

chapter nine

Scientific research methods

Key knowledge and skills

This knowledge includes:

- experimental research: construction of research hypotheses; identification of operational independent and dependent variables; identification of extraneous and potential confounding variables including individual participant differences, order effects, experimenter effect, placebo effects, artificiality, demand characteristics, and non-standardised instructions and procedures; ways of minimising confounding and extraneous variables including type of experiment, counterbalancing, single and double blind procedures, placebos, type of sampling procedures, and standardised instructions and procedures; evaluation of different types of experimental research designs including independent-groups, matched-participants, repeated-measures; reporting conventions
- sampling procedures in selection and allocation of participants; random sampling; stratified sampling; random-stratified sampling, convenience sampling; random allocation of participants to groups; control and experimental groups
- techniques of qualitative and quantitative data collection: case studies; observational studies; self reports; questionnaires; interviews; brain imaging and recording technologies
- statistics: measures of central tendency including mean, median and mode; interpretation of *p*-values and conclusions; reliability, including internal consistency;

validity, including construct and external; evaluation of research in terms of generalising the findings to the population

- ethical principles and professional conduct: the role of the experimenter; protection and security of participants' rights; confidentiality; voluntary participation; withdrawal rights; informed consent procedures; use of deception in research; debriefing; advantages and limitations of the use of non-human animals in research in terms of generalisation and conclusions; role of ethics committees.

These skills include the ability to:

- formulate research questions and construct testable hypotheses
- design and conduct investigations using experimental and non-experimental methods such as observation studies and case studies
- collect, record and summarise both quantitative and qualitative data
- analyse and interpret data, and draw conclusions consistent with the research question
- evaluate the validity and reliability of research investigations including potential confounding variables and sources of error and bias
- work independently and collaboratively as appropriate and within identified research constraints
- adhere to current occupational health and safety codes and ethical guidelines for conducting psychological investigation.

SCIENTIFIC RESEARCH METHODS

Steps in psychological research

Experimental research

- Experimental and control groups
- Hypothesis
 - Experimental
 - Research (operational)
- Variables
 - Operational
 - Independent
 - Dependent
- Extraneous and confounding
 - Artificiality
 - Demand characteristics
 - Non-standardised instructions and procedures
 - Individual participant differences
 - Order effect
 - Experimenter effect
 - Placebos and placebo effect
- Experimental research designs
 - Independent-groups
 - Matched-participants
 - Repeated-measures

Sampling procedures

- Probability sampling
 - Random
 - Stratified
 - Random-stratified
- Non-probability sampling
 - Convenience
- Assigning participants to groups
 - Random allocation

Data-collection techniques

- Qualitative and quantitative data
- Case studies
- Observational studies
- Self-reports
 - Questionnaires
 - Interviews
- Brain imaging and recording

Statistics

- Measures of central tendency
- Statistical significance and p -values
- Validity and reliability
- Generalisation

Ethical principles and professional conduct

- Experimenter's role
- Participants' rights
 - Confidentiality
 - Voluntary participation
 - Withdrawal rights
 - Informed consent procedures
 - Use of deception
 - Debriefing
- Use of non-human animals
- Role of ethic committees

Psychology and the scientific method

Most people agree that biology, chemistry and physics are all sciences. However, many people refer to psychology as 'just common sense'. This reference is unfounded, as psychologists employ the same **scientific method** used in biology, chemistry or any other scientific experiments – and this chapter looks in detail at the many methods and conventions psychologists use in research investigations. Without psychological research, we would probably understand little about the brain and how it functions – and about human behaviour, intelligence, personality, creativity and cognitive development.

The scientific method involves testing the truth of a proposition by means of careful measurement and controlled observation (see Figure 9.1). It consists of a series of orderly and systematic steps that are used to plan, conduct, interpret and report research. In psychological research, seven steps are used; these are shown in Figure 9.2 and outlined below.



Figure 9.1 Applying the scientific method to the study of behaviour requires careful observation. Here, a psychologist videotapes a session in which a child's problem-solving abilities are being tested.

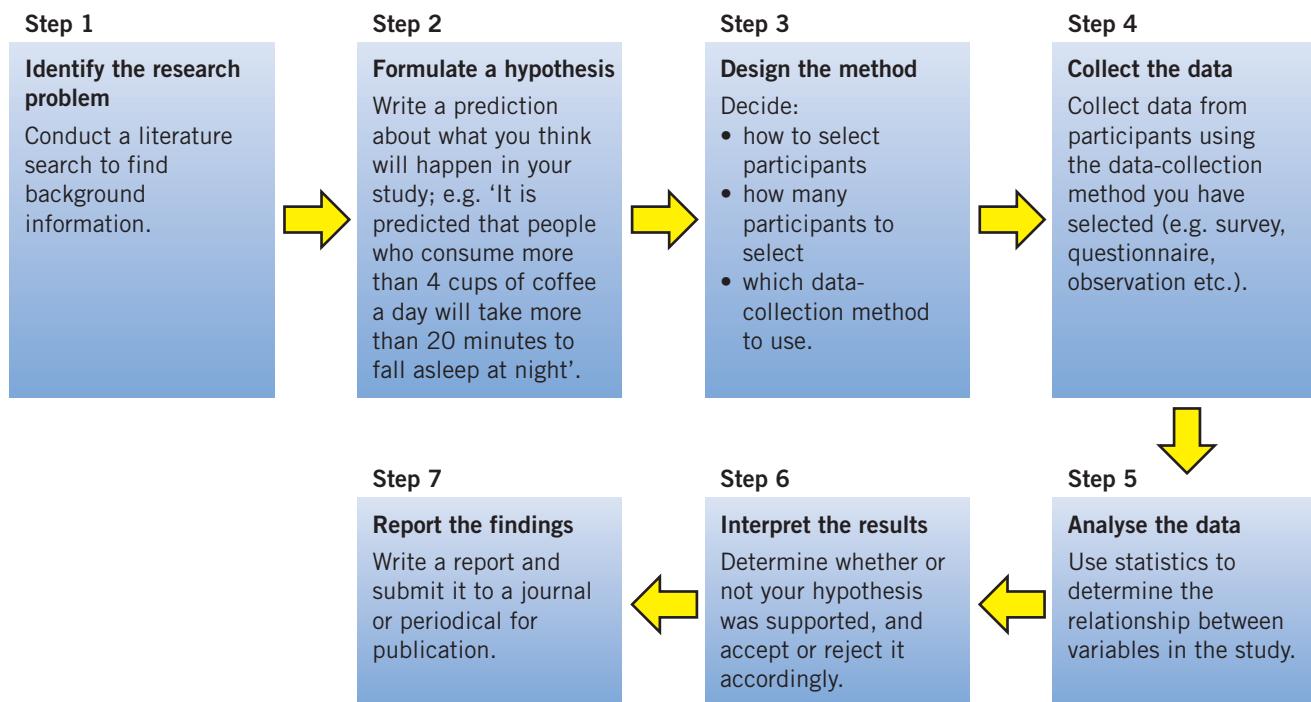


Figure 9.2 Summary of the steps involved in psychological research

scientific method

A data-gathering method that involves testing a hypothesis by means of careful measurement and controlled observation

STEP 1: IDENTIFY THE RESEARCH PROBLEM

The first step in psychological research is to identify the area in which to conduct a study. For example, a researcher may be interested in investigating the causes of eating disorders in adolescent females. To do this, the researcher first needs to conduct a *literature search*, which involves finding and reading relevant research and literature on the topic. Usually, past research is reported in periodicals and journals, which you can find in university libraries or on the Internet. In the case of eating disorders, a literature search might show that factors such as low self-esteem and previous use of weight-control measures (for example, laxatives and vomiting) may be related to future incidence of eating disorders in adolescent females. At this point, a testable general research question needs to be developed. In this case it might be something like: 'Do factors such as self-esteem and laxative use predict the likelihood of adolescent females developing eating disorders?'

