Rosenthal, R., & Rubin, D. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences, 1*(3), 377–386.	The expectancy effect is present and influences participant behaviour in many research contexts.
Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta- analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin, 119*, 111–137.	Investigated whether the level of conformity in Asch's classic study has changed over time and whether it is related cross-culturally to individualism–collectivism. Found that collectivist countries tend to show higher levels of conformity than individualist countries.
Brierley, B., Shaw, P., & David, A. S. (2002). The human amygdala: a systematic review and meta-analysis of volumetric magnetic resonance imaging. *Brain Research Reviews, 39*(1), 84–105.	Measured the normal size of the human amygdala and how it changes with age. Found no significant difference in left–right size, the amygdala is significantly larger in men, and an inverse relationship (i.e. negative correlation) between size and age.
Segerstrom, S.C., & Miller, G.E. (2003). Psychological stress and the immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin, 130*(4), 601–630.	Major findings are three-fold. (1) Stress alters immunity, (2) Short-term stress actually 'revs up' the immune system, an adaptive response preparing for injury or infection, but long-term or chronic stress causes too much wear and tear, and the system breaks down, (3) The immune systems of people who are older or already sick are more prone to stress-related change.
Else-Quest, N.M., Hyde, J.S., & Lin, M.C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin 136*(1), 103–127.	Girls around the world are not worse at maths than boys. Boys are more confident in their maths abilities, and girls from countries where gender equity is more prevalent are more likely to perform better on maths assessment tasks.
Tannenbaum, M., et al., (2015). Appealing to fear: A meta-analysis of fear appeal effectiveness and theories. *Psychological Bulletin, 141*(6), 1178–1204.	Fear-based appeals appear to be effective at influencing attitudes and behaviours, especially among women and when they contain recommendations for one-time only (versus repeated) behaviours.
Firth, J., Cotter, J., Elliott, R., French, P., & Yung, A. (2015). A systematic review and meta-analysis of exercise interventions in schizophrenia patients. *Psychological Medicine, 45*(7), 1343–1361.	Interventions that implement a sufficient dose of exercise (e.g. 90 min of moderate-to-vigorous exercise per week) in supervised or group settings, can be feasible and effective interventions for improving both physical and mental health outcomes in people with schizophrenia.

(continued)

UNCORRECTED PAGE PROOFS

(continued from previous page)

Meta-analysis	Findings
Borges, G., Bagge, C., Cherpitel, C., Conner, K., Orozco, R., & Rossow, I. (2017). A meta-analysis of acute use of alcohol and the risk of suicide attempt. *Psychological Medicine,47*(5), 949–957.	Acute use of alcohol is associated with increased likelihood of a suicide attempt, particularly at high doses.
Maijer, K., Begemann, M., Palmen, S., Leucht, S., & Sommer, I. (2018). Auditory hallucinations across the lifespan: A systematic review and meta-analysis. *Psychological Medicine*, *48*(6), 879–888.	Auditory hallucinations are quite common (up to 1 in 10 individuals) in the general population, with children and adolescents reporting these experiences significantly more often compared with adults and elderly.

Figure 1.70 A meta-analysis combines and summarises the results of different research studies on a given topic.

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.30

eBook*plus*
Word copy of table

Review questions

1. What are self-reports?
2. Complete the following table to summarise three measures or techniques for collecting self-report data.

Self-report	Description	Advantages	Limitations
Questionnaires			
Interviews • structured • semi-structured • unstructured			
Rating scales			

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. *Empirical evidence* is data collected through formal observations and/or carefully controlled experiments.

Data can take different forms. The type of data collected is determined by the specific kind of research method or procedure used. For example, questionnaires and interviews often provide data in the form of words, whereas data collected in experiments is usually provided 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'. For example, you will collect primary data when you undertake an experiment to test a research hypothesis for the practical investigation which is one of the Unit 4 SACs. When using a self-report measure, the primary data will be the participants' responses. Their original responses may also be called *raw data* because they 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 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, if you access data in a journal, book or at a website to complete a learning activity or SAC, 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 in 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.

studyon

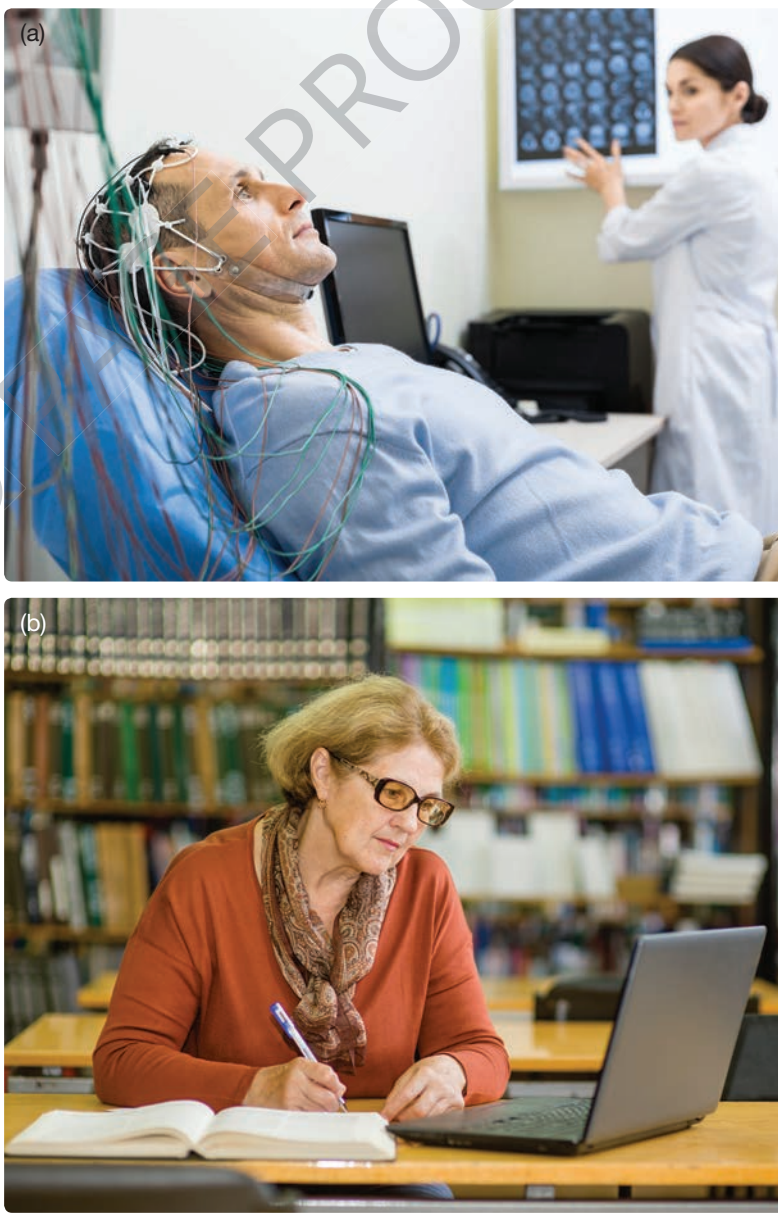


Figure 1.71 (a) Primary data is data collected directly by the researcher (or through others) for their own purpose. (b) Secondary data is data that has been collected by someone other than the original user for their own purpose.

UNCORRECTED PAGE PROOFS

Qualitative and quantitative data

Primary and secondary data may be qualitative, quantitative or a combination of both.

Qualitative data is information about the 'qualities' or characteristics of what is being studied. They are descriptions, words, meanings, pictures and so on. It can describe any aspect of a person's mental experiences or behaviour; more specifically, what type, what category, what something is like or how something is experienced.

Qualitative data could be in the form of written or verbal comments by participants, audio or video recordings, or notes of participants' comments made by the researcher. Anything a person thinks, feels or does can be a source of qualitative data. For example, a researcher may collect and analyse drawings in order to study what the onset of menstruation is like for adolescent girls (Banyard & Grayson, 2000).

Alternatively, a researcher may conduct research on the advantages and limitations of government-funded mental health services provided over the internet and telephone. The researcher may collect data by conducting interviews or holding small-group discussions with individuals who have recently used one or more of these services. Participants may be asked to give examples of when they have used a service and describe their experience without any constraint, other than occasional questions by the researcher to ensure their responses are relevant, have enough detail and have been clearly understood.

Quantitative data is numerical information on the 'quantity' or amount of what is being studied; that is, how much of something there is. This type of data is usually expressed in the form of units of measurement or numbers, such as raw scores, percentages, means, standard deviations and so on. For example, the height or age of a participant is considered quantitative data as both of these characteristics can be expressed in units of measurement (centimetres or years). Similarly, the percentages of participants who respond with 'Agree' or 'Disagree' to interview questions, or the mean time taken to solve a problem in an experiment, are quantitative data.

All types of mental experiences and behaviours can be described in quantitative terms, as quantities or numbers. For example, a survey question might ask participants to use a five-point scale to rate the level of stress caused by different events or the effectiveness of different stress-management strategies.

Psychologists use many different tests to measure various mental processes and behaviours and most of these also provide quantitative data. There are tests to measure intelligence, personality traits and all kinds of aptitudes, interests and abilities. Answers are often totalled to yield a score that can be interpreted and applied to the person or group who did the test.

Similarly, data collected by devices used to record the electrical activity of the brain when awake or asleep are measurements and numerical values that are best described as quantitative data.

Experiments can produce qualitative data as well as numbers. For example, in his classic experiments on obedience to authority in the 1960s, Milgram described the behaviour of his participants in some detail (qualitative data), as well as measuring the extent to which they were prepared to obey the experimenter (quantitative data).

Although different, quantitative and qualitative data are not mutually exclusive and are not often used separately. Qualitative data are typically expressed in the form of words, but they can be converted into a quantitative form. For example, participants' responses to open-ended interview questions about

Figure 1.72 This child's drawings may be qualitative data on how she is feeling following a recent family trauma.

eGuide*plus*

Weblink
ABS — What are quantitative and qualitative data?

UNCORRECTED PAGE PROOFS

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.

The majority of studies referred to in this text use quantitative data rather than qualitative data. This reflects the preference for quantitative data in most psychological research. Generally, psychologists tend to prefer quantitative data because using numbers increases the precision of results and the ease with which the results can be communicated. Quantitative data also enables more precise and detailed analysis through the use of statistical procedures and tests. These are also the reasons why qualitative data are often converted into quantitative data.

Figure 1.73 This participant is required to indicate food preferences shown on the computer screen. Are the data collected qualitative or quantitative?

BOX 1.22 Objective and subjective data

The terms 'objective' and 'subjective' are also used to refer to the way that data are collected and the way they are described and explained.

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).

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 tool 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).

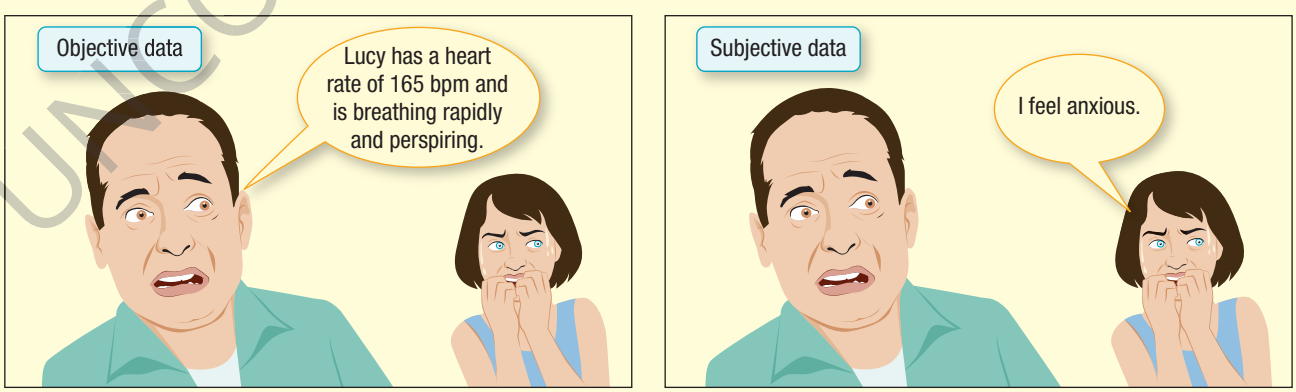


Figure 1.74 Objective and subjective data

UNCORRECTED PAGE PROOF

BOX 1.23 Scales of measurement

The process of assigning numbers to objects and events that have been studied in research is called *measurement*. Essentially, measurement is no different from the process of assigning words to objects and events when using language. Furthermore, in the same way that language has rules that are followed in its correct use, so too does measurement have certain rules. For example, a measurement rule may use the symbol + (plus) for combining numbers. This means adding one thing to another when the symbol for the addition rule is correctly applied.

Researchers use different scales or 'types' of measurement, when collecting data. In turn, the particular scale used determines the way in which the data can be described and interpreted. For example, a researcher investigating which leadership style is most effective might record how often (the number of times) employees follow instructions given by three different managers, each of whom uses a different leadership style. Alternatively, observers might watch the behaviour of employees in various work situations and rate each interaction between employees and their managers on a scale from extremely negative (a score of one) to extremely positive (a score of 10). In both cases, the researchers are using rules to assign numbers to people's behaviour so that the data can be described and interpreted. However, the rules for assigning the numbers in each example differ. Which rule and therefore measurement scale is used will depend on what the researcher wants to find out and how they collect their data.

Four different scales of measurement used with data obtained from psychological research are called nominal, ordinal, interval and ratio scales.

Nominal scale: Numbers are attached to categories as labels ('names') to identify and differentiate between the data categories (and scores do not actually indicate amounts) e.g. data about eye colour could be organised into the categories of (1) blue eyes, (2) green eyes and (3) brown eyes. This is not unlike the way a footballer is identified by the number on their jumper. But the numbers indicate only that one footballer is qualitatively different from another. In research, if an experiment compares males and females, then this IV involves a nominal scale.

Ordinal scale: Ranks ('orders') categories of data according to some 'more or less' dimension e.g. from highest to lowest, best to worst, or always to never. Thus, the numbers of an ordinal scale show a relationship between numbered items and represent more than simply a category. The AFL ladder uses an ordinal measurement scale to indicate current performance of teams relative to one another. But the same amount does not separate categories in an ordinal scale e.g. 1st on the AFL ladder may be 4 points ahead of 2nd and 2nd may be 12 points ahead of 3rd. In research, if an experiment compares level of obedience to officers with different ranks, then this IV involves an ordinal scale.

Interval scale: Each number indicates an actual quantity and an equal amount separates all adjacent scores (i.e. equal intervals). Temperature in degrees Celsius is an example of an interval scale. The scale includes zero but it is not a true zero – it does not mean none of the variable is present; for example, 0 °C does not mean 'no temperature' or the complete absence of heat. In research, if the experimental conditions were based on whether participants were in a positive, negative or neutral mood, then this IV reflects an interval scale.

Ratio scale: Each number is an actual quantity, an equal amount separates adjacent numbers, and 0 truly means 'absolutely nothing there' (or none of the variable is present). Comparisons can be made between two numbers (or 'bits of data'); for example, it can be said that 100 kg is twice as heavy as 50 kg. In research, if the experimental conditions compare effects of 5, 10 and 15 mg of a new anxiety drug, then this IV involves a ratio scale.