STEP 2: FORMULATE A HYPOTHESIS

A **hypothesis** is a testable prediction of the relationship between two **variables**. A variable is any condition (event or characteristic) that can change. In other words, a hypothesis is an educated guess about what the researcher thinks the results of the experiment will be, based on information gathered in the literature search. That is, it is not just a guess made without prior knowledge or investigation.

The hypothesis is developed before the research is conducted and is written as a specific statement. An example of a hypothesis for the research into eating disorders might be: 'It is predicted that adolescent girls with low self-esteem are at higher risk of developing an eating disorder than adolescent girls with high self-esteem.'

STEP 3: DESIGN THE METHOD

The third step in research is to design the method of undertaking the research. The research method determines how the researcher will test the hypothesis. The type of research method chosen depends on the type of topic or problem being investigated. While designing the method, the researcher needs to determine:

- which **participants** will be studied
- how many participants will be studied
- how participants will be selected
- how participants will be allocated to various groups in the study.

Participants are the people or animals whose behaviour, characteristics or responses are investigated and measured as part of the experiment. Participants' responses generally form a study's **data** – the observed facts that constitute the results. (Note that 'data' is plural; 'datum' is singular.)

After choosing the participants, the researcher needs to decide how the data will be collected. For example, will you use a controlled experiment, an observational study or a case study? For the eating disorders research, if you chose to survey participants using a *questionnaire* (see page 312) to investigate the hypothesis, your design might be to give every female aged between 11 and 14 years who was attending any of three Melbourne secondary schools a survey about their eating behaviours – such as use of laxatives, frequency of vomiting and binge eating – and their feelings about their body image. You could plan to measure each participant's height and weight. Two years later, you could return to the same three schools and gather the data again from the same girls to see if their original feelings about their body image have affected current behaviour between the first survey and the current time.

Alternatively, you could design the study to use an *experimental method* (see page 296) to test the relationship between self-esteem and eating disorders. The participants from all three schools might be divided into two groups. All participants could be given questionnaires and interviews to initially gauge their eating habits and feelings about body image. One group might be given a series of training seminars on how to improve self-esteem. This could be followed up with more interviews and surveys, to gauge the effect of the seminars on their self-esteem and see if this has any effect on their eating behaviour.

STEP 4: COLLECT THE DATA

When undertaking the research and collecting the data, the researcher needs to follow the plans developed in Step 3. To measure participants' responses, a researcher can use any of a variety of data-collection techniques, including questionnaires, direct observation, psychological tests or examination of files and documents. For example, for the eating disorders research, a researcher might choose to gather data about the participants' levels of self-esteem using a questionnaire, and might gather data about their eating behaviours using an interview.

STEP 5: ANALYSE THE DATA

Data analysis involves objectively organising, summarising and representing raw data in a coherent and logical manner. **Raw data** are the actual data collected from a study. There are usually large amounts of raw data that are then 'broken down' into a smaller set or sets of numbers (for example, an average score in a set of scores).

STEP 6: INTERPRET THE RESULTS

Interpreting results involves forming conclusions about what the data show. A **conclusion** is a decision or judgement about the meaningfulness of a study's results. **Inferential statistics** are statistics used to make inferences and conclusions about the data, and

they are often used to interpret a study's results. Such statistics are a way of explaining the meaning of the data.

For example, if statistical analysis of the data from the eating disorder research clearly indicated that females who suffered from eating disorders had low self-esteem two years before the onset of the disorder, the data might be interpreted as supporting the hypothesis. You could conclude that female adolescents aged 11–14 years who have low self-esteem are more likely to suffer from an eating disorder in the future than adolescent girls who have high self-esteem.

STEP 7: REPORT THE FINDINGS

The report describes the background information used to formulate the hypothesis, gives information about the participants (such as age and gender), describes the procedures and materials used, and provides data analysis and interpretation (including graphs, tables and statistical tests). Reports also outline any problems or limitations encountered while conducting the study, and indicate how these may have affected the results. In addition, reports present conclusions based on the study's findings, as well as a list of references that were used to plan the study and prepare the report.

Once a research report is published, other researchers use it in their literature searches and further investigations. Publication also enables the general public to benefit from research findings.

REPORTING CONVENTIONS

Psychologists use a set format when reporting their research findings. It is important to note that this is a guideline only and often needs to be modified to suit a particular investigation.

Each formal report has two important characteristics. First, it needs to provide enough information about the study – for example, sampling method, materials used, procedure and data analysis – so that another researcher who reads the report is able to replicate the study. Second, the language used is very important: it needs to be clear, concise and written in the third-person and past tense.

Figure 9.3 on the next page shows the structure of a research report. Following are descriptions of each section in the formal report.

Title

The title briefly identifies what the investigation is about. Often it is easier to write a title after you have written the aim.

Abstract

An abstract is a brief summary (about 100 words) of the entire report. It includes a statement about the aim, participants, method, results and main conclusions.

Introduction

The introduction should be between 200 and 600 words. It provides background information obtained from literature searches on the topic. You may want to briefly describe the aim, method, results and conclusions of a couple of previous studies in the research area. You may also need to define specific terms mentioned and discuss relevant theories and concepts.

The introduction should lead to the current study's purpose. It includes the *aim* – a general, non-directional statement about what you want to investigate. The aim is followed by the research *hypothesis*, which also identifies the *independent* and *dependent variable(s)* (see pages 296–8).

Method

The method should be approximately 150–200 words and divided into three subsections that clearly describe how the investigation was conducted. It should provide the following information:

- Participants – How many? From which population was the sample drawn? How were participants selected and allocated to groups?
- Materials – Describe all materials used (for example, a questionnaire, word list A, word list B).
- Procedure – A detailed description of the steps in the experiment, presented in logical sequence. Other information required includes how the independent variable was manipulated, the researcher's role(s), and the instructions given to participants.

hypothesis

A testable prediction of the relationship between two variables

variable

Any event, condition or characteristic that changes (varies) or can be made to change

participants

The people or animals whose behaviour, characteristics or responses are investigated and measured as part of an experiment

data

The observed facts that constitute the results of an experiment

raw data

The actual data collected from a study, before it is sorted or analysed

conclusion

A decision or judgement about the meaningfulness of the results of a study

inferential statistics

Statistics that allow an experimenter to make inferences and conclusions about data; they are often used to interpret results of a study