In addition, any measurement scale may be either continuous or discrete.

A *continuous* scale allows for fractional amounts; it 'continues' between the whole number amounts and so decimals makes sense. For example, the variable of age is continuous because it can be represented as specifically as 16.74 years old.

In contrast, some variables involve a *discrete* scale, which can be measured only in whole-number amounts. For example, being male or female, or being in Year 10 or Year 11 are discrete variables, because you can be in one group or you can be in another group, but you can't be in between.

Usually, researchers assume that nominal and ordinal variables are discrete, and that interval or ratio variables are continuous.

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.31

Review questions

1. Define and distinguish between the following data types with reference to examples that are not used in the text.
(a) primary and secondary
(b) qualitative and quantitative

2. Distinguish between quantitative and qualitative data with reference to characteristics of your psychology class.
3. Complete the following table to apply your understanding of different data types.

eBook*plus*
Word copy of table

Research investigation	Primary (P) or Secondary (Q)	Qualitative (Ql) or Quantitative (Qn)
(a) A researcher compares the detail in paintings by people with a phobia and people with schizophrenia		
(b) A researcher compares the differences in visual perceptual abilities of kittens with and without damage to the visual cortex in the brain		
(c) A researcher observes how much time male and female adolescents take to get ready for a deb ball		
(d) A researcher reviews a YouTube mini-documentary showing participant responses during an experiment on the effects of playing violent video games		
(e) A researcher uses diary records kept by people hospitalised with a mood disorder to study their mental experiences		
(f) A researcher compares how infants who can walk independently respond when left alone with a stranger in a laboratory setting with infants who can crawl but are unable to walk independently		
(g) A researcher analyses the emotional content of a blog on the ethics of animal research		
(h) A researcher analyses participant scores on a test of recall in a study on long-term memory decline and ageing		
(i) A researcher uses open-ended questions to investigate how people feel when stressed		

4. Give an example of a strength and a weakness of quantitative and qualitative data.

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.32

Media analysis/response — psychological research reporting in the mass media

Option A
Select a contemporary newspaper, magazine or internet article that reports psychological research and has been published in the last calendar year (but is not a formal research report published in a journal).

Make a copy of the article (or the relevant excerpt/s) so that it can be presented on an A3 sheet of paper or within PowerPoint or similar. 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.
7. Comment on your choice of article for the purpose of this Learning Activity, e.g. why it captured your attention, aspects to which you wanted to draw the attention of others.

Option B
Select a contemporary electronic media item that reports psychological research, such as a TV program, movie, YouTube clip, Ted talk or podcast.

Make a copy of the item (or relevant extracts), its URL or another link that would give teacher access.

Write a one paragraph outline of the item, then answer the questions for Option A. Use screen shots, clips or other excerpts or links to exemplify answers or key points.

ORGANISING, PRESENTING AND INTERPRETING DATA

When data has 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 similar situations. In addition, suggestions for further research and evidence are often made.

To support all these requirements, researchers use statistics to analyse and describe the data they collect. They also use statistics to help interpret and give meaning to the results obtained from the research. Statistics are essentially mathematical procedures.

Two main types of statistics are used in psychology. **Descriptive statistics** are used for analysing, organising, summarising and presenting data obtained for a specific sample. They include calculations such a percentages and mean scores, and preparation of tables and graphs that help 'describe' the data. **Inferential statistics** are used for interpreting and giving meaning to the results. Like descriptive statistics, inferential statistics involve the use of mathematical procedures. However, unlike descriptive statistics, inferential statistics involve judgments about the results, especially conclusions and generalisations that may be made.

Descriptive statistics

Suppose that a researcher is interested in whether memory declines with age. In order to investigate this, some previously unseen information may be given to ten 10-year-olds, ten 25-year-olds, ten 40-year-olds, ten 55-year-olds, ten 70-year-olds and ten 85-year-olds. The research participants would be required to learn the information and then complete a memory test so that their memory could be assessed. In all, there would be 60 bits of data (test scores) about the memory of participants in different age groups. How could the researcher make sense of all these different bits of information so that meaningful conclusions about memory and age might be drawn?

The first step would be to use descriptive statistics to analyse, organise, summarise and describe the data so that it can be interpreted. It is difficult to draw conclusions about whether memory declines with age by looking at 60 individual scores (see Table 1.7). Thus, in order to compare the memory scores of people in the six different age groups to determine whether there has been a decline in memory with age, the data for each age group could be summarised and presented like that in Table 1.8.

The table provides some order to the data by organising the individual scores into age groups, but comparison of scores across the different age groups is still difficult so Table 1.7 is not particularly useful (nor suitable for inclusion in a report).

TABLE 1.7 Individual participants' scores on a test of memory

Age (years)	Participants' scores
10	14, 11, 9, 10, 15, 16, 14, 12, 13, 11
25	14, 16, 16, 18, 13, 17, 14, 15, 17, 8
40	17, 15, 12, 16, 19, 10, 18, 14, 13, 18
55	10, 18, 13, 14, 15, 14, 12, 19, 12, 10
70	13, 10, 12, 16, 7, 15, 9, 12, 11, 8
85	6, 14, 12, 10, 11, 9, 16, 10, 8, 13

To better enable the scores from different age groups to be compared, a single number that summarises the data for each age group could be calculated. For example, the researcher could calculate the *mean* score on the memory test for each age group. The mean scores could be used to describe the 'average' performance on the memory test for each age group and would enable the researcher to compare the different age groups, as shown in Table 1.8. A graph such as a bar chart could also be prepared to visually represent the results, as shown in Figure 1.75 on the next page. In this way, trends, patterns and relationships in the data may become apparent.

Generally, the specific type of descriptive statistic used depends on the kind of research and on the type of data collected. Some descriptive statistics are more suited to particular research and data than others.

TABLE 1.8 Mean scores for each age on a test of memory

Age (years)	Mean scores
10	12.5
25	14.8
40	15.2
55	13.7
70	11.3
85	10.9

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

UNCORRECTED PAGE PROOFS

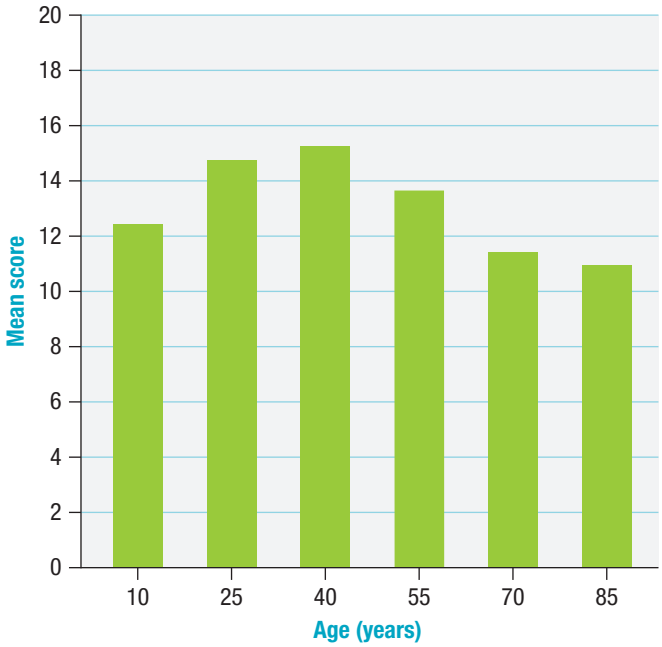


Figure 1.75 Frequency distribution of scores on a memory test

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. A commonly used table in psychology is called a frequency distribution.

A *frequency distribution* is a way of organising data to show how often ('frequently') a value or measure (such as a score) occurs in a set of data. A frequency distribution can also be presented as a graph.

Table 1.9 is an example of a frequency distribution for scores obtained by males and females in an experiment. It shows all the possible values of what has been measured (organised into groups or categories called *class intervals*) and the number of times each value occurs in the set of data (the number of individuals in each class interval). In a frequency distribution, the scores are often arranged either from the highest to the lowest score or from the lowest to the highest score, so that data are presented in an orderly, logical way.

TABLE 1.9 Example of a frequency distribution of scores by males and females

Scores	Males	Females
20–24	0	0
15–19	1	2
10–14	3	5
5–9	4	2
0–4	2	1
Total	10	10

When there is a large number of scores, it is often useful to organise the scores into class intervals, then total the number of scores for each class interval. The class interval can be any size within the range of scores, but the size of each class interval should be consistent across all scores. Intervals of five or ten units are typically used. If an interval of five is used (as in Table 1.9), then the difference between one interval and the next is five; that is, 0–4, 5–9, 10–14, and so on.

Note that the labelling of the table (e.g. its title and headers) is just as important as its contents. Some conventions ('standards'), for tables used in psychology are:

- All tables should be 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 that explains what the table is about without being too long, e.g. *Individual Participant Scores on a Test of Memory*
- The table number and title should be on separate lines with the table number above the title e.g. Table 1 *Individual Participant Scores on a Test of Memory*
- 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'.

Although tables are an effective means of recording data, they may not be the best way to show trends, patterns or relationships. Graphical representations are more suited to this purpose.

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). When drawing a graph for experimental research data, the IV is represented on the horizontal (x) axis and the DV is represented on the vertical (y) axis.

Graphs are best used to determine and communicate trends, patterns or relationships 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, and how one variable may be related to or change in relation to another.

Various types of graphs display data in different ways. The kind of graph used depends mainly on the nature of the research that was conducted (e.g. experimental or self-report) and the type of variables investigated (e.g. categorical or numerical) as described in Box 1.4 on page 26.

UNCORRECTED PAGE PROOFS

Among the more commonly used graphs in psychology are bar charts and line graphs. For example, bar charts (and pie charts) can be used to display data in which one of the variables investigated is categorical and line graphs can be used to display data in which both the IV and DV are continuous. In addition, as shown in Box 1.12 (page 61) on correlational research, lines of best fit can be used to illustrate the underlying relationship between variables and scatter plots can be used to show an association between variables (VCE Advice for teachers: Psychology). 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

A bar chart is a graph that uses a series of discrete (separate) bars or rectangles next to, but not touching one another, to enable comparisons of different categories of data (i.e. 'categorical data'), as shown in Figure 1.76 below. 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 axis is used to show the frequency with which each category occurs (e.g. how often, how much).

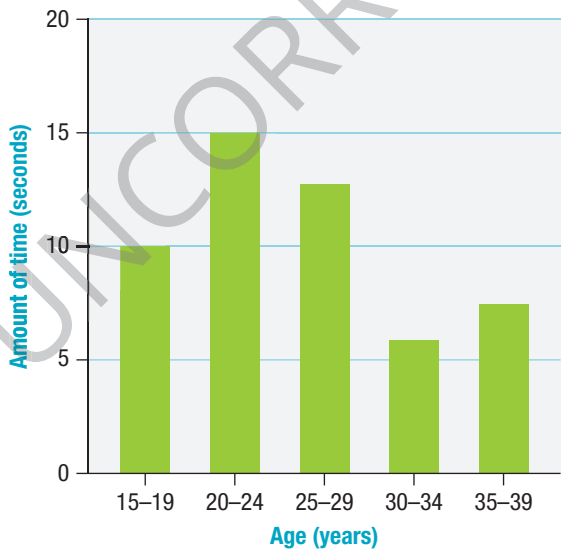


Figure 1.76 An example of a bar chart

One important feature of a bar chart is that each of the categories shown in the graph is discrete 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. In addition, the bars can start from either the *x*-axis (vertically) or from the *y*-axis (horizontally).

Sometimes a bar graph is used to represent values from two categories; for example, scores obtained by age group (e.g. amount of time to solve a problem) and by sex. This is shown in Figure 1.77a. The data for two categories can also be presented within a single bar. This is shown in Figure 1.77b which combines results for males and females within single bars.

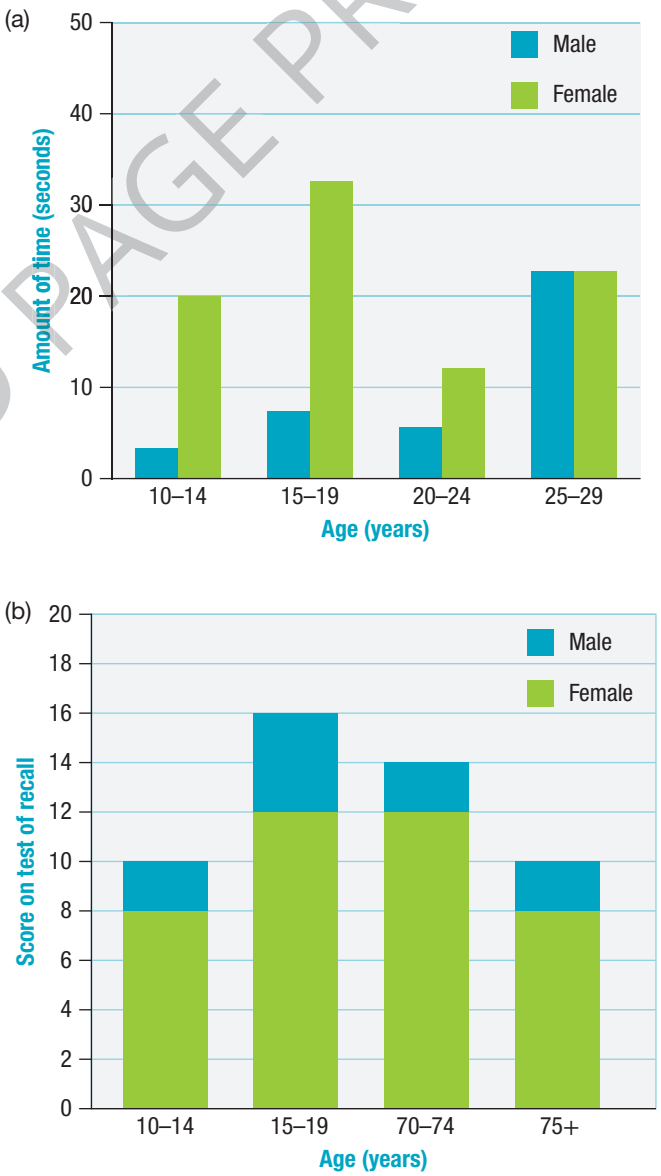


Figure 1.77 Two bar graphs showing different ways of presenting data for males and females of different ages

UNCORRECTED PAGE PROOFS

Line graphs

A line graph uses points connected by lines to show how one variable changes as another variable changes. For example, Figure 1.78 below shows how performance on a speed and accuracy test (e.g. matching symbols with numbers) changes in relation to the number of hours of sleep a person has had. You can see at a glance that the number of errors (the measure of performance) was the greatest when only one hour of sleep was obtained and that the number of errors tended to decrease as the amount of sleep obtained increased.

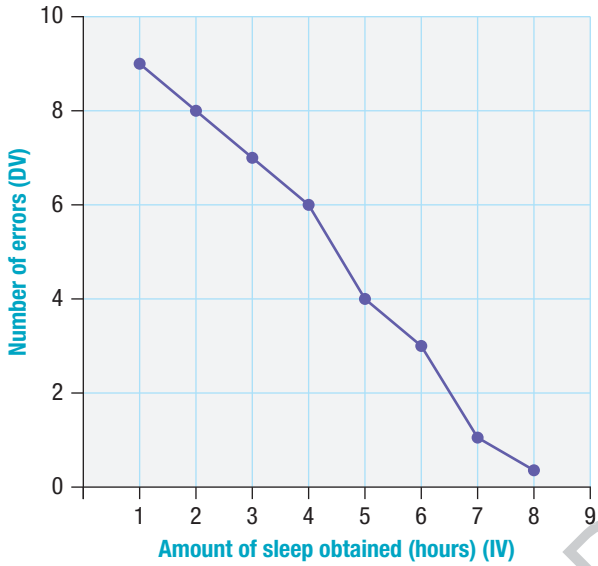


Figure 1.78 A line graph showing the results of an experiment investigating the effect of amount of sleep (IV) on performance on a speed and accuracy test (DV)

When used to show the results of an experiment, the IV is usually plotted on the horizontal (*x*) axis, with the numerical value of the data increasing as you go along this horizontal axis from left to right.

As shown in Figure 1.78, a line graph that describes the relationship between amount of sleep obtained and test performance would list the amount of sleep in hours on the *x*-axis in intervals; for example, beginning at zero, then one, two, three, four hours and so on. One important feature of a line graph is that the variable plotted on the *x*-axis is continuous; that is, a series of progressively increasing values can be listed.

The vertical (*y*) axis usually has the DV (ie. the measure of performance) plotted along it. A line graph that described the data from the experiment on the amount of sleep obtained and test performance would record the test scores (e.g. a total correct score or number of errors) along the *y*-axis in intervals, beginning at zero. This is also shown in Figure 1.78.

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.

A number of different sets of data can also be plotted on the one graph. For example, in Figure 1.79 there are three sets of data showing age-related performance on a problem-solving task following different amounts of sleep deprivation. To identify the results of different age groups, a different coloured line has been used for each set of data. Note too the use of dots to identify the points of intersection between data for the X and Y axes.

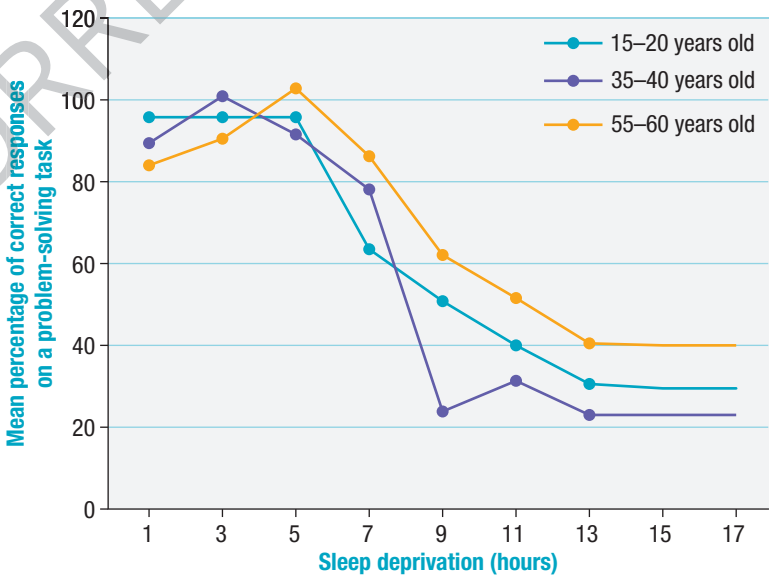


Figure 1.79 A line graph showing three sets of data so comparisons can be made

UNCORRECTED PAGE PROOFS

BOX 1.24 Histogram and pie chart

A *histogram* is like a bar chart but the bars are drawn so that they touch each other because the data are continuous.

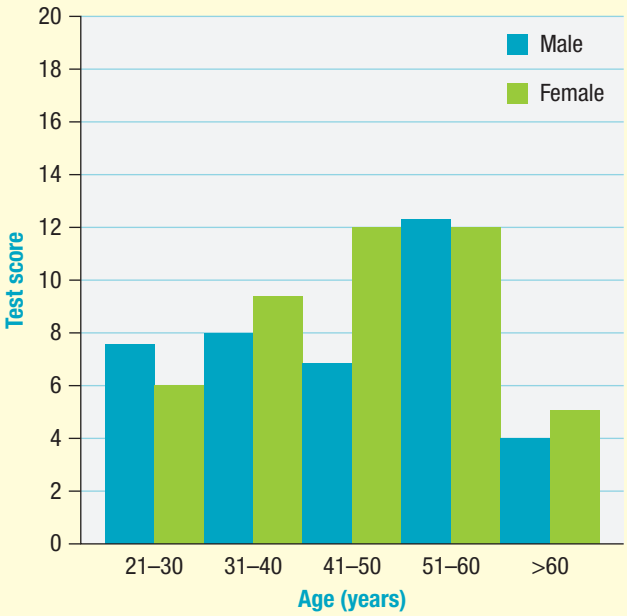


Figure 1.80 An example of a histogram

A *pie chart* (also called a *pie graph*) shows the proportions of scores, values or cases within a set of data. The differences between the categories in a set of data are represented by different-sized 'slices of pie'. This type of graph is best used for categorical data and to compare different parts of the same whole.

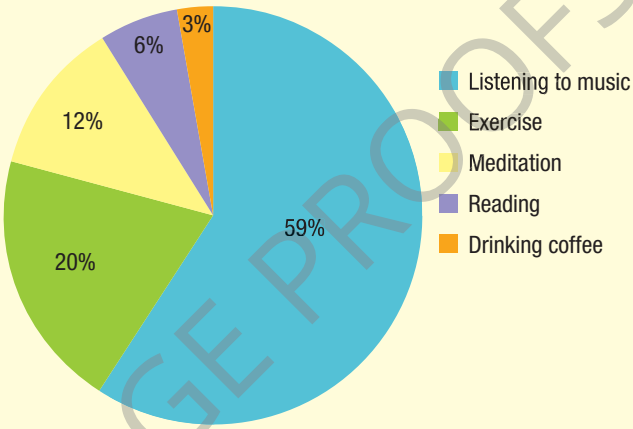


Figure 1.81 An example of a pie chart showing relaxation techniques used by participants when alone

UNCORRECTED PAGE PROOFS

Percentages

Suppose you conduct an observational study to find out whether more year 7 boys engage in disruptive behaviour in the classroom than do year 7 girls. You want to obtain quantitative data, so you work out a list of observable behaviours that you consider to be disruptive in the classroom. Your list includes behaviours such as distracting other students, calling out in class and being out of the seat without permission.

You observe two different classes from a store room at the back where you have concealed your presence. Whenever you see a boy or girl demonstrating one of the disruptive 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 engage in a disruptive behaviour and four of 16 girls observed are disruptive on at least one occasion.

On the basis of these results, more boys than girls were disruptive at least once. 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 engaged in disruptive behaviour, then making a comparison.

A percentage is a descriptive 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 percentage 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:

$$\frac{90\ (\text{subtotal})}{100\ (\text{total})} \times \frac{100}{1} = \frac{90 \times 100}{100} = \frac{900}{100} = 90\%$$

For the data obtained in the observation study described above:

$$\text{Boys: } \frac{6\ (\text{subtotal})}{25\ (\text{total})} \times \frac{100}{1} = \frac{6 \times 100}{25} = \frac{600}{25} = 24\%$$

$$\text{Girls: } \frac{4\ (\text{subtotal})}{16\ (\text{total})} \times \frac{100}{1} = \frac{4 \times 100}{16} = \frac{400}{16} = 25\%$$

This means that the proportion of boys (calculated 'out of 100') who were disruptive in the classroom 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 sociocultural group (such as gender, age, educational qualifications, country of birth and language spoken at home).

Mean as a measure of central tendency

Data are often summarised by determining a single numerical score that can describe the data for the whole group(s). This score, called a **measure of central tendency**, indicates the 'central' or 'average' value in a set of scores. When a measure of central tendency is calculated, it often provides a typical score for a set of scores. 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 together all the scores and dividing the total by the number of scores. For example, if ten rats were put into a maze, the length of time (in seconds) it might take each rat to reach the end of the maze could be as listed below:

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

The mean ($\bar{X}$ or M) score for the group is calculated by adding the individual scores together ($\Sigma = 195$), then dividing the total by the number of individual scores ($N = 10$). In this example, the mean is 19.5 seconds. The formula for the mean is shown below.

$$\bar{X}(\text{mean}) = \frac{\Sigma\,(\text{sum or total of all scores})}{N(\text{number of scores})}$$

When scores in a set of data cluster closely around a central score, the mean is a suitable and fairly accurate indicator of the typical score. It is representative of the scores. If, however, the scores are widely spread, unevenly distributed or cluster around extreme values, then the mean can be misleading. For example, 'outliers' are scores (or values) that 'lie' well 'outside' most of the other scores in a set of data. An outlier is typically much smaller or much larger than most scores. A few very high or very low values (sometimes even one depending on the number of scores) can inflate or deflate the mean, thereby making it an inaccurate indicator of the typical score. For example, in the set of scores 52, 58, 63, 17, 54, 61 and 92, both 17 and 92 lie a long way from the other results and are therefore outliers.

When it seems that one or more outliers have a big effect on the mean, then the outliers should be analysed and accounted for, rather than being automatically dismissed. Checking the measure, the specific result or some other aspect of the data collection procedure may be useful in further examining an outlier (VCE Advice for Teachers: Psychology). Alternatively, another measure of central tendency may be considered, such as the mode and median described in Box 1.25 below.

UNCORRECTED PAGE PROOFS

BOX 1.25 Median and mode as measures of central tendency

When scores in a set of data cluster closely around a central score, the mean is a fairly accurate indicator of the 'typical' score; that is, it is representative of the scores. If, however, the scores are very widely spread, unevenly distributed or cluster around extreme values, then the mean can be misleading. For example, a few high or low values ('outliers') within a relatively small set of data may inflate (increase) or deflate (decrease) the mean. In such cases, another measure of central tendency will be a more accurate measure of the 'typical' score and would therefore be used. Two other measures of central tendency which can be considered are the median and the mode.

Median

Another way of obtaining a score that may represent the central point in a set of scores is to arrange the scores in order of size and select the score that falls in the middle as being typical of the whole set of scores. This score is called the median.

The *median* is the middle score (or mid-point) of a set of scores. For example, the time taken (in seconds) for each child to complete a jigsaw puzzle in rank order (from lowest to highest) is:

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

In this example the median is 18. When there is an even number of scores, the median is the average of the two middle scores. For example, if the two middle scores are 20 and 21, the median would be 20.5.

The median is a particularly useful descriptive statistic if there is a limited amount of data, but if there is a large amount, determining the median can be time consuming and often impractical. The median is also a useful statistic when many very high or very low scores occur in the set of scores because the median is not affected by extreme scores. For example, the test scores shown in Table 1.10 on the next page were obtained when a psychology teacher gave her class of 10 students a test on research methods in psychology.

(continued)

(continued from previous page)

TABLE 1.10 Test scores

Rank	%
1	98
2	91
3	91
4	60
5	59
6	57
7	57
8	57
9	56
10	54
Total	680
Mean	68
Median	58

The calculation of the mean score on the test does not provide an accurate impression of the average score on the test, because the inclusion of three very high scores inflates the mean figure. In situations such as this, the median is a more accurate reflection of the 'typical' score on the test as it is closer to the majority of scores in the set of data.

Mode

A third measure of central tendency is the mode. In everyday language, the word mode means 'common'. This term accurately describes what the statistical mode is; that is, the *mode* is the most frequently occurring score in a set of scores. Using the scores again for the children completing the jigsaw puzzle:

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

the mode would be 17 because it occurs three times.

The mode is infrequently used in statistics because it is often not typical or representative of a complete set of data. For example, if a set of scores is 1, 1, 6, 7, 8, 10, the mode would be 1, which is not a representative score of the entire group. If one of the scores of 1 is changed to 10, the mode shifts completely to the opposite end of the scale. Thus, a single score can alter the mode dramatically, which is in contrast to the median, and to a lesser extent the mean, where individual score changes tend to have less of an effect.

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 1.82 below 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 they are 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*).

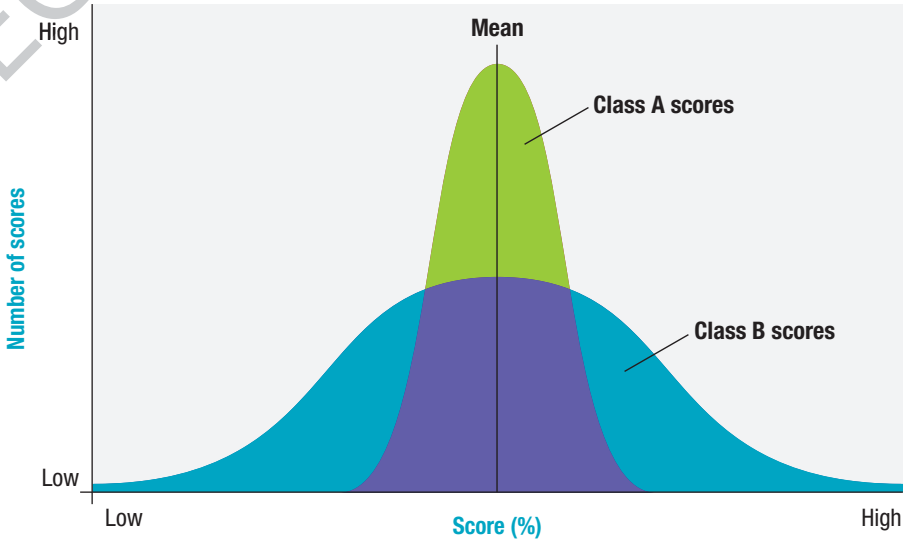


Figure 1.82 Class A and Class B both have the same means. The distribution of scores for Class A (green) shows low variation, as indicated by the clustering of scores around the mean. The distribution of scores for Class B (blue) shows high variation, as indicated by a greater spread of scores from the mean.

UNCORRECTED PAGE PROOFS

The **standard deviation (SD)** 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, as shown in Figure 1.83, curve C. The higher the standard deviation, the greater the variation there is among the scores, as shown in Figure 1.83, curve A.