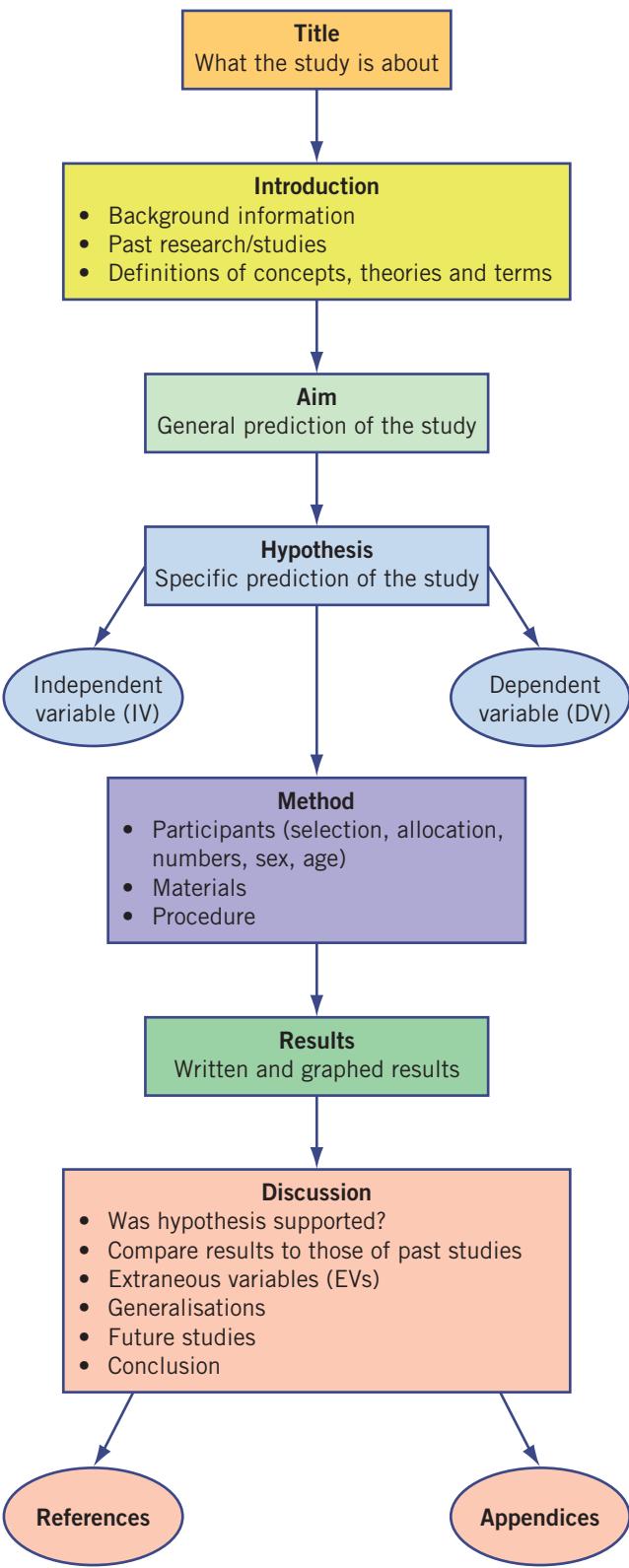


Figure 9.3 The structure of a research report

Results

The results section should be approximately 150–200 words and it should provide a *summary* of the *main* results. Results can be displayed in tables, charts or graphs, or as mathematical calculations. When using charts, graphs and tables, correct labels must be used, a descriptive title must be stated, and results should be described in a couple of short sentences.

If mathematical or statistical calculations are performed as part of the study, these also need to be reported in the results section. Lists of raw data are *not* included. Raw data are attached as appendices (see page 295).

Discussion

The 200–600-word discussion section describes and explains the results, and explains the conclusions that have been drawn. A statement is given about whether the hypothesis was supported or rejected, based on the findings. Any unplanned-for variables that may have affected the results are identified and explained and suggestions are given about how to control them in future research. Results of your study are compared with the results of the research mentioned in the introduction. Finally, the discussion should state whether the results of the sample can be generalised to the population from which the sample was drawn.

References

The references list should include all references cited in your report. Psychologists use the referencing method specified by the American Psychological Association (APA) (1994). Any study or other publication in psychology will contain in-text reference citations, and a reference list that is generally at the end of the paper.

CITING REFERENCES IN TEXT

All references cited in the text must appear in the reference list and, conversely, every entry in the reference list must be cited somewhere in the text. You need to use the Author–Date format when referencing material within text; that is, give the author's surname and the date of publication for the item you are citing. For example: '... Smith and Watson (1990) found that ...', or '... can be seen from these results (Smith & Watson, 1990)'.

If there are more than two authors, you name all the authors the first time you cite them in the text, then in subsequent citations you use the first author's name and the words *et al.* followed by the date. For example, you would use 'Smith, Watson, Percy and Jones (1999)' the first time you referred to their publication, and then in subsequent mentions you would use 'Smith et al. (1999)'.

THE REFERENCE LIST

The following examples show how publication details for books, journals and Internet sites are presented in the reference list. A reference list should be in alphabetical order according to the author's surname.

Books

- **Author:** The author's surname is given first, followed by their initials (for example, 'Green, A. H.'). All names of multiple authors are cited in the order they appear on the book's title page (which may not necessarily be alphabetical order).
- **Year of publication:** The publication year is enclosed in brackets, followed by a full stop (for example, 'Green, A. H. (1998)').
- **Title:** The title is italicised (or, if this is not possible, underlined), and followed by a full stop (for example, 'Green, A. H. (1998). *The beginnings of perception*.')).
- **City of publication:** The city of publication is followed by a colon (for example, 'Green, A. H. (1998). *The beginnings of perception*. Melbourne: Barretts Press.').
- **Name of publisher:** The publisher's name is followed by a full stop (for example, 'Green, A. H. (1998). *The beginnings of perception*. Melbourne: Barretts Press.').

An example of a full reference list entry for a book by a single author is as follows:

Garner, D. M. (1991). *Manual for the Eating Disorder Inventory*. Odessa, Florida: Psychological Assessment Resources Inc.

An example of a full reference list entry for a book by multiple authors is as follows:

Booth, W. C., Colomb, G. G., & Williams, J. M. (1995). *The craft of research*. Chicago: University of Chicago Press.

Journals and periodicals

- **Author:** The author's surname is given first, followed by their initials (for example, 'Brown, T. B.'). All names of multiple authors are cited in the order they appear on the book's title page (which may not necessarily be in alphabetical order).
- **Year of publication:** The publication year is enclosed in brackets, followed by a full stop (for example, 'Brown, T. B. (1999)').
- **Title of the article:** The article's title is followed by a full stop (for example, 'Brown, T. B. (1999). From sensation to perception.').
- **Title of the journal:** The journal title is italicised (or, if this is not possible, underlined), and followed by a comma (for example, 'Brown, T. B. (1999). From sensation to perception. *Psychological Studies*.').
- **Volume of the journal:** The journal's volume number and issue number (where relevant) are italicised (or, if this is not possible, underlined) and followed by a comma (for example, 'Brown, T. B. (1999). From sensation to perception. *Psychological Studies*, 23,'). If the journal has an issue number as well as a volume number,

the issue number follows the volume number in brackets, with no spaces.

- **Page numbers:** Page numbers are followed by a full stop (for example, 'Brown, T. B. (1999). From sensation to perception. *Psychological Studies*, 23, 145–52.').

An example of a full reference list entry for a journal article by multiple authors is as follows:

Striegel-Moore, R. H., Solberstein, L. R., Frensch, P., & Rodin, J. (1989). A prospective study of disordered eating among college students. *International Journal of Eating Disorders*, 8, 499–509.

An example of a full reference list entry for a journal article by a single author, but which shows the journal issue number as well as the volume number, is as follows:

Pollner, M. (1998). The effects of interviewer gender in mental health interviews. *Journal of Nervous and Mental Disease*, 186(6), 369–73.

Internet sites

Use the following format when referencing material from an Internet site:

Dewey, R. A. (2002). Psych Web by Russ Dewey. Retrieved 25 January 2010 from <http://www.psywww.com>.

Appendices

An appendix (plural appendices) is a section of material that does not fit in the body of a piece of work. The appendices section therefore includes all extra materials that did not fit into the body of the report (for example, the raw data). Make sure that any appendices are referred to in the body of the report (for example, 'See Appendix 1 for a table of the raw data.').

CHECK YOUR UNDERSTANDING 9.1

- 1 Match each term with its definition.

a Hypothesis	i The section of a formal written report that summarises the entire report
b Results	ii A judgement about what the results mean
c Conclusion	iii A testable prediction about the results of the experiment
d Abstract	iv The data yielded by an experiment
e Discussion	v The section of a formal written report that states whether the results of the experiment support or reject the hypothesis
- 2 Indicate whether the following statements are true (T) or false (F).

a There are six steps in the scientific method employed for psychological research.	b In a formal written report of an experiment, the introduction comes before the abstract.
---	--

- c The method section of a formal written report of an experiment identifies participants, materials and procedures.
- d Raw data appears in the results section of a formal written report.
- 3 Which of the following statements is true of a hypothesis?
- A A hypothesis is formulated after the research design method is chosen.
- B A hypothesis is formulated after the research problem has been identified.
- C A hypothesis is formulated after the results are analysed.
- D A hypothesis is formulated after the conclusion(s) has been drawn
- 4 Fill in the gaps with the correct terms.
- a In experiments, events or characteristics that change or can be made to change are called _____.
- b Statistics that allow you to draw conclusions about data are called _____ statistics.
- c When conducting experiments, psychologists follow the _____ method.
- 5 In a psychology experiment, the steps, in order, of the scientific method are: 1 Identify the research problem; 2 _____; 3 _____; 4 Collect the data; 5 Analyse the data; 6 _____; 7 Report the findings.

EXPERIMENTAL GROUP AND CONTROL GROUP

A psychological experiment usually has two categories of groups. The **experimental group** is the group (or groups) exposed to the **independent variable (IV)**. The independent variable is the variable controlled by the researcher in order to gauge its effect on the **dependent variable (DV)**. The dependent variable is the condition or the aspect of participant behaviour that is being measured. For example, psychologists interested in testing the effect of a new drug on sleeping habits would give the drug (IV – the variable being manipulated) to the experimental group to take. In turn, the DV would be the amount of hours of uninterrupted sleep obtained by participants – that is, the behaviour measured after manipulation of the IV. (Independent and dependent variables will be discussed in more detail on page 298.)

The **control group** is the group that is exposed to the **control condition** – that is, where the IV under investigation is absent from the conditions experienced, but all other conditions are the same as those of the experimental group. The control group provides a standard that the performance of the experimental group can be compared with, in order to determine if the treatment (IV) has had an effect on behaviour (DV). For example, in the experiment into the effect of a new drug on sleep, the control group may be given a sugar pill or no pill, rather than the drug. All other conditions experienced by the control group and the experimental group would be the same. The amount of hours of uninterrupted sleep obtained by the control group would then be compared to those of the experimental group, to see if the drug taken by the experimental group has had any effect on sleep.

Suppose you notice that you seem to study better while you are listening to music. This might suggest the hypothesis that music improves learning. We might test this idea by forming an experimental group, where people study with music. Meanwhile, we would also form a control group, where people study without music. The sample from which we drew both groups might consist of Year 10 students of average intelligence and a B-average in their studies. (Sampling techniques will be examined later in this chapter.) We might then compare the two groups' scores on a test after they study under the different conditions. Without a control group it would be impossible to tell whether music had any effect on learning. The control group provides a point of reference against which to compare the experimental group's scores. Figure 9.4 shows that the difference between the control and experimental group is exposure to the independent variable.

Sometimes there is no control group in an experiment, but instead there are two or more experimental groups that are differentiated by two or more levels of the IV. For example, a study may investigate the sleep patterns of people with high and low anxiety. The two experimental groups in this case are *high anxiety* and *low anxiety*.

Experimental research: Where cause meets effect

The **experimental method** is the most widely used research method in psychology. The experimental method is a scientific research method that uses participants in a formal trial to confirm or disconfirm a hypothesis. In an **experiment**, data is gathered under controlled conditions to test a hypothesis by exposing participants to a treatment and observing and measuring its effect to determine whether the treatment influences or causes a change in the aspect of their behaviour that is of interest.

Psychologists favour experiments because experiments allow them to control the conditions experienced by participants. Under controlled conditions, psychologists can better determine the effects of a treatment – in other words, identify cause-and-effect relationships between variables. For example, in a controlled experiment investigating the effects of caffeine on sleep, researchers can manipulate the strength and amount of coffee that participants consume. Researchers can then test sleep quality in a sleep laboratory that enables them to control variables such as noise and light, so that these do not affect the results. The data from such an experiment give us clear information about cause-and-effect relationships between variables.

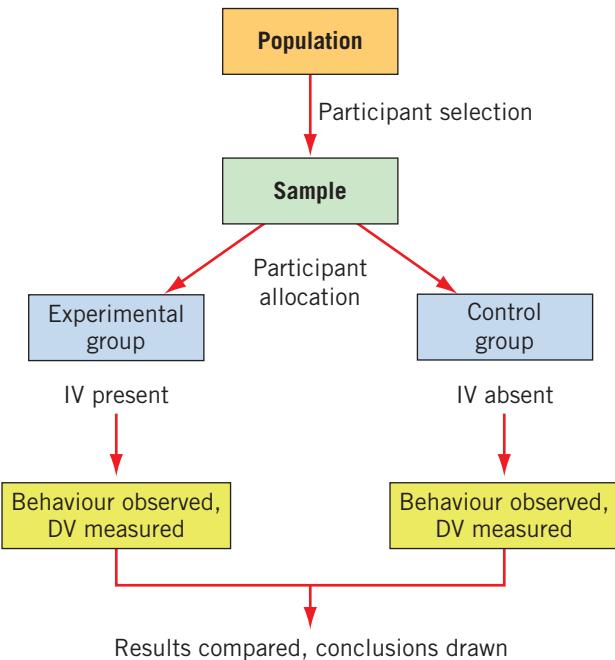


Figure 9.4 The control group is not exposed to the IV. The experimental group is exposed to the IV. Results from each group are then compared to determine the effect of the IV.

Additionally, sometimes the same group is exposed to the experimental condition and the control condition at different times. This means that only one group is used twice. In this experimental set-up we do not say there is a control group and an experimental group; instead, we say there is a *control condition* and *experimental condition* experienced by the same group of people at different times.

THE RESEARCH HYPOTHESIS

Researchers must state a prediction about the relationship between variables in an experiment that can be tested. This prediction about the results is known as an **experimental hypothesis**. An experimental hypothesis is usually a broad and general prediction about the *direction* of the relationship between variables in an experiment.

'Direction' refers to whether the variables increase or decrease in relation to one another. For example, you might predict that consumption of alcohol will decrease driving ability. This is obviously a very general statement: it does not tell us what *levels* of alcohol or what *type* of alcohol are being referred to, or how 'driving ability' might be being measured.

A **research hypothesis** (also known as an **operational hypothesis**) must achieve more than an **experimental hypothesis** (see Figure 9.5). A research

experimental method

A scientific research method that uses participants in a formal trial to confirm or disconfirm a hypothesis

experiment

A research method that involves gathering data under controlled conditions to test a hypothesis by exposing participants to a treatment and observing and measuring its effect

experimental group

In a controlled experiment, the group of participants exposed to the independent variable

independent variable (IV)

The condition that an experimenter systematically manipulates (changes or varies) in order to gauge its effect on another variable (the dependent variable)

dependent variable (DV)

The condition in an experiment or aspect of the participant's behaviour that is affected by changes in the independent variable (IV); it is used as a measure of the IV's effect

control group

In a controlled experiment, the group of participants exposed to all conditions or variables except the independent variable

experimental hypothesis

A broad and general prediction about the direction of the relationship between variables in an experiment – i.e. whether the variables increase or decrease in relation to one another

research hypothesis

A hypothesis that operationalises the variables by precisely defining and describing how each variable is measured, and predicts the exact effect the IV is expected to have on the behaviour of the population from which the sample has been selected

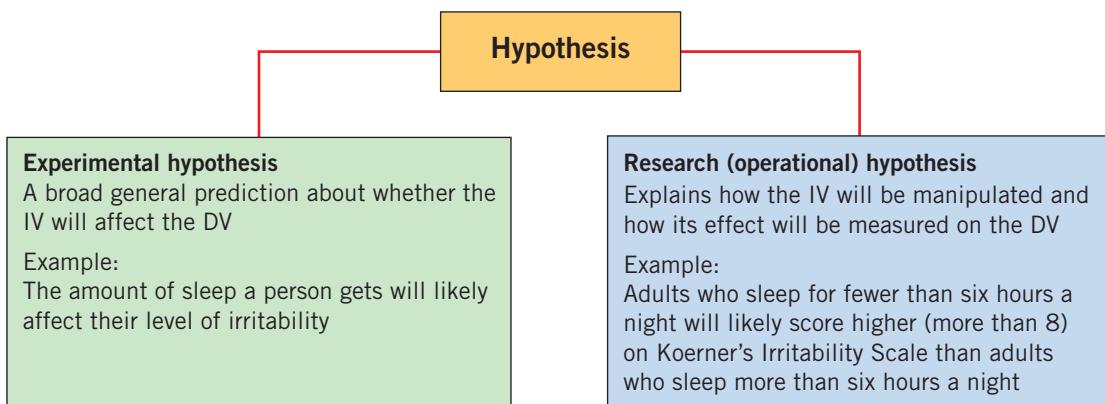


Figure 9.5 Basic differences between an experimental hypothesis and a research (operational) hypothesis

hypothesis clearly identifies the variables and states what *operations* (procedures or manipulations) will be used to observe and measure the variables.