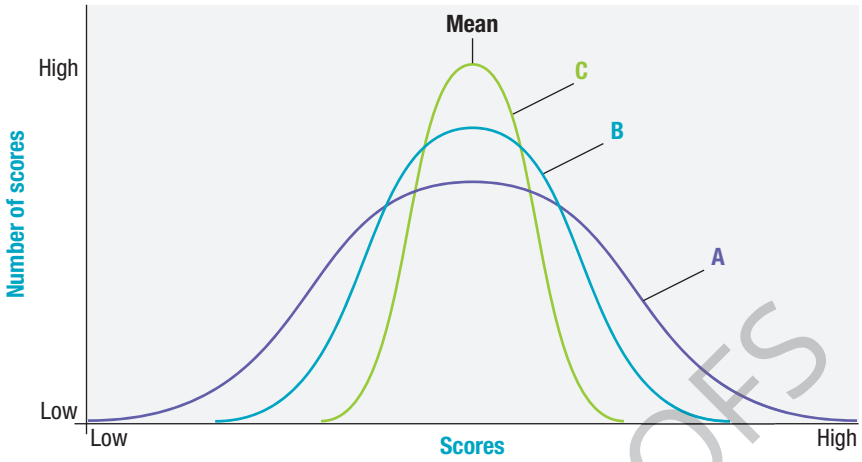


Figure 1.83 This graph shows three distributions of scores, each with a different standard deviation. The purple curve has the highest standard deviation and the orange curve has the lowest standard deviation.

The standard deviation is a particularly useful descriptive statistic in that it provides a point of comparison between the means and the spread of two or more different sets of score. 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.

In sum, when considering standard deviations, it is important to recognise that:

- although two or more different sets of scores (or data sets) may have the same mean, they may not have the same degree of variation (or 'spread') in the data; and
- a higher standard deviation represents a greater variation (or 'spread') in a set of scores (and vice versa).

Note also that, for a normal distribution of any set of scores, 68.26% of the scores lie within one standard deviation of the mean and 95.44% of the scores lie within two standard deviations of the mean. These and other standard deviation values are shown in Figure 1.84 below.

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.

Although you need to be able to calculate a mean in VCE Psychology, calculation of standard deviation or any other measure of variation is beyond the scope of the VCE Psychology Study Design (VCE Advice for teachers: Psychology).

studyon

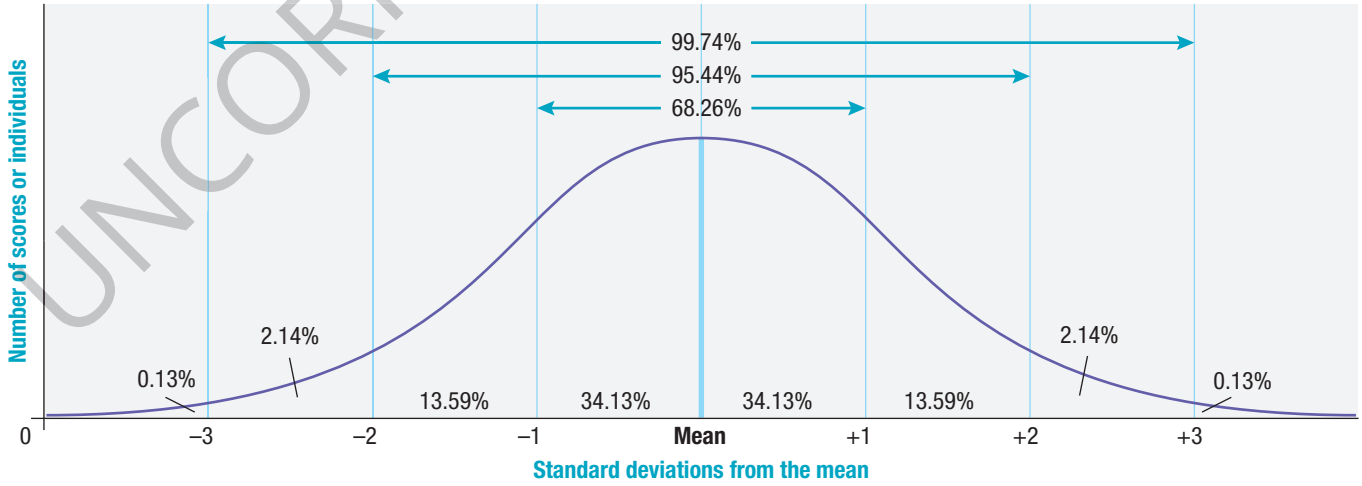


Figure 1.84 Standard deviations in a normal distribution. When standard deviations are represented on the X (horizontal) axis of a normal distribution curve, the percentage of scores falling between the mean and any given point on the axis is always the same.

UNCORRECTED PAGE PROOFS

BOX 1.26 Skewed distributions of scores

The normal distribution curve shown in Figure 1.84 is a 'theoretical ideal' and is rarely perfectly achieved in reality. Often, the scores or other values are unevenly distributed and cluster to the left or the right ends of the graph. In such cases, the spread is called a *skewed distribution* as there is a lack of balance or symmetry. The skew of the graph — whether it is positive or negative — is linked to the direction of its 'tail'.

Positive skew

If the number of words 12-month-old children spoke were plotted, it is highly likely that many of the scores would cluster towards the lower end (left) of the graph producing a positively skewed distribution (or spread). This is shown in Figure 1.85 below.

When the distribution of scores has a *positive skew*, there is a disproportionate number of low scores. The 'tail' of the graph tapers in a positive direction towards the higher scores.

Negative skew

In contrast with 1-year-olds, if the number of words 16-year olds know were plotted, many of the scores would cluster at the higher end (right) of the graph producing a negatively skewed distribution. This is shown in Figure 1.86 below.

When the distribution of scores has a *negative skew*, there is a disproportionate number of high scores. The 'tail' of the graph tapers in a negative direction, towards the lower scores.

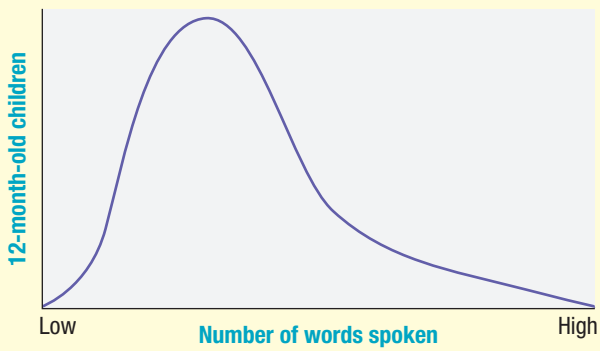


Figure 1.85 Positive skew — a curve skewed to the left, indicating that there is a clustering of a relatively large number of low 'scores'

Figure 1.86 Negative skew — a curve skewed to the right, indicating that there is a clustering of a relatively large number of high 'scores'

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.33

Review questions

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

2. (a) What information does the standard deviation provide about the distribution of scores?
(b) Explain why the standard deviation is useful for comparing two or more data sets obtained from research that have very similar or even identical means.

3. What percentage of scores lie within one standard deviation of the mean? Two standard deviations?

4. For each of the following examples, indicate whether you believe there is likely to be high or low variability among the data from Australian samples.
(a) age when infants first walk unassisted
(b) age when infants utter their first sound
(c) age when people go on their first date
(d) speed at which people successfully navigate a complex maze
(e) goal-shooting accuracy of professional netballers who usually play a goal attack or goal shooter position but under a condition of high arousal

5. (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 75%.
Class B: mean 75%.
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 75%; standard deviation: 0.1
Class B: mean 75%; standard deviation: 2.5.
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 summarising and organising data so that it can be comprehended. For example, they can show patterns that emerge from data or provide a single number such as a mean or standard deviation that can help describe the data. However, they do not in themselves indicate whether the data are meaningful.

In order to interpret the data and find out whether the results are meaningful, inferential statistics can be used. Inferential statistics enable researchers to 'infer', or to make judgments or draw conclusions on the basis of some evidence, such as the data collected through research. For example, an inference from data may involve deciding whether manipulation of an independent variable alone is responsible for the changes in the dependent variable. Inferential statistics can help the researcher do this with some degree of confidence.

Of initial interest to the researcher for any type of study is whether the data support the hypothesis that was tested. Furthermore, when results have been obtained for a sample, inferential statistics can be used to help decide whether these would occur in the population from which the sample was drawn.

In most experiments there will usually be a difference in the mean scores of the experimental and control groups. Suppose, for example, a researcher wanted to find out if a new study technique called SupaStudy improved performance on a Psychology test. The researcher conducted an experiment using two groups of Units 3 and 4 Psychology students. The students were randomly allocated to either of two groups. Group A was taught the SupaStudy technique and Group B was not. Both groups attended lunchtime sessions three times a week for a month during which they were taught a new topic in Psychology. At the end of the topic, both groups were given the same test. The mean test score for Group A was 79% and the mean test score for Group B was 73%. On the basis of these results, what legitimate conclusion can the researcher draw about the effect of using SupaStudy on Psychology test performance?

Assuming that the researcher controlled all variables that may have affected the results obtained, other than the use or non-use of SupaStudy, can the researcher draw a conclusion that SupaStudy 'works'? More specifically, can it be concluded that the difference in the mean scores is due to the presence or absence of the IV — the use of SupaStudy?

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, by chance, there were a few individuals in the control group who were far more able or motivated than students in the experimental group who used SupaStudy. This may have inflated the control group's mean score. Perhaps, by chance, there were a few individuals in the experimental group who were distracted when learning SupaStudy, missed some vital details and did not apply the technique correctly. This may have deflated the experimental group's mean score. Perhaps the classroom conditions were slightly better for one group and marginally impacted on performance. Perhaps the experimenter was tired and slightly impatient when instructing the experimental group participants. Perhaps the effects of these apparently controlled variables 'added up' and combined in such a way as to cause the difference.

There's no doubt that 79% is a higher score than 73%, but the difference does not seem to be very large. Is the difference in the scores large enough to claim that SupaStudy had a real effect or is the difference in the scores simply due to chance factors? What is an *acceptable* difference? How big does the difference between the mean scores of the two groups need to be in order to say that the difference is not due to chance?

Figure 1.87 These children were participants in a research investigation on sleep patterns conducted by VCE students. They were part of a sample of eight participants who all attended the same eastern suburbs preschool. Can the results obtained from this sample be generalised to all preschoolers at the centre in the year of the study (i.e. the population)? To all children enrolled at the preschool the following year? Can the results be generalised to wider populations, such as all preschoolers in the eastern suburbs? In the Melbourne metropolitan area? In Victoria? Other populations? Inferential statistics can be used to make these types of judgments about results.

UNCORRECTED PAGE PROOFS

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 (see Box 1.27 below).

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 research mean. All conclusions must be based on evidence (i.e. the results), be consistent with the evidence and relevant to what was actually investigated and take account of potential limitations of the research.

One type of conclusion relates to whether or not the hypothesis is supported on the basis of the results obtained. This requires careful examination of the results so that an objective ('unbiased') judgment can be made. Although the results alone may indicate that the hypothesis is supported, the DV and therefore results may have been influenced in a significant way by variables other than (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 (or chance factors).

The conclusion about the hypothesis is expressed as a statement in a written report that describes the investigation and its findings. In psychology, a hypothesis may be supported or it may be refuted (rejected), but it cannot be said to be 'proven' true or correct. Generally 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 or judgment 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

UNCORRECTED PAGE PROOFS

BOX 1.27 Statistical significance and *p* values

Statistical significance

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.

***p* value**

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.

extended to apply to an entire class of objects, events or people (the sample's 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 about the results and making generalisations, 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
- all conclusions are relevant to what was actually investigated
- any influential extraneous variables or 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.

In many cases, psychology researchers use university students enrolled in psychology courses as participants in their experiments. Some researchers believe that the results of these experiments cannot be generalised beyond the sample; for example, by applying the results to all students enrolled in a psychology course at a university, to students in other courses, to young adults, or to adults in general. However, other researchers believe that it is reasonable to assume that a relationship between an IV and a DV observed in one group of people is likely to be seen in other groups, as long as the relationship is strong and the sample of participants is not particularly unusual. For example, results from a sample of people with an addictive disorder or an intellectual disability cannot be generalised as easily as data obtained from university students.

The extent to which results can be generalised also depends on the topic studied. Many psychologists believe that researchers who study topics such as sensory processes and biological or physiological factors that underlie behaviour can more easily generalise their results quite widely. The phenomena that these researchers study are usually not affected by individual differences. For example, the visual system is basically the same for all fully developed people (without relevant nervous system damage), therefore, differences between individuals in characteristics such as intelligence, beauty, personality, age, weight, political preferences, willingness to volunteer, and so on, will not affect it. Consequently, it is widely believed that it is reasonable to generalise findings on these topics beyond the populations from which they were drawn, as long as such generalisations are tentative.

Researchers who study topics such as personality, social behaviour, attitudes, consciousness and learning or memory strategies tend to be less likely to generalise their findings beyond the sample's population. In these areas, the effect of the IV is often influenced by the individual characteristics of the person participating in the experiment. For example, a new technique for learning mathematics may be effective for university students but not for secondary school or primary school students (Wood, 1981).

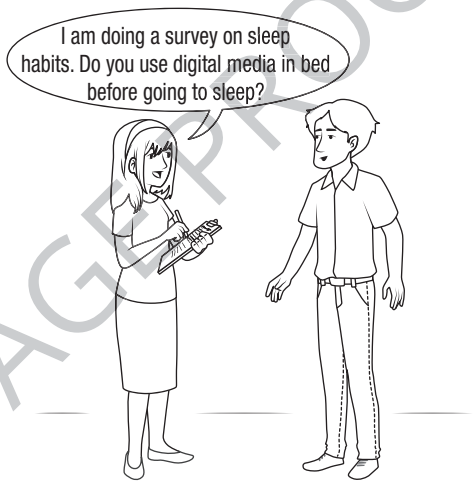


Figure 1.88 On the basis of a sample of one, can the results be generalised?

UNCORRECTED PAGE PROOFS

study on

LEARNING ACTIVITY 1.34

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. What kind of judgment is made about the research hypothesis after the results have been obtained and analysed?
4. What is the meaning of the term conclusion?
5. Why must potential extraneous and confounding variables be considered when drawing conclusions from results obtained in a study?
6. What is the meaning of the term generalisation?
7. What are the important considerations when drawing conclusions and making generalisations from the results of a study?
8. Distinguish between the terms conclusion and generalisation as they apply in research studies.
9. How would you answer the question in the Figure 1.88 caption above?

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 the results are of value and use. Reliable and valid results can be achieved when the research, its data collection procedures and measurement tools are reliable and valid.

Although researchers may refer to a study or results as either 'reliable' or 'not reliable', 'valid' or 'not valid', reliability and validity are not necessarily 'present-or-absent' features of a research design, its measurement tools or the results. Instead, they are considered to vary in degree on a scale ranging from low to high. Both are ways of assessing the quality of a research design, the specific procedures or tools used to collect data, and the results obtained.

Reliability

Reliability refers to the extent to which the results obtained from a research study are consistent, dependable and stable. This means that each time a behaviour or event is measured under the same conditions, the procedure(s) used should produce very similar results at the least. For example, if your body temperature was measured with an oral thermometer while you were laying down in a resting state and the reading was double-checked immediately, you should expect the same result. Similarly, if you conducted an experiment on a group of participants and repeated it again with the same group or a similar group under the same conditions, you should expect the results to be very similar.

Conducting an experiment is a more complicated process than measuring your body temperature level because it involves human participants and the strict control of many variables, so it is not likely, or expected, that the results will be identical each time the experiment is conducted. The main reason the results of an identical study are unlikely to be exactly the same when replicated is due to individual participant differences within another sample.