Because the research hypothesis provides a testable prediction of the outcome, it must also predict the relationship between variables. Therefore, it must state the effect the IV is expected to have on the behaviour of the *population* from which the *sample* has been selected. A population is the larger group of research interest from which a sample has been drawn, and a sample is the group of participants selected from, and representative of, a population of research interest (see Figure 9.6). An example of a research hypothesis that operationalises the IV and DV would be: It is predicted that adults who sleep fewer than six hours a night will be more likely to score high (more than 8) on Koerner's irritability scale than will adults who sleep more than six hours a night.



Figure 9.6 A sample (in this case, the cordial in the glass on the right) is a representative subset of the population of research interest (in this case, the cordial in the jug).

In psychology, there is rarely a 100 per cent chance of something happening. There may be a 99 per cent chance that one variable is affecting another variable, but we do not know for sure. Therefore, when writing hypotheses, psychologists usually avoid using absolutes such as 'never', 'always' or 'will have'. Instead, they use terms such as 'more likely' or 'less likely'.

'Try it yourself 9.1' contains an activity on writing a hypothesis. View 'Videolink: Hypothesis testing' to see an example of hypothesis testing.

VIDEO
Hypothesis testing

TRY IT YOURSELF 9.1

Writing an operational hypothesis

In order to transform an experimental hypothesis into an operational/research hypothesis, we need to complete the following steps.

- 1 Write an experimental hypothesis (a broad, general statement about how one variable will affect another variable) for a topic you are interested in, such as the effect of alcohol on driving ability.

- 2 Identify all the variables in the experimental hypothesis.
- 3 Describe how each variable will be measured in the study.
- 4 Replace the variables in your experimental hypothesis with the description of variables in Step 3 so as to form the research hypothesis.

It may help to model your work on the following example.

- 1 Sample hypothesis: It is predicted that consumption of alcohol will decrease driving ability.
- 2 Variables: alcohol consumption and driving ability
- 3 Alcohol consumption = five standard alcoholic drinks in one hour; Driving ability = number of errors made on a driving simulator
- 4 It is predicted that the consumption of five standard alcoholic drinks in one hour will decrease driving ability as measured by the number of errors on a driving simulator.

OPERATIONAL VARIABLES: IVs AND DVs

Research hypotheses predict the relationship between variables, but we have not yet examined in detail exactly what a variable is. Earlier we noted that a *variable* is any event or characteristic that changes (varies) or can be made to change and that might affect an experiment's outcome. Variables can be external factors that differ among different people, such as the amount of caffeine drunk in a day or amount of exercise engaged in over a week. Yet individual factors, such as gender, ethnic background, height and blood type, are also classified as variables for the purpose of psychological research.

We also noted earlier that when psychologists engage in experimental research they define or describe an event in *operational* terms – or in terms of the procedures used to observe and measure it. When psychologists conduct research they are often interested in observing and measuring abstract concepts such as aggression or intelligence. In order to do this, they must operationalise the concept, or make it physically measurable or testable. Therefore, when we identify an *operational variable*, we are defining or describing exactly *what* that variable is and exactly *how* it will be measured. For example, when investigating intelligence we might measure the variable of intelligence by the number of puzzles solved in an hour. We have therefore operationalised the variable by defining exactly how it will be measured.

Independent variables (IVs) and dependent variables (DVs) are forms of operational variables. As you know, an IV is a condition that an experimenter systematically manipulates (changes or varies) in order to gauge its effect on another variable. An IV

is the suspected cause of differences in behaviour or results between an experimental and control group. The IV is said to cause a change in another variable (the DV).

A dependent variable (DV) is the condition in an experiment, or aspect of the participant's behaviour, that is affected by changes in the IV. The DV is used as a measure of the IV's effect – it reveals the effects that exposure to the IV has had on behaviour. Such effects are often revealed by measures of performance, such as test scores. DVs 'depend' on the effects of another variable – the IV – as to how much they change and the way in which they change. For example, a researcher examining the effects of a new treatment therapy for schizophrenia might divide a sample into two groups. The experimental group would be given the treatment (IV) while the control group would receive no treatment, then the effects of the treatment would be measured (DV).

To demonstrate the difference between an IV and a DV, imagine you were a researcher involved in testing the effectiveness of a new memory drug. You ask participants to learn a list of nonsense syllables (such as BIX, CFI and WOL) and then test their retention of the words by asking them to write down as many of them as they can recall. The two variables being tested would be the new drug and the retention of the nonsense syllables – but which is the IV and which is the DV? The IV is the variable manipulated by the experimenter, so, because the experimenter is manipulating the amount of drug given to the participants, the IV is the memory drug. The DV is the resulting behaviour of the experiment, so, if the results are the number of words remembered, then the DV must be the number of words remembered.

CHECK YOUR UNDERSTANDING 9.2

- 1 Fill in the gaps with the correct terms.
 - a In order to test a possible cause-and-effect relationship, a simple psychological experiment creates two groups: the _____ group and the _____ group.
 - b Experiments are conducted to prove or disprove a _____.
 - c The results of an experiment show the relationship between two _____.
- 2 Which of the following statements is incorrect about psychological experiments?
 - A The control group is exposed to the IV.
 - B The control group is used as a standard against which the performance of the experimental group is compared.
 - C The experimental group is exposed to the IV.
 - D The experimental group's performance is measured against the standard set by the control group.

- 3 In the following examples of research, identify the IV and the DV.

- a The effect of alcohol consumption on speed of reflexes
- b The effect of taking steroids on aggression levels
- c Recording a car's mileage level when using premium grade petrol, then recording it again using standard grade petrol

- 4 The factor that is being manipulated in an experiment is called the:

- A hypothesis.
- B control group.
- C dependent variable.
- D independent variable.

- 5 Indicate whether the following statements are true (T) or false (F).

- a A research hypothesis involves a broad prediction about the relationship between the IV and the DV.
- b A sample is the large group of research interest.
- c A population is a subset of a sample.

EXTRANEous AND CONFOUNDING VARIABLES

An **extraneous variable** is a variable other than the IV that might cause unwanted changes in the DV. Extraneous variables compromise a study because, when analysing the results, they make it difficult to determine whether any change in the DV was caused *solely* by the IV and no other factor. Since extraneous variables can affect the results of the experiment, the experimenter must ensure that they are eliminated, or that their influence is controlled. When this elimination occurs, the extraneous variable is described as a **controlled variable**.