Even when a researcher repeats their experiment with the same participants it is unlikely that identical results will be obtained due to participant differences in responses to the IV manipulation and/or DV measurement, even if these are only minor differences. For example, consider the likelihood of your responding exactly the same way as you did in the original experiment if you are a participant in an identically repeated experiment, especially an experiment for which responses are relatively unrestricted. However, researchers believe that if the results of an experiment are to be considered reliable, they should be *similar* (i.e. fall within a very narrow range of values) each time the experiment is repeated under the same conditions on other occasions.

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.

In addition, experiments that use human judgment may not always produce reliable results due to potential bias. For example, in an experiment where participant responses are recorded by observers, there is a possibility that observers may rate certain responses differently depending on what they are attending to, their concentration span, the time of day, their mood, and so on, even if a checklist is used. Consequently, such an experiment may be inherently less reliable than one using a stopwatch where time is crucial. It is also more likely that experiments that use small sample sizes or an insufficient number of trials will produce results that are not reliable when compared with similar experiments that use large samples and an adequate number of trials.

Validity

Validity refers to the extent to which the research study and its procedures measure what the research intended to measure. Basically, the research design and the specific procedures used should match the requirements of the investigation to produce results that are relevant to the aims of the research. If a research study has a high level of validity, this means that the investigation has produced results that accurately measure the behaviour or mental process that was claimed to have been measured. For example, an experiment on some aspect of memory would use a test that actually measured that aspect of memory and not something else such as intelligence, language ability, numerical skills or emotional state.

Validity also relates to the conclusions (including any generalisation) the researcher makes about a study. In this case, the results are considered 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 other colas, the research has a high level of validity only if the new drug really works or if the participants really did prefer Coca-Cola (Stangor, 2004).

As with seeking reliability, researchers always attempt to conduct valid research and to draw accurate conclusions from their data. Yet often, despite a

UNCORRECTED PAGE PROOFS

Figure 1.89 If a researcher concludes that participants in a taste-preference study preferred Coca-Cola over other colas, the research has a high level of validity only if the participants really did prefer Coca-Cola.

researcher's best intentions, their research has low validity or is 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 lacking validity 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. Consequently, when researchers evaluate the validity of a study, they consider both the research and its implementation.

Note that a measure can be reliable even though it is not valid, but a measure cannot be valid unless it is reliable. 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 (even if inaccurate). Similarly, it is possible to obtain a reliable measurement for skull size using a stretched cloth tape measure, but that would not be a valid measure for intelligence.

Figure 1.90 Researchers seek to conduct investigations that are both reliable and valid. Reliability is a necessary condition for validity.

BOX 1.28 Replication of procedures: repeatability and reproducibility

Experimental data and results must be more than one-off findings and should be repeatable and reproducible to draw reasonable conclusions.

Repeatability refers to the closeness of agreement between independent results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus and/or same laboratory).

Reproducibility refers to the closeness of agreement between independent results obtained with the same method on identical test material but under different conditions (different operators, different apparatus and/or different laboratories). The purposes of reproducing experiments include checking of claimed precision and uncovering of any systematic errors that may affect accuracy from one or other experiments/groups. Reproducibility is often used as a test of the reliability of an experiment.

Source: VCE Advice for teachers: Psychology. Retrieved from http://www.vcaa.vic.edu.au/Pages/vce/adviceforteachers/psychology/measurementterms.aspx

UNCORRECTED PAGE PROOFS

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, both its design and implementation. Strengths and limitations of different types of investigations and their procedures and results can be discussed in terms of these aspects of validity. As with reliability and validity in general, these types of validity also vary in degree; that is, a research study is considered to be more or less valid rather than valid or not valid in either or both these aspects. However, without internal validity, there is no external validity.

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 due to some other factor. For example, in an experiment, the researcher needs to be confident that the measured change in the DV was produced solely by the IV and not by any extraneous or confounding variable, nor due to chance. It can therefore be inferred that there is a causal relationship between the IV and DV.

When a study is said to have internal validity, then the measurement tools and procedures used for the research measured what they were supposed to measure. Internal validity is maximised when the research is free from the presence of confounding variables. If a study has gaps or 'flaws' in its procedures, such as the use of a sampling procedure that resulted in an unrepresentative sample when it was important to have a representative sample, then it may be considered as lacking in internal validity if the biased nature of the sample produced one or more confounding variables.

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

Internal validity can be improved through a number of methods, especially use of a research design that is appropriate for testing the research hypothesis and

UNCORRECTED PAGE PROOFS

Figure 1.91 The term *ecological validity* is specifically used to refer to the extent to which the findings of a research study are able to be generalised to everyday, common real-life behaviours and natural settings. Studies that take place under highly controlled artificial conditions are often closely evaluated for ecological validity. Consider Zimbardo's (1971) Stanford Prison Experiment. The study took place in a very realistic prison environment set up in the basement of the university's psychology building and participants displayed realistic behaviour. However, it was an artificial environment and not a real prison. To what extent did this situation reflect real life in a real prison? Did the results generalise to prisons in the outside world? How likely is it that participants simply behaved in ways they thought they should be behaving in that setting and acted out their assigned roles? Do you think the research is high or low in ecological validity?

by controlling relevant extraneous variables to ensure none become confounding variables; for example, by using appropriate sampling procedures for selection and allocation of participants, counterbalancing, single and double blind procedures, placebos, and standardised instructions and procedures.

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 and over time. 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 results 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 at another point in time.

Similarly, if an observed effect may actually be found only under certain conditions (e.g. in a laboratory) for specific groups of participants (e.g. university students obliged to participate as a course requirement), then a study *may* be lacking in external validity. In addition, the results should not be time-dependent; that is, the results should apply across time and be found in the future if the research were to be replicated under the same conditions.

Generally, the bigger and more representative a sample is of its population, the more confident the researcher can be in generalising from the sample to the population. Conducting an experiment in a real-world setting natural to the research question of interest and therefore more like an event in 'real life' can also improve external validity.

Internal and external validity are related. Internal validity is a precondition of external validity, which means that a study cannot have external validity without internal validity. Furthermore, a study that is said to have external validity is also said to have internal validity. It does not necessarily follow, however, that an effect observed in a strictly controlled laboratory experiment with a high level of internal validity will also have the same effect in a real-world situation.

LEARNING ACTIVITY 1.35

Review questions

1. Distinguish between validity and reliability in research.
2. Is it essential that the results of an experiment can be replicated in order for the experiment to be considered reliable? Explain your answer.
3. Give an example of when results would not be considered reliable.
4. Under what circumstances can it be said that the conclusions or findings of research are 'valid'?
5. (a) Distinguish between internal and external validity with reference to an example.
   (b) List three procedures that could adversely impact on the internal validity of an experiment.
   (c) List three procedures that could improve the external validity of an experiment, with reference to a relevant example.
   (d) Explain, with reference to an example, whether a study can have high internal validity but not external validity.
6. Explain, with reference to an example, why reliability is possible without validity but validity requires reliability.
7. Find an example of an online test on some aspect of how people may think, feel or behave. This could include personality and intelligence tests.
   (a) Name the test and insert a copy of the URL where it can be located.
   (b) Comment on the validity and reliability of the test.

LEARNING ACTIVITY 1.36

Designing an experiment and interpreting the results

Outline the design of an experiment that could be undertaken to test the following hypothesis: 'If a newborn infant has extra contact with their mother soon after birth, then the bond formed between the mother and infant will be stronger.' Twelve mothers who are about to give birth have volunteered to be participants in your experiment for two years. In designing your experiment, make sure that you address the following:

- What groups will be used?
- What are the independent and dependent variables?
- How will these variables be operationalised?
- What is your research hypothesis for the experiment?
- What potential extraneous or confounding variables will it be important to control?
- What type of experimental design will you use? Explain your choice of design.
- When considering your results, what question will you ask about your experiment's:
  – internal validity?
  – external validity?
- If you find an acceptable difference in the results for the two groups, what conclusion would you draw? What generalisation would you make?
- How could you determine the reliability of your results?

UNCORRECTED PAGE PROOFS

ETHICS IN PSYCHOLOGICAL RESEARCH AND REPORTING

Is it appropriate to expose human participants to stressful conditions in order to study bodily changes when stressed? Is it appropriate to deprive participants of sleep for a prolonged period in order to study the effects? Is it appropriate to deceive participants and misinform them of what an experiment is about in order to control responses that may not ordinarily occur? Is it appropriate to test a new medication by giving a placebo treatment that is known to not work to participants in a placebo group who are unwell and genuinely need the new medication, and are therefore intentionally allowed to remain unwell? Should all participants be fully informed of the purpose of the research before they agree to participate? Should participants have the right to withdraw from an experiment at any time without giving a reason for wanting to do so? 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 judge 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 others (NHMRC, 2007).

All societies and cultures 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 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 or 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*. This document is simply referred to as the National Statement. The National Statement is regularly updated, with the May 2015 version incorporating most updates.

studyon

Figure 1.92 Ethical standards for human research ensure all participants are given the respect and protection due to them, irrespective of who they are.

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.

UNCORRECTED PAGE PROOFS

1. **Research merit and integrity**: Research that has *merit* is worthwhile and conducted in an appropriate way 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.

   Integrity generally refers to honesty and 'doing the right thing'. 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 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**: Justice involves fairness. Research that is *just* has a concern for the use of fair procedures and fair distribution of the costs and benefits of the investigation. 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**: This involves recognition that each human being has value in himself or herself. *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 people with an intellectual disability who depend on others.

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

eBook*plus*

Weblinks
- National Statement
- APS Code of Ethics

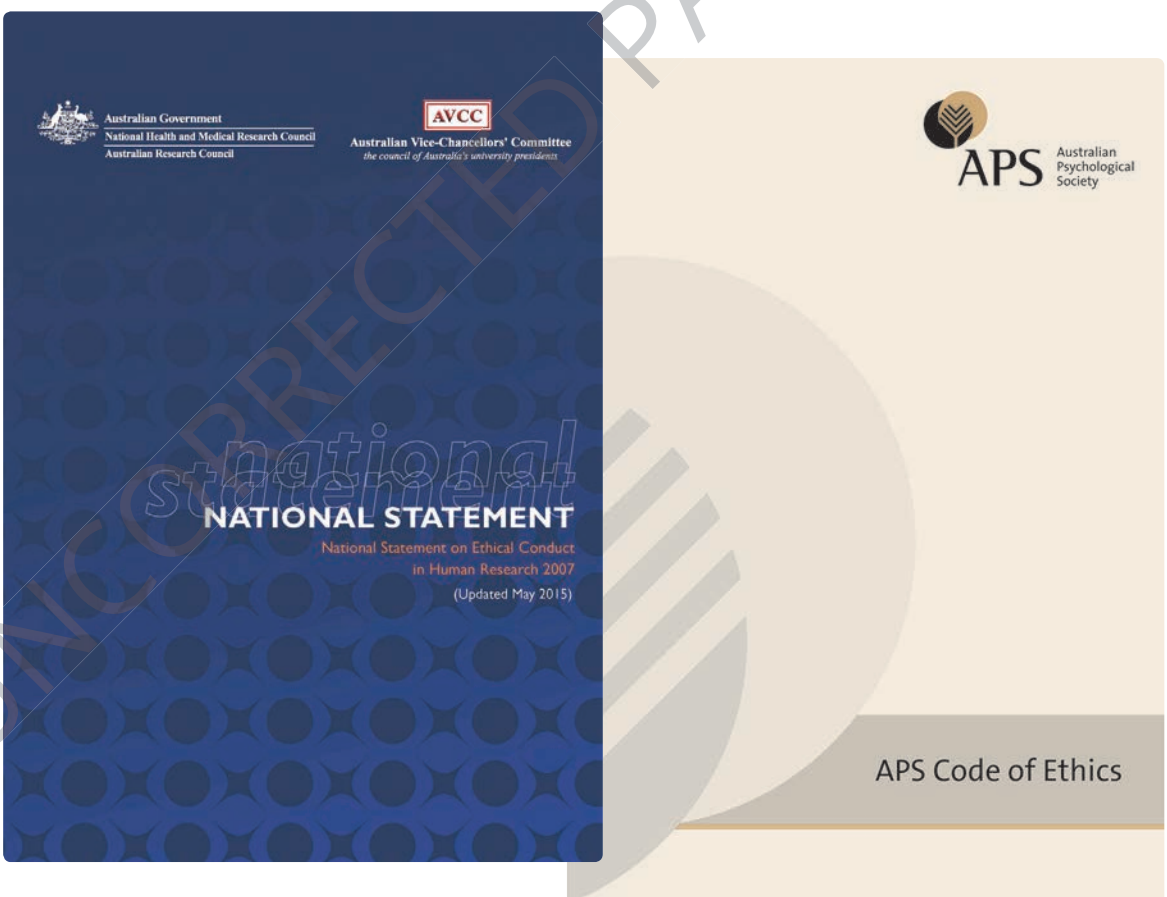


Figure 1.93 All psychological research with human participants must follow the guidelines in the *National Statement*. The guidelines in the *APS Code of Ethics* are consistent with those in the *National Statement*.

UNCORRECTED PAGE PROOFS

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 proposal 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 a HREC 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.

UNCORRECTED PAGE PROOFS

Figure 1.94 Human research ethics committees are established to review research proposals for approval purposes, and then monitor the conduct of the research to ensure all relevant ethical guidelines are adopted and followed. Review meetings are usually informal.

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, 2017).

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
- *Data quality* — ensure information is accurate, complete and up to date
- *Data security* — protect the information (e.g. from loss or unauthorised access) and destroy or permanently de-identify personal information if no longer needed.

eBook*plus*

Weblink
Australian Privacy Principles

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 ethical 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 (kept safely) 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 (or anyone else) 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 stated previously, 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 opt out of 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.

UNCORRECTED PAGE PROOFS

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 (such as intellectual or mental health status). There should be an opportunity for prospective participants to ask questions about the research.

It is essential that participants have the competence to give informed consent. A wide variety of symptoms, diseases, injuries and other conditions can affect a person's ability to understand information and researchers must take this into account when seeking informed consent and providing documents or information relevant to informed consent.

For participants who are legally unable to give informed consent (e.g. children and people impaired by 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 (e.g. 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 1.96 on the next page. Two copies are made so that one can be kept by the researcher and one by the participant.

There are specific ethical issues associated with informed consent for mental health research. These are examined in Chapter 11.

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. Its use also means that the relationship between the researcher and participant is not open and honest. However, deception is considered 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 if deception was used, 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 each 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. In extreme cases, participant wellbeing may be monitored after the research; for example, participants may receive questionnaires, be asked to complete diaries, and/or have follow-up meetings with a researcher.

Figure 1.95 The researcher is responsible for the proper ethical conduct of their investigation.

UNCORRECTED PAGE PROOFS

SAMPLE ONLY

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, ..., give my informed 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 I have any questions, comments or complaints to make on this research, I can 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: .....

UNCORRECTED PAGE PROOFS

Figure 1.96 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.

eBook*plus*

Weblinks
Templates for psychology research consent forms

BOX 1.29 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/e72syn.htm
- the National Privacy Principles in the *Privacy Amendment (Private Sector) Act 2000* (Cwlth), www.privacy.gov.au/
- 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:

- the *Prevention of Cruelty to Animals Act 1986* and its Regulations 2008
- the *Australian Code of Practice for the Care and Use of Animals for Scientific Purposes 2014* (8th edition), www.nhmrc.gov.au/guidelines/publications/ea28

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*
- *Electrical Safety Act 1998*

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.

Source: © VCAA, *VCE Psychology Study Design* (June 2017 update), pp. 8–9.

UNCORRECTED PAGE PROOFS

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], 2017a).

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; the role of the brain in behaviour; 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, 2017a; Bennett, 2012; Puente, 2017).

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

Figure 1.97 Animals are used in a wide variety of psychological research studies. Research with animals is governed by NHMRC ethical guidelines.

- 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.
- 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 many species of 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 researcher 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.

UNCORRECTED PAGE PROOFS

- 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 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 standards and 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 *Australian Code for the Care and Use of Animals for Scientific Purposes 2013 (8th ed)*.

The purpose of the Code is 'to promote the ethical, humane and responsible care and use of animals for scientific purposes'. An obligation to respect animals is central in the Code. According to the Code (p.1), 'This obligation brings with it a responsibility to ensure that the care and use of animals for scientific purposes is ethically acceptable, balancing whether the potential effects on the wellbeing of the animals involved is justified by the potential benefits to humans, animals or the environment. The use of animals for scientific purposes must have scientific or educational merit; must aim to benefit humans, animals or the environment; and must be conducted with integrity. When animals are used, the number of animals involved must be minimised, the wellbeing of the animals must be supported, and harm, including pain and distress, in those animals must be avoided or minimised.'

Figure 1.98 Psychologists must ensure that research animals are well cared for, humanely treated and experience minimal pain and suffering.

eBook*plus*

Weblink
NHMRC Code for care and use of animals

Figure 1.99 Although many important benefits have been achieved through animal research and that there are ethical obligations for their care and use, many people remain vigorously opposed to all animal research. This Animal Liberation Victoria activist covered herself in fake blood and strapped herself to a giant vivisection board as part of three days of protest at an international animal research conference held in Melbourne. Undoubtedly, the use of animals in research has been, and continues to be, a highly charged, controversial issue.

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.37

Review questions

1. Explain the meaning of ethics in relation to research.
2. What is the main purpose of ethical standards for research with human participants?
3. (a) Name and describe the four values that should be reflected in all human research.
   (b) Give an example of a research practice that would be:
   (i) consistent with each value
   (ii) inconsistent with each value.
4. What are three essential informed consent procedures?
5. What is the ethical responsibility of a researcher who does not want to fully inform prospective participants of the true purpose of an experiment because it may influence how they respond during the study?
6. If a participant became very upset during an experiment, what is the responsibility of the researcher?
7. Explain what privacy is and the ethical relevance of the Australian Privacy Principles.
8. (a) What is an ethics committee?
   (b) List three of its roles or responsibilities.
9. Give two advantages and limitations of animal use in psychological research.
10. List three ethical guidelines to be followed for research with animals.
11. Construct a table that summarises the ethical standards and guidelines for conducting human research. Use the table to develop an ethics checklist that could be used for your practical activities and research investigation in Units 3 and 4.

LEARNING ACTIVITY 1.38

eBook*plus*
Word copy of table

Applying ethical values

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

	Statement	Ethical value
1.	The process of recruiting participants is fair.	
2.	The researcher ensures all members of the research team are properly qualified to undertake their respective responsibilities.	
3.	The participant is given a chance to have their questions about the research answered.	
4.	The researcher does not unnecessarily invade privacy.	
5.	The researcher is confident that what will be learnt from their study justifies the risks of discomfort to participants.	
6.	The researcher is based on a thorough study of the current literature, as well as previous studies.	
7.	Where there are no likely benefits to participants, the risk to participants should be lower than would be ethically acceptable where there are such likely benefits.	
8.	There is no exploitation of participants in the conduct of research.	
9.	There is due regard for the customs and cultural heritage of all individual participants.	
10.	Research outcomes are made accessible to participants in a way that is timely and clear.	

UNCORRECTED PAGE PROOFS

LEARNING ACTIVITY 1.39

Identifying ethical issues

Consider the following fictitious examples of research studies that may breach one or more ethical standards or guidelines and identify the ethical issue(s) raised, if any, in each example.

Study 1
A psychology lecturer at a university was studying techniques for reducing fear of spiders. He asked a research assistant to telephone students in the first-year psychology course he was teaching to determine their willingness to participate. The researcher was unaware that the assistant told participants that they had to participate.

Study 2
A researcher was interested in factors influencing cheating. She gave participants an exam, then collected and photocopied their answers. The participants were not informed about the photocopying. The answers were returned unmarked and the participants were given the opportunity to cheat while marking their own papers. The answers were collected again and compared with the photocopies.

Study 3
Another researcher investigated cheating by concealing themself and three colleagues in a projection booth in an auditorium during an exam. From this vantage point, high above all students, the researchers used binoculars to observe cheating behaviours of students in different quadrants of the room. Each observer used a checklist to record inappropriate head movements, exam paper switching, note checking, note passing and other suspicious exam behaviours within the quadrants.

Study 4
An experiment was conducted to assess driver reaction to a stressful situation. Each participant was asked to drive a car past a construction site. The researcher rigged a human-looking dummy in such a way that it would be propelled in front of the car, making it impossible for the participant to avoid hitting it. The participants reacted as expected. When they learned that the situation was faked, they informed the researcher of their displeasure. Despite their complaints, the researcher continued testing further participants (adapted from Wood, 1981).

Study 5
A VCE Psychology student was required to undertake a research investigation to satisfy the course requirements. The student researcher replicated an experiment on learning that involved classical conditioning of an eye-blink response using two preschool children and two adults as participants. The student researcher thought that the adult participants' knowledge of the conditioning procedure would affect the results in an unwanted way and decided not to seek their informed consent. The student researcher also based their decision on the belief that the conditioning procedure was physically and psychologically harmless. The student researcher did, however, obtain informed written consent from both parents of each child.

Study 6
A researcher conducted an observational study to investigate behaviour in public rest rooms. This research method was expected to obtain more valid and reliable data than could be accessed through a self-report measure. A team of male and female researchers concealed themselves in vacant toilet stalls of the respective restrooms and observed behaviours of men and women (adults only), such as flushing vs non-flushing, hand washing vs hand drying, mirror checking, clothing adjustments, littering and graffiti writing.

LEARNING ACTIVITY 1.40

Reflection

Comment on whether the potential benefits of using non-human animals in psychological research justify their use. Does this apply to all types of research or only those studies for which there is a risk of harm? Only certain animal species? When should and should not animals be used in psychological research, if at all?

REPORTING CONVENTIONS

The final and very important stage in the research process involves preparation of a detailed written report on the research study and its findings. This is done for two main 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.

Generally a written report provides 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 research (including the results) and, if required, to replicate the research
- reporting conventions are used.

Reporting conventions are well-established and widely recognised standards for how a report is written and presented. Reporting conventions

UNCORRECTED PAGE PROOFS

determine aspects of the report such as writing style, structure of the report, headings, presentation of tables and graphs, and formats for referencing.

For example, the writing style, or language, used in a psychological research report is like that of all scientific reports. The language is formal, clear and concise. It is 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, commonly called the 'APA manual' and its conventions are commonly referred to as 'APA format'.

The APA format is widely recognised and used by psychological researchers throughout the world to guide their preparation and presentation of written and poster reports. These conventions are also used by psychology students for writing research reports conducted as part of their university studies. VCAA does not mandate specific reporting conventions for VCE Psychology.

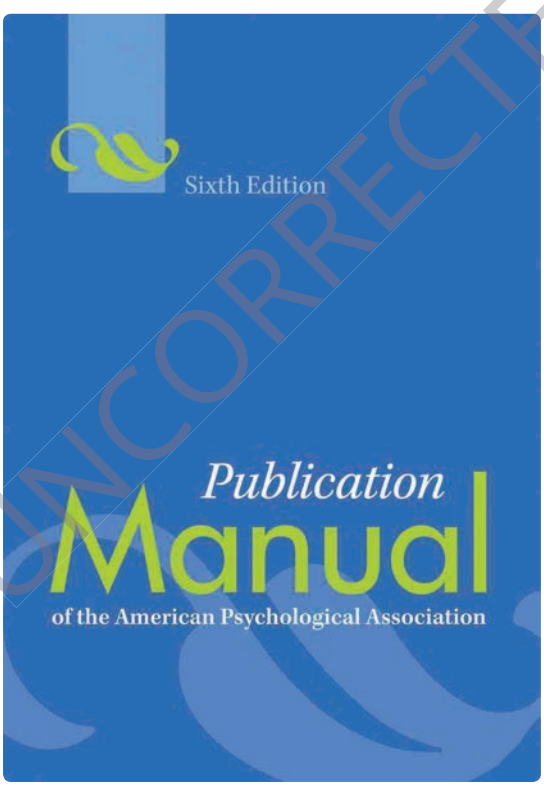


Figure 1.100 The cover of the APA manual

The following guidelines for writing a research report and referencing are based on the APA manual. Guidelines for preparing a poster report are then outlined. The poster is less detailed and is commonly used to provide a summary of the key features of a research study for display and discussion at conferences or meetings with other psychologists likely to be interested in the research and its findings.

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 a 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.

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).

Considering the nature of its contents, 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.

UNCORRECTED PAGE PROOFS

The introduction is often written in a way that leads the reader to a statement of the aim of the research and the hypothesis that 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, usually refers to the independent and dependent variables (but not necessarily operationalised).

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, 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.

Procedure

This is a detailed description of the research method and design. This information should be presented in a way that another researcher could conduct the same study just by reading your description.

Relevant information may include the roles of the researcher, how participants were recruited, whether they were placed into groups or whether this was achieved by using random allocation, what participants were asked to do, the setting and duration 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 on 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 research hypothesis is supported or refuted on the basis of the results obtained. If the results do not support the hypothesis, then an explanation is given.

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

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

Finally, the practical applications of the findings to the real world are considered. 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 1.30 on the next page.

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 at a meeting or conference. It may also be used for reporting research conducted by students in psychology courses.

UNCORRECTED PAGE PROOFS

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, shown below. More comprehensive information about the poster is available in the VCE Advice for teachers: Psychology digital document at the VCAA website.

In Unit 4, there is a SAC requiring 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. Scientific poster templates available on the internet may be used provided that the mandated poster sections (Title, Introduction, Methodology, Results, Discussion, Conclusion, References and Acknowledgments) are included.

Section	Content and activities
Title	Question under investigation is the title
Introduction	Explanation or reason for undertaking the investigation, including a clear aim, a hypothesis and/or prediction and relevant background psychological concepts
Methodology	Summary that outlines the methodology used in the investigation and is authenticated by logbook entries
	Identification and management of relevant risks, including the relevant health, safety and ethical guidelines followed in the investigation
Results	Presentation of collected data/evidence in appropriate format to illustrate trends, patterns and/or relationships
Discussion	Analysis and evaluation of primary data
	Identification of outliers and their subsequent treatment
	Identification of limitations in data and methods, and suggested improvements
	Linking of results to relevant psychological concepts
Conclusion	Conclusion that provides a response to the question
References and acknowledgements	Referencing and acknowledgement of all quotations and sourced content as they appear in the poster

Source: © VCAA, *VCE Psychology Study Design* (June 2017 update), p.13.

UNCORRECTED PAGE PROOFS

BOX 1.30 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'. 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 essay, 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 Smith (2011), participants were required…

Two authors: A similar result was reported by Voulos and Jones (2014), who found that…

Three to five authors: List all authors (separate the names with commas) and publication date e.g. Black, White, and Grey (2015) studied the effects of…

Six or more authors: First author + et al. + date e.g. According to Hemming et al. (2016), the frontal lobes…

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.

(continued)

(continued from previous page)

At the end of a sentence

One author: Participants who are rewarded are more likely to repeat the response for which they are rewarded (Canasta, 2010).

Two authors: When individuals are anxious they tend to seek the company of other people (James & Mahir, 2012).

Three to five authors: The human stress response has physiological and psychological components (Stavros, Wilson & Pink, 2015).

Six or more authors: . . . therefore, the amygdala has a crucial role in emotion (Hemming et al., 2016).

Citing a reference within another source

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 Jackson, 2012, p. 142) was concerned about . . .

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 Schlink (2010) explained the observation in terms of "the interaction between short-term memory and long-term memory and . . . decay of memory traces (p. 92)".

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 formats such as those in the following chart. There are some minor variations to suit VCE Psychology. Generally, the goal is to provide sufficient information about a source to enable it to be efficiently located.

Type	Format	Example
Book	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.) 2. *Year of publication*. (Enclosed in brackets, followed by a full stop) 3. *Title of book*. (Italicised and followed by a full stop, but no full stop if there is an edition number.) 4. *Edition*. (If a second or subsequent edition, 'edition' is abbreviated, enclosed in brackets and followed by a full stop.) 5. *City of publication* (and state if city is not well known, followed by a colon) 6. *Name of publisher*. (Followed by a full stop. Omit unnecessary terms such as Publishers, Co., and Inc. Retain the words Books or Press.)	**Book with one author:** Baddeley, M. (2011). *Memory*. London: Psychology Press. **Book with two or more authors:** Red, J., & White, N. (2016). *Psychology of life*. New York: W.M. Freeman. **Book with another edition:** Grivas, J., & Letch, N. (2018). *Psychology: VCE Units 3 and 4* (7th ed.). Milton, Qld: John Wiley & Sons. **Online book:** If available only online, the URL (full http address) takes the place of the publisher location and name but it is not underlined, active or followed by a full stop e.g. Grivas, J., & Letch, N. (2018). *Psychology: VCE Units 3 and 4* (7th ed.). Retrieved from http://content.jacplus.com.au/Psychology VCE Units 3 & 4
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 on 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 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.) 7. *Name of publisher*. (Followed by a full stop. Omit terms such as Publishers, Co., and Inc. Retain the words Books or Press.)	**One author:** Purpley, M. (2016). Neuroplasticity throughout the lifespan. In N.O. Dixon & S. Lim (Eds.), *Biological basis of behavior* (pp. 103–115). New York: Academic Press. **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. Purpley, M. (2016). Neuroplasticity throughout the lifespan. In N.O. Dixon & S. Lim (Eds.), *Biological basis of behavior* (pp. 103–115). Retrieved from http://store.elsevier.com/Academic-Press/IMP_5/