If you conducted research on the effect of coffee consumption on performance of learning tasks, your IV would be the amount of coffee drunk and your DV would be the score or performance level on a

population

The larger group of research interest from which a sample in a research study has been drawn

sample

The group of participants in a research study selected from, and representative of, a population of research interest

operational variable

A variable defined or described in terms of the procedures used to observe and measure it

extraneous variable

In an experiment, a variable other than the IV that might cause unwanted changes in the DV

controlled variable

An extraneous variable whose influence has been eliminated from an experiment so that it cannot affect results; it has been controlled

learning task, such as recall of a list of numbers. In this experiment an extraneous variable might be alcohol consumed by the participants, since this too might affect performance on the learning task. If you wanted to control or eliminate this variable, you might only choose participants for your sample who did not drink alcohol at all, or monitor participants for an amount of time leading up to the experiment to make sure they did not consume alcohol in that time.

Extraneous variables may be present without the researcher knowing, and they may not be identified until after the experiment is complete (if they are identified at all). These **uncontrolled variables** are important because they can confuse the interpretation of results. This means that our assumption that the manipulation of the variable X (the IV) has or has not affected variable Y (the DV) may be inaccurate.

VIDEO

Psychology experiment

When an extraneous variable actually succeeds in confusing (or *confounding*) the results, it becomes a **confounding variable**. Confounding variables actually change the cause-and-effect relationship between the IV and the DV. When researchers plan their experiment, it is crucial that any potentially confounding variables that might be associated with the experiment are identified and controlled, so that they become controlled extraneous variables (which do not have a chance to affect results) rather than confounding variables (which do end up affecting results). ‘Focus on research: Experimental cola’ demonstrates why. ‘Videolink: Psychology experiment’ outlines a simple experiment.

FOCUS ON RESEARCH

Experimental cola

There has always been intense rivalry between the brands Pepsi-Cola® and Coca-Cola®. Consequently, executives at Pepsi-Cola conducted a taste test to determine if Coca-Cola drinkers preferred the taste of Coca-Cola or Pepsi-Cola. Regular Coca-Cola drinkers were asked to taste two samples of cola – one being Pepsi-Cola, the other Coca-Cola – and to identify which one they preferred (Sdorow, 1993).

In order to hide the identity of each of the colas, Pepsi-Cola was in a cup labelled ‘M’ and Coca-Cola was in a cup labelled ‘Q’. The results indicated that the majority of participants preferred the cola labelled ‘M’: Pepsi-Cola. Pepsi-Cola was then advertised as the cola Coca-Cola drinkers prefer (Sdorow, 1993).

Coca-Cola executives decided to repeat the taste test to see if it yielded the same results. This time they filled both cups (‘M’ and ‘Q’) with Coca-Cola. Results indicated that most participants still preferred the cola in the cup labelled ‘M’, even though the same cola was in each cup! They concluded that the Pepsi-Cola experiment did not demonstrate that Coca-Cola drinkers prefer Pepsi-Cola to Coca-Cola, but instead that Coca-Cola drinkers prefer the letter ‘M’ to the letter ‘Q’. The effect of the letters on the cups (being an extraneous variable) had been confounded with the type of cola (the IV). The IV did not cause the change in

the DV, which was the participants’ preference for Pepsi-Cola (Sdorow, 1993).

QUESTION

- How might researchers test whether participants prefer Pepsi-Cola or Coca-Cola without the study results being affected by extraneous variables? Design an experiment to reduce or eliminate any extraneous variables.

We will now take a look at some reasons why extraneous and confounding variables occur.

Artificiality

The fact that many research experiments are conducted in a laboratory setting that has been expressly organised for the experiment gives the researcher control over variables. However, as control increases, the environment in which data are collected becomes more unnatural or artificial. Participants know they are in an unnatural environment and this knowledge may have an unwanted effect on their behaviour and therefore an unwanted effect on the results. If participants alter their natural behaviour, the study’s results will not be an accurate representation of their behaviour. So, in trying to control as many variables as possible, experimenters may actually introduce another variable: **artificiality**. Artificiality is the unwanted effect on participant behaviour created by the unnatural environment in which the experiment is conducted.

For example, if a researcher wanted to examine the effects of a new drug on sleep patterns, participants would be required to sleep overnight in a sleep laboratory so that their sleep patterns could be monitored. However, some people do not sleep ‘normally’ when they are not in their own bed, so the results of such a study may be affected by the unfamiliarity of the sleep laboratory.

One way of reducing artificiality is to conduct the study in a non-laboratory environment. If a *field study* was used, participants would be observed and their behaviour measured in their natural setting. This increases the chance that the data will accurately represent their true behaviour.

For example, psychologists may be interested in whether people are more likely to help someone in distress if they have just seen someone else receiving help. To test this, they might set up a situation where a man is changing a flat tyre for a woman beside a busy road. From the road, the man would be clearly seen changing the tyre while the woman looks on (see Figure 9.7). Two blocks further up on the same road, psychologists might then set up another flat tyre situation. This time, a woman might be standing beside a stopped car. A spare tyre and a jack might be placed beside the car so that it can easily be seen that the woman needs her tyre changed. The psychologists would then record how many people stop to help the woman with the flat tyre, having just seen a man helping another woman two blocks earlier.



Figure 9.7 Psychologists interested in helping behaviour might set up a field experiment involving a woman needing help changing a flat tyre.

It must be remembered, however, that conducting experiments outside of the laboratory makes other extraneous variables harder to control.

Demand characteristics

If the participants' knowledge of a study's aim causes them to behave in a way that is not normal for them, then the study has been affected in an unwanted way by **demand characteristics**. For example, participants might deliberately behave in a way they think will please the researcher or, alternatively, they may deliberately try to spoil or bias the results to compromise the study's reliability and validity.

Demand characteristics can be overcome by concealing the real aim of the experiment from the participants. Such deception, however, would need the approval of the ethics committee to whom the researcher originally submitted the plan of their research study. This will be examined later when we discuss ethical principles of psychological research (page 319).

Non-standardised instructions and procedures

Test **standardisation** involves two things. First, it means that when a test is given to a number of participants, often at different times and in different places, standard procedures are used so that there is a consistency in procedure. The instructions, answer forms, amount of time in which participants must respond, the way tests are scored and so forth are the same for all participants in standardised tests. By standardising instructions and procedures, researchers have a greater chance of ensuring that the study's results are valid and reliable. Non-standardised tests may result in scores or responses that misrepresent participants' true characteristics.

Second, standardisation involves finding the **norm**, or average score, made by a large group

of people such as those for whom the test was designed. Without standardisation, we couldn't fairly compare the scores of people taking the test at different times. And without norms, there would be no way to tell if a score is high, low or average. For example, without norms, a score on an intelligence test would be a meaningless number. Norms are established by giving the test to samples of hundreds or thousands of people who are representative of the people for whom the test is designed. If a test is to be used in Victoria, for example, samples might include representative proportions of gender groups, age groups, and city and country dwellers across the state. A standardised test is typically appropriate for specific groups and specific purposes. They have norms that can be described in terms of numbers that represent arithmetical averages (means) and distribution.

The use of non-standardised instructions and procedures would remove the systematic basis for making inferences about people. Unless a test is reliable and valid, it cannot measure behaviour accurately, and unless it has been standardised, there is no way to determine the meaning of an individual's score. Therefore, reliability and validity are important criteria for judging a test's value, and standardisation is essential in judging its utility. (Reliability and validity are discussed later in this chapter; pages 317–18.)

VIDEO

ANOVA – Sources of variance in an experiment

'Videolink: ANOVA – Sources of variance in an experiment' outlines the concept of 'ANOVA' (analysis of variance) in experimental research, which relates to the topic of standardisation. This is part A of the video – if you have time, you may wish to view part B.

Individual participant differences

Some extraneous variables that can influence the DV occur as a result of individual differences in the personal characteristics of participants. Collectively,

uncontrolled variable

An extraneous variable whose influence has not been eliminated from an experiment because the experimenter was not aware of it

confounding variable

An uncontrolled variable that has had an unwanted effect on the DV and might be confused with the effect of the IV

artificiality

The unwanted effect on participant behaviour created by the unnatural environment in which an experiment is conducted

demand characteristics

When participants' knowledge of the aim of a study causes them to behave in a way that is not normal for them; this affects the results of the study

standardisation

Establishing standards for administering a test and interpreting scores

these are known as **participant variables**. Participant variables include individual characteristics such as memory, motivation, mood, age, gender, prior experience, personality, expectations, ethnicity, religion and ability.