UNCORRECTED PAGE PROOFS

Type	Format	Example
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.) 6. *Page numbers*. (Followed by a full stop)	**One author:** Mandalis, D.S. (2016). Stress and athletic performance. *Australian Journal of Psychology*, *10*, 334–345. **Two or more authors:** Blackmore, A.B., & Chan, I. (2012). Memory decline and aging. *Developmental Psychology, 19* (2), 121–128.
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 publisher*. (Followed by a full stop.)	**Print copy:** SANE Australia. (2014). *Dealing with a traumatic event when you have a mental illness* (Fact sheet, 37). South Melbourne, Vic.: Author. **Online:** beyondblue. (2014).*Specific phobias* (Fact sheet). Retrieved March 23, 2017, from http://resources.beyondblue.org.au/prism/file?token=BL/0508
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) 4. *Page numbers*. (Followed by a full stop)	**If you know the author:** Paggio, I. (2016, July 20). Improve your memory. *The Sunday Age*, p. 11. **If you do not know the author:** Video games promote violence. (2017, January 4). *Herald Sun*, p. 6. **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. Black, M. (2016, December 12). Study shows practice is beneficial. *The Age*. Retrieved from http://www.theage.com.au
Dictionary or other text reference with no author	1. *Title* (Italicised) 2. *Edition*. (If a second or subsequent edition, 'edition' is abbreviated (e.g. 2nd ed.), enclosed in brackets and followed by a full stop.) 3. *Year of publication* (Enclosed in brackets, followed by a full stop.) 4. *City of publication* (and state if city is not well known, followed by a colon) 5. *Name of publisher*. (Followed by a full stop. Omit unnecessary terms such as Publishers, Co., and Inc. Retain the words Books or Press.)	**In the reference list:** *APA dictionary of psychology* (2nd ed.). (2015). Washington, DC: American Psychology Association. **Within text citation:** The APA Dictionary of Psychology (2015, p. 65) defines . . . as . . . or . . . can be defined as . . . (APA Dictionary of Psychology (2015, p. 65). *Note capitalisation of all major words for a within text citation*.
Internet (including YouTube and Pod casts)	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. Retrieved from 5. *When retrieved if date of publication is not available or if the material is not subject to frequent change*. (Year, month, date followed by a comma and the word from) *Otherwise, simply state* Retrieved from 6. *URL*. (Not underlined, active or followed by a full stop)	**Site when publication date is known:** mindhealthconnect. (2015). *Phobias*. Retrieved from http://www.mindhealthconnect **When publication date is not known:** Smartie, J. (n.d.). *How I felt when I saw a spider*. Retrieved 2017, January 12, from http:// XXXXXXXXX **You Tube:** Mount Sinai Hospital (2013, April 18). *Deep brain stimulation surgery to treat Parkinson's Disease at Mount Sinai Hospital* (Video file). Retrieved from https://www.youtube.com/watch?v=MEBdXbZ5CDM **Pod cast:** Hammond, C. (2010, August 11). *Case study: HM — The man who couldn't remember* (Audio podcast). Retrieved from http://www.bbc.co.uk/programmes/b00t6zqv

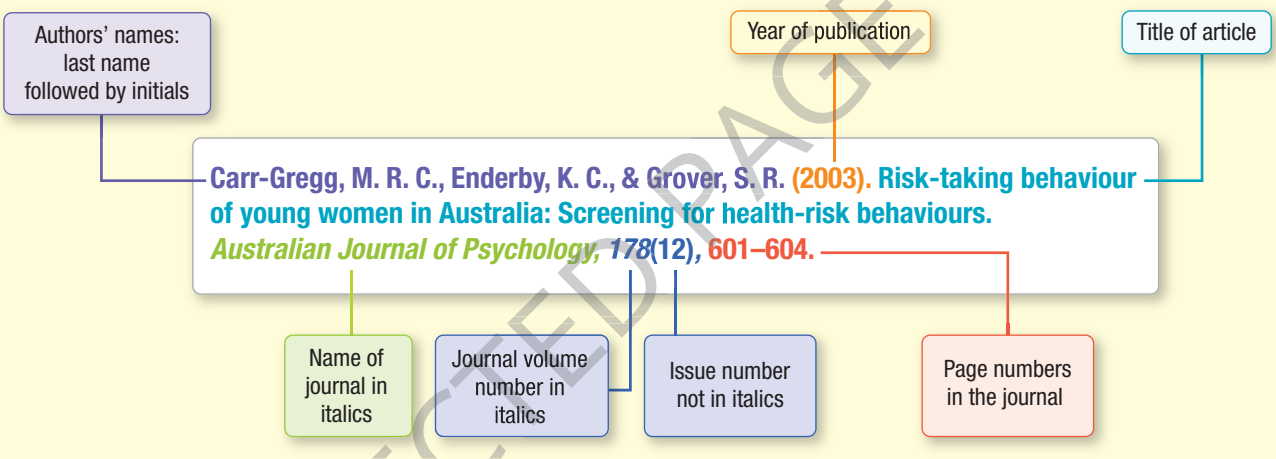
(continued)

UNCORRECTED PAGE PROOFS

(continued from previous page)

Type	Format	Example
Motion picture (movie), TV, 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 motion picture's country of origin, where it was primarily made and the name of the production company)	**Movie:** Mamin, S.A. (Producer), & Samir, M. (Director). (2010). *Old and young* (Motion picture). Australia: Billabong Productions. **TV program:** Smith, J. (Executive producer). (2016, December 12). *The 7pm project* (Television broadcast). Melbourne,Vic: Network TEN. **DVD:** Hays, P.A. (2015). *Culturally responsive cognitive behavior therapy over time* (DVD). Available from http://www.apa.org/pubs/videos/4310945.aspx
Personal communication or interview e.g. class notes and data, letters, personal interviews, telephone conversations, email	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.	J. Smith (interview, March 12, 2017).

Example 1. Journal article (print copy)



Example 2. Website

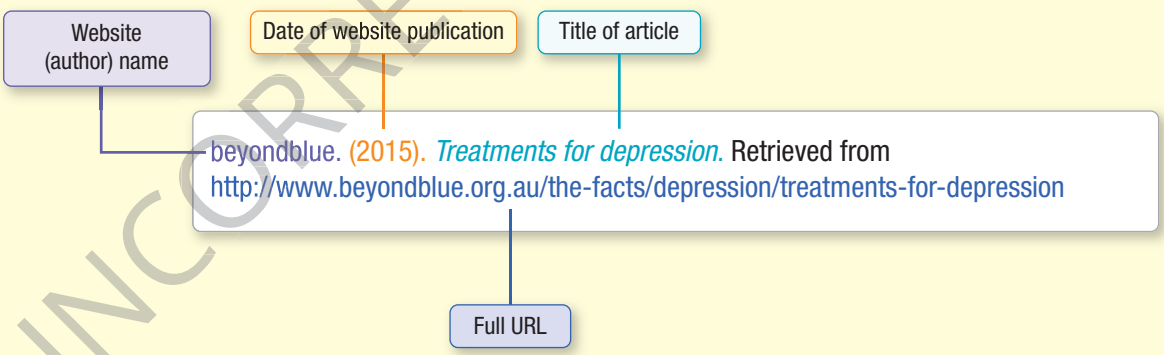


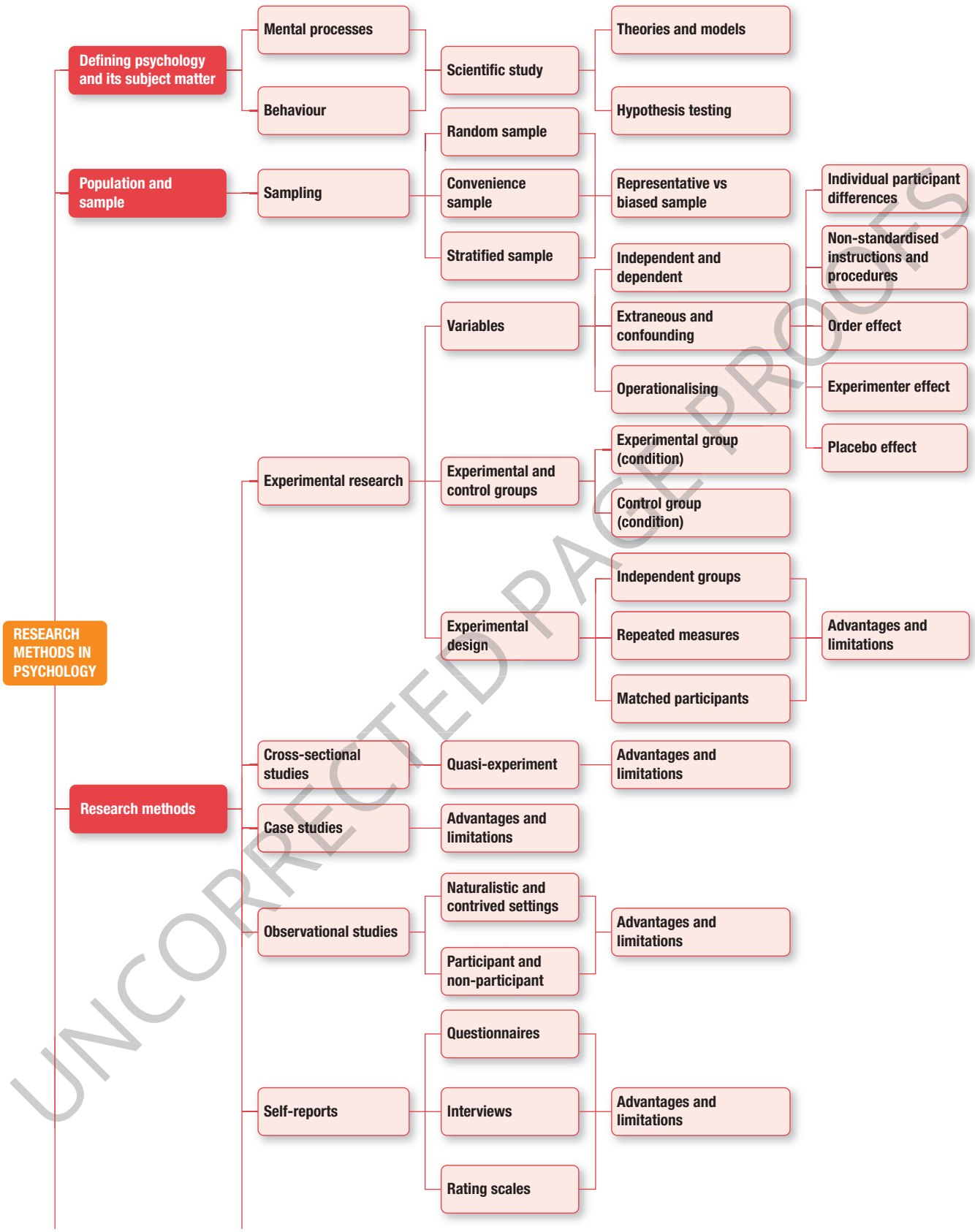
Figure 1.101 Examples of conventions used by psychological researchers for referencing. Note the use of commas, full stops and italics. Colour is used here for illustration purposes and is not a convention. Note also that researchers now include digital object identifier numbers for references that are available electronically. These have been excluded from references at the back of this text.

eGuide*plus*

Weblink
Digital Object Identifier (DOI) number

UNCORRECTED PAGE PROOFS

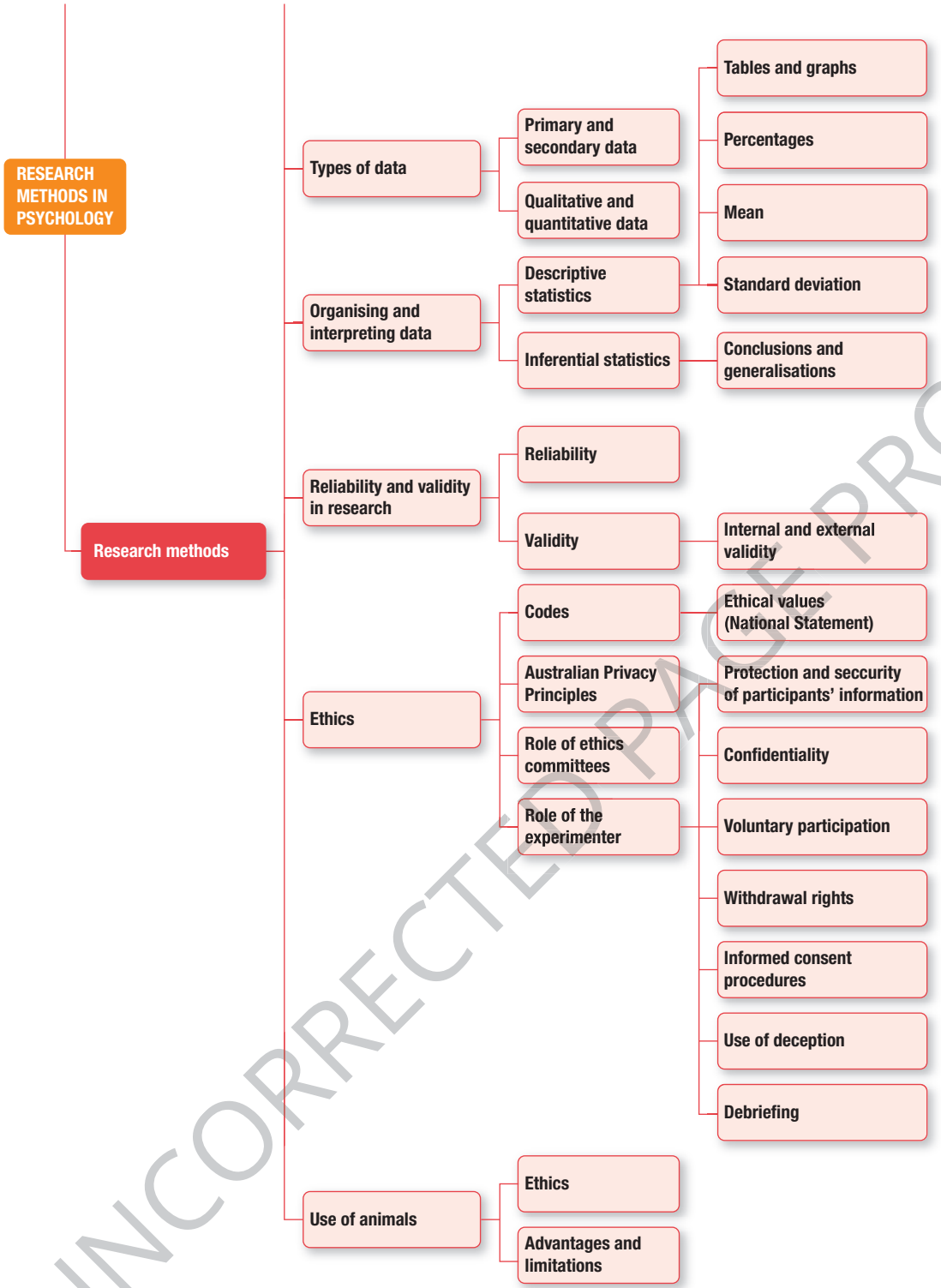
CHAPTER SUMMARY



(continued)

UNCORRECTED PAGE PROOFS

(continued from previous page)