All participant variables have the potential to confound the results of an experiment; therefore, it is important that the researcher controls them. One way of achieving this is to ensure that all participants are as similar as possible in terms of the personal characteristics that are relevant to the experiment when they are selected for the sample and when they are allocated to either the experimental or control groups. Thus, the researcher must choose their research design (see pages 304–5) carefully.

Order effect

Another possible extraneous variable can come in the form of the **order effect** (also known as the practice effect). Order effect may be experienced when the experimental group and the control group are made up of the same participants – that is, the one group experiences two (or more) different conditions or tests. Order effect occurs when prior knowledge of a task or situation influences a participant's performance and therefore influences the results of the experiment. That is, when a participant learns from one condition or test, this knowledge influences (positively or negatively) his or her performance on the next condition or test. Order effect can result in *improved* performance on a second test due to practice; alternatively, a participant's performance might be *impaired* in the second condition because of boredom or fatigue due to previous experience with the same task.

Increasing the time period between completion of the first condition and the second condition helps to overcome the effects of boredom, fatigue or practice when using the same participants for both conditions. For example, participants may be in the control condition one week, then in the experimental condition a week later. If this method of minimising order effect is inappropriate or impossible, the researcher can then *counterbalance* the conditions.

Counterbalancing alters the order that participants experience each condition. In the first instance, half the participants are exposed to the control condition and the other half are exposed to the experimental condition. In the second instance, this is reversed: those who experienced the experimental condition first now experience the control condition, and those who experienced the control condition first now experience the experimental condition. Counterbalancing helps to balance order effects over the experiment so that each effect occurs equally in both conditions.

Experimenter effect

Extraneous variables may also include differences in how the experimenter treats the participants. The **experimenter effect** refers to changes in participants'

behaviour that are caused by the unintentional influence of the experimenter rather than the IV itself. This is a common problem in psychological research: in essence, experimenters run the risk of finding what they expect to find.

The experimenter effect also applies outside the laboratory. Psychologist Robert Rosenthal (1973) reported an example of how expectations can influence people. He randomly assigned 100 prep-grade children of equal ability to five different maths classes. The children's teachers did not know about this random placement. Instead, each teacher was told that his or her students had unusually high or low ability. Students in the 'high ability' classes improved much more in their maths scores than those in the 'low ability' classes, even though all the classes had students of equal ability.

In this case the teachers apparently communicated their expectations subtly to students – the teachers expected the 'low ability' students to do poorly and the 'high ability' students to do well. Most likely they communicated these expectations through tone of voice and body language, or by giving different levels of encouragement or criticism. Their 'hints' created **self-fulfilling prophecies** that affected their students. Self-fulfilling prophecy refers to a prediction that prompts people to act in ways that make the prediction come true. In short, some people sometimes become what we expect of them (Jussim & Eccles, 1992; Madon, Jussim & Eccles, 1997).

A **double-blind procedure** can be used to counteract the experimenter effect. A double-blind procedure involves neither the experimenter nor the participants knowing which experimental condition the participants have been allocated to. In other words, they are not informed of, or are 'blind' to, which participants are experiencing which conditions and this prevents experimenters from unconsciously influencing participant behaviour. Typically in these experiments someone other than the experimenter controls the IV. In this way, it is possible for the experimenter not to know which participants were exposed to the IV until after testing, therefore

 **VIDEO**
Double-blind clinical trials

eliminating experimenter influence on the results. 'Videolink: Double-blind clinical trials' demonstrates the double-blind procedure.

Placebo effect

Imagine that an experiment is conducted to see whether a particular stimulant affects learning. Before they are due to start studying, members of the experimental group take an amphetamine pill, while members of the control group take nothing. How much each participant learnt will be assessed later. Does this experiment seem valid? The study is actually seriously flawed, and as such it could not be said that any differences in the amount learnt will have been caused by the drug.



Figure 9.8 Sugar pills are often used as placebos, and work by altering a person's expectations about their own emotional and physical reaction to the pill

The **amount** of drug (i.e. a quantity or no quantity) was not the only difference between the groups – the added difference was the fact that experimental group participants swallowed a pill, and control group participants did not. It might be that those who swallow the pill expect to do better, and this alone might affect their performance, even if the pill does not. So, without both groups having swallowed a pill in this experiment, it is impossible to tell whether the drug alone has affected learning.

But if we don't want to actually give the drug to both groups, how can we ensure conditions for both groups are the same? In this case we would use a **placebo**. A placebo is a fake treatment that has no active effect. Inert substances such as sugar in pill-form and salt water by injection are common placebos. So, in the example described above, both groups would swallow a pill but the control group would be given a placebo while the experimental group would be given the drug.

In turn, the **placebo effect** refers to changes in behaviour caused by the *belief* that one has been exposed to a treatment that will affect them in some way. This effect can be very powerful. For instance, it has been found that a saline injection is 70 per cent as effective as morphine in reducing pain. This is why doctors sometimes prescribe placebos – especially for complaints that seem to have no physical basis. Placebos have been shown to have an effect on pain, anxiety, depression, alertness, tension, sexual arousal, cravings for alcohol and many other processes (Kirsch & Lynn, 1999).

Placebos work by altering people's expectations about their own emotional and physical reactions (see Figure 9.8). These expectancies in turn influence bodily activities. For example, placebos that relieve pain do so by causing the pituitary gland to release endorphins. These powerful chemicals are similar to painkilling opiate drugs such as morphine (Ter Riet, de Craen, de Boer & Kessels, 1998). Thus, if a placebo has any effect on participants' behaviour, the effect must be due to participant expectations rather than to the effect of an active chemical.

To control the placebo effect, researchers often use a **single-blind procedure**, which is a procedure designed so that participants do not know if they

participant variables

Individual differences in the personal characteristics of research participants that, if not controlled, can confound the results of the experiment

order effect

Where prior knowledge of a task or situation influences a participant's performance, which in turn influences the results of the experiment; also known as the practice effect

counterbalancing

A method used to control order effect, where half the participants in an experiment are exposed to the control condition first and the other half are exposed to the experimental condition first; this is then reversed in the second instance

experimenter effect

Changes in participants' behaviour that are caused by the unintended influence of the experimenter rather than the IV

self-fulfilling prophecy

A prediction that prompts people to act in a way that makes the prediction come true

double-blind procedure

An experimental procedure where neither the experimenter nor the participants know which experimental condition the participants have been allocated to

placebo

A fake treatment that has no active effect, such as a fake pill or injection

placebo effect

Changes in behaviour caused by the belief that one has been exposed to a treatment that will affect them in some way

single-blind procedure

An experimental procedure where participants do not know which experimental condition they have been assigned to, but the experimenter does

are receiving a real drug or a placebo. In such experiments, all participants receive a treatment (for example, a pill or injection). People in the experimental group are given an active treatment that can affect them in some way, while those in the control group are given a placebo. Because participants are ‘blind’ as to whether they received the active treatment, expectations regarding the treatment do not differ across the groups. Any difference in behaviour between the groups is most likely caused by the active treatment. Note that while participants do not know which group they belong to, the experimenter does – this is how a single-blind procedure differs from a double-blind procedure.