UNCORRECTED PAGE PROOFS

KEY TERMS

beneficence (ethics) p. 111
biased sample p. 39
case study p. 69
chance p. 103
cohort effect p. 67
conclusion p. 104
confidentiality (ethics) p. 113
confounding variable p. 24
control condition p. 22
control group p. 22
convenience sampling p. 43
counterbalancing p. 50
cross-sectional study p. 66
debriefing (ethics) p. 114
deception (in research) p. 114
dependent variable (DV) p. 19
descriptive statistic p. 94
double blind p. 51
ethics p. 110
ethics committee p. 112
experiment p. 17
experimental condition p. 22
experimental group p. 22
experimenter effect p. 32
external validity p. 109
extraneous variable p. 23
generalisation p. 104
hypothesis (in research) p. 15
independent groups experiment p. 55
independent variable (IV) p. 18
individual participant differences p. 28
inferential statistic p. 94
informed consent p. 113
integrity (ethics) p. 111
internal validity p. 108
interview p. 81
justice (ethics) p. 111
matched participants experiment p. 57
mean p. 99
measure of central tendency p. 99
measure of variation p. 100
mental process p. 5
merit (ethics) p. 111
model p. 16
naturalistic observation p. 74
non-participant observation p. 75
non-standardised p. 30
observational study p. 72
operationalising a variable p. 20
order effect p. 31
participant observation p. 75
participant variable p. 28
placebo p. 35
placebo effect p. 35
population p. 14
primary data p. 89
privacy (ethics) p. 113
qualitative data p. 90
quantitative data p. 90
questionnaire p. 81
random allocation p. 47
random sampling p. 40
rating scale p. 83
reliability p. 106
repeatability p. 107
repeated measures experiment p. 56
replication p. 10
reporting convention p. 120
representative sample p. 39
reproducibility p. 107
research hypothesis p. 15
research merit and integrity (ethics) p. 111
respect for human beings (ethics) p. 111
sample p. 14
sampling p. 39
secondary data p. 89
self-report p. 81
single blind p. 51
standard deviation p. 101
standardised instructions p. 53
standardised procedures p. 53
stratified sampling p. 42
theory p. 16
validity p. 106
variable p. 17
withdrawal right (ethics) p. 113

UNCORRECTED PAGE PROOFS

LEARNING CHECKLIST

Complete the self-assessment checklist below, using ticks and crosses to indicate your understanding of this chapter's key knowledge (a) before and (b) after you attempt the chapter test on pages 133–39. Use the 'Comments' column to add notes about your understanding.

eBook*plus*
Word copy of checklist

Key knowledge I need to know about research methods	Self-assessment of key knowledge I understand *before* chapter test	Self-assessment of key knowledge I need to revisit *after* chapter test	Comments
• Sample and population			
– sample vs population			
– representative vs biased sample			
– random sample			
– convenience sample			
– stratified sample			
• research hypothesis			
– key features			
– how to formulate			
• Experimental research			
– independent and dependent variables			
– operationalising variables			
– experimental and control groups (conditions)			
– random allocation			
○ procedure			
○ purpose			
– extraneous and confounding variables			
○ individual participant differences			
○ non-standardised instructions and procedures			
○ order effects			
○ experimenter effect			
○ placebo effect			
– Ways of minimising extraneous and confounding variables			
○ participant selection and allocation			
○ counterbalancing			
○ single and double blind procedures			
○ placebos			
○ standardised instructions and procedures			
○ use of an appropriate experimental research design			

UNCORRECTED PAGE PROOFS

Key knowledge I need to know about research methods	Self-assessment of key knowledge I understand *before* chapter test	Self-assessment of key knowledge I need to revisit *after* chapter test	Comments
• Experimental designs			
– key features			
– independent groups			
– repeated measures			
– matched participants			
– advantages and limitations			
• Cross-sectional studies			
– key features			
– advantages and limitations			
• Case studies			
– key features			
– advantages and limitations			
• Observational studies			
– key features			
– natural and contrived settings			
– participant and non-participant			
– advantages and limitations			
• Self-reports			
– key features			
– questionnaires			
– interviews			
– rating scales			
– advantages and limitations			
Types of data			
• Primary and secondary data			
• Qualitative and quantitative data			
Organising, presenting and interpreting data			
• Descriptive statistics			
– tables			
– graphs (especially graphing IV and DV)			
– percentages			
– mean			
– standard deviation			
• Inferential statistics			
• Conclusions and generalisations			

(continued)

UNCORRECTED PAGE PROOFS

Key knowledge I need to know about research methods	Self-assessment of key knowledge I understand *before* chapter test	Self-assessment of key knowledge I need to revisit *after* chapter test	Comments
• Reliability and validity in research			
– reliability			
– validity			
○ internal validity			
○ external validity			
• **Ethics in psychological research and reporting**			
– National Statement values			
– Australian Privacy Principles			
– Role of ethics committees			
– Role of the experimenter			
– Ethical standards			
○ protection and security of participants' information			
○ confidentiality			
○ voluntary participation			
○ withdrawal rights			
○ informed consent procedures			
○ use of deception			
○ debriefing			
• Use of animals in psychological research			
○ advantages and limitations			

studyon

Unit 3 > Area of study 1 > Topics 1–5

Concept screens and practice questions

UNCORRECTED PAGE PROOFS

CHAPTER 1 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

Which of the following procedures is an essential feature of any type of psychological experiment?

A. random allocation
B. independent groups
C. double blind
D. counterbalancing

Question 2

A random sample of VCE students in a school could be achieved by selecting

A. every tenth student walking out of a VCE assembly.
B. all students who walk to school.
C. all students who are enrolled in three or more science studies.
D. all students whose VCE candidate number ends with an even number.

Question 3

Operationalising the variables for an experiment involves

A. strictly controlling all variables that can impact on the dependent variable.
B. defining and explaining how the independent and dependent variables will be measured.
C. deciding on the importance of all the experimental variables.
D. identifying all potential extraneous and confounding variables.

Question 4

A structured observational study will involve

A. naturalistic observations.
B. observations in a contrived setting.
C. the non-participant procedure.
D. a pre-prepared system to guide and record observations.

Question 5

In an experiment, the group that receives the treatment is called the _____ group, whereas the group that does not is called the _____ group.

A. independent; dependent
B. control; experimental
C. dependent; independent
D. experimental; control

Question 6

Random allocation and random sampling

A. are avoided by researchers as they are haphazard procedures.
B. differ in that random allocation is used to place participants in groups and random sampling is used to select participants for the research.
C. differ in that random sampling is used to place participants in groups and random allocation is used to select participants for the research.
D. are both used to select participants for an experiment.

Question 7

A matched participants experiment involves

A. allocating each member of a matched group to a different experimental condition.
B. matching participants to the experimental condition to which they will be allocated.
C. allocating each participant to the same experimental conditions.
D. randomly selecting participants, then randomly allocating to a different experimental condition.

Question 8

A researcher selects participants by randomly sampling different groups from a target population. The researcher believes that the sex and religious beliefs of participants will be influential on the results, so the researcher ensures these characteristics are proportionally represented in the sample.

This type of sampling procedure is best described as

A. biased.
B. random.
C. stratified.
D. stratified random.

Question 9

Ethical standards in psychological research are intended to ensure that

A. participants are responsible for the research.
B. participants can comment on the results whenever they want to.
C. the rights and wellbeing of the researcher are safeguarded.
D. the rights and wellbeing of participants are not compromised in any way.

UNCORRECTED PAGE PROOFS

Question 10

Heart rate can be an independent variable in an experiment because

A. everyone has a heart.
B. heart rate can be measured.
C. a researcher can manipulate heart rate.
D. participants can manipulate their heart rate.

Question 11

For a normal distribution, about _____ of data values lie within one standard deviation of the mean and about _____ of data values lie within two standard deviations of the mean.

A. 13%; 34%.
B. 34%; 13%.
C. 34%; 68%
D. 68%; 95%

Question 12

A researcher gives vitamin C to one group of research participants and a placebo to another group to measure the effect of vitamin C on the common cold.

The frequency of colds is

A. the independent variable.
B. the dependent variable.
C. an extraneous variable.
D. a confounding variable.

Question 13

An experiment was conducted to assess the effectiveness of a new technique for learning Greek words. One group used the learning technique and another group did not. Both groups were then given the same test of recall of Greek words.

The results showed that the group using the learning technique recalled more Greek words than did the group who did not use the learning technique. In this experiment, _____ is the independent variable, whereas _____ is the dependent variable.

A. number of Greek words correctly recalled; using the learning technique
B. using the learning technique; number of Greek words correctly recalled
C. number of Greek words learned; number of Greek words correctly recalled
D. number of Greek words correctly recalled; number of Greek words learned

Question 14

A researcher intentionally arranged the order in which the conditions of a repeated measures experiment were experienced. This was done to control a practice effect that was expected to occur.

This procedure is most commonly referred to as

A. manipulation.
B. control.
C. counterbalancing.
D. experimenter bias.

Question 15

In a repeated measures experiment, each participant is exposed to

A. all conditions of the experiment.
B. the independent variable only.
C. the independent variable repeatedly.
D. the dependent variable repeatedly.

Question 16

Generalising from of the results of research involves

A. determining the reliability and validity of the results.
B. establishing a cause–effect relationship between the independent and dependent variables.
C. applying the findings obtained from the sample to its population.
D. drawing a conclusion about whether the results support or do not support the hypothesis.

Question 17

Which of the following is an example of a self-report?

A. notes kept by a person with a phobia of research interest
B. the researcher's raw data collected for their study
C. the researcher's formal report on their study
D. the researcher's records in an observation checklist

Question 18

Which of the following researcher behaviours would be considered unethical?

A. informing participants about the results of the experiment
B. preventing a participant from opting out midway through the experiment
C. checking up on the age of a participant when there is doubt that the participant may not be old enough to give informed consent
D. publishing the results of the experiment without obtaining informed consent from the participants

Question 19

A researcher investigated the effect of a new drug designed to improve memory. A placebo was used to control

A. participant expectations.
B. experimenter expectations.
C. experimental error.
D. ethical issues.

Question 20

Before conducting an experiment, a researcher identified all extraneous variables with the potential to affect the dependent variable, then refined the experiment's design to control the influence of these variables. The researcher did this to help ensure that

A. the independent variable could be manipulated.
B. the dependent variable could be measured.
C. the experiment would be reliable.
D. there would be no confounding variables.

UNCORRECTED PAGE PROOFS

Question 21

The standard research procedure for ensuring control over participant variables and experimenter effects is the use of

A. operationalised variables.
B. double blind.
C. standardised variables.
D. single blind.

Question 22

A researcher uses test scores as a measure of their dependent variable. The test scores are best described as _____ data.

A. primary
B. secondary and quantitative
C. primary and quantitative
D. primary and qualitative

Question 23

A general description or explanation of a set of observations or findings about behaviour and/or mental processes which seem to be related is best described as a

A. research finding.
B. research hypothesis.
C. generalisation.
D. theory or model.

Question 24

A common feature of cross-sectional studies is

A. random allocation to groups.
B. use of an independent groups design.
C. one time only study of age-related differences.
D. long-term study of age-related differences.

Question 25

If research procedures are standardised, then

A. the results will be valid.
B. the research will be conducted ethically.
C. all participants will understand what the experiment is requiring of them.
D. the procedures used in a specific condition will be the same for all participants.

UNCORRECTED PAGE PROOFS

SECTION B

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

Question 1 (1 mark)

What is the most essential feature of a research hypothesis?

Question 2 (1 mark)

What is the main purpose of using a control group in an experiment?

Question 3 (6 marks)

(a) Distinguish between an extraneous variable and a confounding variable. 2 marks

(b) Explain why confounding is evident in the following data. 2 marks

Participants		Group 1 (IV present)	Group 2 (IV absent)
Number		50	50
Sex	Males	40	10
	Females	10	40
Mean age (years)		35	70
Mean score on an intelligence test (IQ)		100	130

(c) Explain why a confounding variable is an unwanted variable in an experiment. 2 marks

Question 4 (2 marks)

How does random allocation in an experiment minimise the likelihood of individual participant differences becoming a confounding variable?

UNCORRECTED PAGE PROOFS

Question 5 (2 marks)

Distinguish between internal and external validity in relation to research.

Question 6 (5 marks)

Read the following extract from a student's report on a research investigation and answer the following questions.

> *To ensure randomisation, questionnaires were handed out at many different places and at different times throughout the day. Moreover, by choosing to sample a relatively large population, we were able to ensure that the average results of many individual results would produce a stable result.*

(a) Identify the sampling procedure. 1 mark

(b) Explain whether or not the researcher actually 'ensured randomisation'. 2 marks

(c) What technical term do psychologists use to refer to 'stable results'? 1 mark

(d) Explain whether 'stable results' were actually achieved. 1 mark

Question 7 (4 marks)

Suppose that a researcher wants to test the hypothesis that participating in psychotherapy will cause a decrease in reported anxiety.

(a) Outline the design of a cross-sectional study that could be used to test this hypothesis. 2 marks

(b) Explain the problem a cohort effect creates when interpreting the results. 2 marks

Question 8 (4 marks)

(a) Give two reasons to explain why variables of research interest are operationalised. 2 marks

UNCORRECTED PAGE PROOFS

(b) Give an example of how each of the following variables of research interest could be operationalised for an experiment.

(i) student satisfaction as an IV _____ 1 mark

(ii) decision making as a DV _____ 1 mark

Question 9 (10 marks)

To test the effectiveness of a new sleeping pill, a researcher conducts an experiment at the participants' homes rather than in a sleep laboratory.

Eighteen volunteer adult participants, who reported that they have been suffering from sleep-onset insomnia (i.e. difficulty falling asleep) for more than a year, are each given a packet of 14 pills and asked to take one each night for 14 consecutive nights, 15 minutes before their usual sleeping time. They are also given a special apparatus to record the time they fall asleep. The apparatus, worn on the body, measures various physiological responses associated with sleep–awake states, has a timing device and has been reported by participants in previous studies as not being uncomfortable in any way.

The participants do not know that they have been randomly allocated to either of two groups. The researcher's assistant is also unaware of the group to which each participant has been allocated. Group 1 has nine participants whose pills are arranged in the pack so that pills 1 to 7 are the new sleeping pills, and pills 8 to 14 look and taste like the sleeping pills but do not contain the sleep-inducing chemical. Group 2 also has nine participants, but their pills are arranged so that pills 1 to 7 are the fake pills and pills 8 to 14 are the new sleeping pills.

The results are shown in the following table.

TABLE 1 Time taken to fall asleep

Group	Mean time (minutes)	
	Sleeping pills	Non-sleeping pills
1	37	64
2	78	31

(a) Operationalise the independent and dependent variables in the experiment:

independent variable _____ 1 mark

dependent variable _____ 1 mark

(b) Explain an ethical standard that should have been followed for this particular experiment. 1 mark

(c) Explain the difference between a placebo effect and an experimenter effect in relation to this particular experiment. 2 marks

(d) Name the experimental research design used by the researcher. 1 mark

UNCORRECTED PAGE PROOFS

(e) Discuss the double blind procedure in relation to this experiment. 1 mark

(f) Discuss counterbalancing in relation to this experiment. 1 mark

(g) Identify a potential extraneous or confounding variable that may have affected the results of the experiment and the extent to which this variable was controlled. 2 marks

UNCORRECTED PAGE PROOFS

eBook*plus*

The answers to the Section A multiple-choice questions are in the answer section at the end of this book and in eBookPLUS. The answers to the Section B questions are in eBookPLUS.

UNCORRECTED PAGE PROOFS