CHECK YOUR UNDERSTANDING 9.3

- 1 Match each term with its definition.
 - a Extraneous variable i A variable other than the IV that has an unwanted effect on the DV, which is confused with the effect of the IV
 - b Confounding variable ii A variable other than the IV that may cause an unwanted change in the DV
 - c Placebo iii When participants and the experimenter don't know which group is exposed to the IV
 - d Placebo effect iv When participants' behaviour is affected in an unwanted way by being in an unnatural setting, such as a laboratory
 - e Artificiality v A fake treatment that has no active effect
 - f Double-blind procedure vi Changes in behaviour caused by the belief that you have been given something that will affect your behaviour
 - 2 Indicate whether the following statements are true (T) or false (F).
 - a Extraneous variables are the variables manipulated by the researcher.
 - b Independent variables are suspected causes for differences in behaviour.
 - c The placebo effect occurs because participants believe they have been given a real treatment that will affect their behaviour.
 - 3 Fill in the gaps below with the correct terms.
 - a In a _____ study, participants don't know if they are given a treatment or a placebo, but the experimenter does.
 - b _____ variables are conditions that a researcher wants to prevent from affecting the _____ of an experiment.
- c A _____ variable's unwanted effect on the DV is confused with the effect of the IV.
- d _____ variables might possibly cause an unwanted change in the DV that could be confused with the effect of the IV.
- 4 If participants make up both the experimental and control groups, they may perform better in the second condition because of prior learning, or they may perform worse because of boredom or fatigue. This unwanted effect can be overcome by:
A the order effect.
B counterbalancing.
C the confounding variable.
D standardising instructions and procedures.
- 5 The term ‘artificiality’ refers to:
A individual differences in participant's personal characteristics that, if not controlled, can confuse the results.
B the effect that a participant's knowledge of the experiment's aim alters their normal behaviour.
C changing the standards for administering a test and interpreting scores.
D an unwanted change in participants' behaviour being caused by conducting an experiment in a laboratory rather than in their natural environment.

EXPERIMENTAL RESEARCH DESIGNS: MINIMISING UNWANTED VARIABLES

The best way to control or eliminate unwanted variables is to ensure that groups of participants are as similar as possible in characteristics believed to be relevant to, or that might otherwise influence, the experiment. This is achieved by carefully selecting the experimental research design and by controlling the way participants are allocated to groups in the experiment. Each of these designs has advantages and limitations, and each is discussed below. Figure 9.9 shows the characteristics of each design.

Independent-groups design

The **independent-groups design** (also known as *between-groups design*) randomly allocates participants to either the experimental group or control group. The independent-groups design is the experimental design used most frequently in psychological research.

The main goal of the independent-groups design is to ensure that each member of the sample has an equal chance of being selected for the control group as for the experimental group. This ensures that the two groups are well matched on important personality characteristics and are thus reasonably similar. A limitation of this design is that equally-sized groups are difficult to achieve with a small

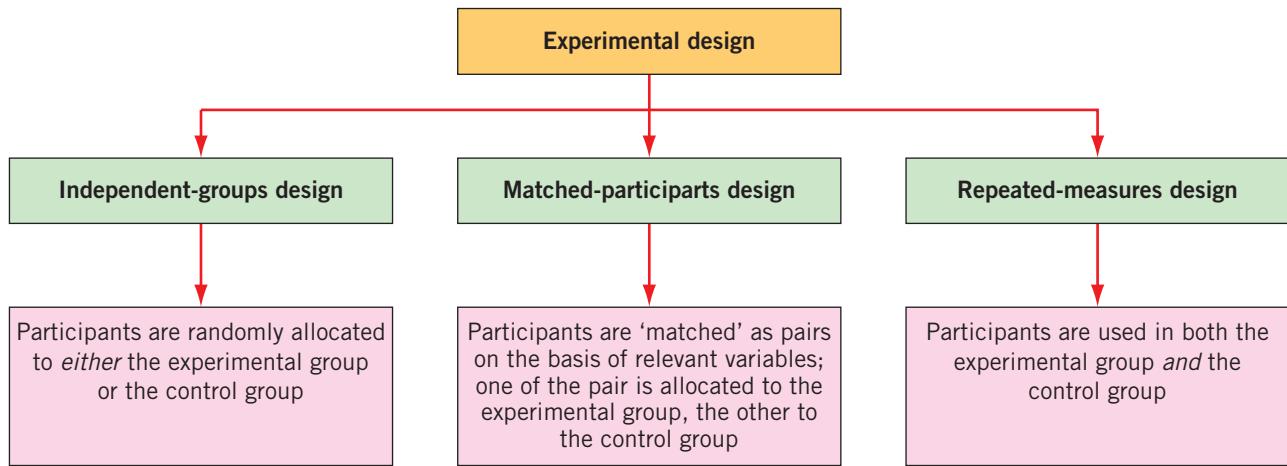


Figure 9.9 A summary of the characteristics of the three different experimental designs

sample size. One way of overcoming this problem is to increase the number of participants in each condition. An advantage of the independent-groups design is that there are no order effects to control, and the independent-groups design is relatively quick and easy to complete.

Matched-participants design

A **matched-participants design** matches (in pairs) participants on the basis of similar characteristics that can influence the DV, with one of the pair being randomly allocated to the experimental group and the other to the control group. Randomly allocating paired individuals ensures that groups are fairly similar in terms of the individual personality characteristics that are of research interest.

While in a matched-participants design the experimental and control groups are usually very similar in terms of important personality characteristics, researchers are unable to perfectly match all participant characteristics. It is therefore likely that other, unconsidered personality characteristics may affect the DV in a matched-participants design. Another limitation of this design is that the process of selecting participants, matching them in pairs and then randomly assigning individuals from each pair to each group is quite time-consuming.

Repeated-measures design

A researcher may choose a **repeated-measures design** (also known as a within-groups design) where the same group of participants makes up both the experimental and control groups. This results in minimal differences in personality characteristics between the experimental and control groups.

In addition, fewer participants are required for the study.

Because the same participants are used for both groups, does this mean a repeated-measures design is problem-free? Not necessarily. One limitation of the repeated-measures design that needs to be controlled is the order effect (see page 302). However, counterbalancing cancels out the order effect because participants are equally exposed to control and experimental conditions.

CHECK YOUR UNDERSTANDING 9.4

- 1 Indicate whether the following statements are true (T) or false (F).
 - a In a matched-participants design, the same group of participants make up both the experimental and control groups.
 - b In a repeated-measures design, participants are exposed to one condition first, after which they are exposed to a different condition.

Continued ▶

independent-groups design

An experimental design where participants are randomly allocated to either the experimental group or control group

matched-participants design

An experimental design where participants are paired (matched) on the basis of similar characteristics that can influence the DV, with one of the pair being allocated to the experimental group and the other to the control group

repeated-measures design

An experimental design method where the same group of participants makes up both the experimental and control groups

- c A repeated-measures design is used to minimise the effect of differences in participant's personal characteristics.
- d A matched-participants design pairs participants on the basis of similar personal characteristics.
- 2 Fill in the gaps with the correct terms.
- a Within-groups design is another name for a(n) _____ design.
- b _____ is used to control possible order effects.
- c A(n) _____ design gives all members of the sample an equal chance of being assigned to either the control or the experimental group.
- d A(n) _____ design would balance out the possible unwanted effects of extraneous variables such as age, gender, ethnicity or religion.
- 3 Which of the following statements is incorrect?
- A An independent-groups design does not result in the order effect.
- B A repeated-measures design can result in the order effect.
- C Counterbalancing offsets the order effect.
- D Random sampling offsets the order effect.
- 4 Which of the following research designs uses counterbalancing to control the effect of an unwanted variable?
- A A repeated-measures design
- B A matched-participants design
- C An independent-groups design
- D All of the above
- 5 The main goal of an independent-groups design is to:
- A give every member of the population an equal chance of being selected for either the control or experimental group.
- B give every member of the sample an equal chance of being selected for either the control or experimental group.
- C control the unwanted effect of extraneous variables associated with the experimenter.
- D control the unwanted effect of extraneous variables associated with the participants.

from which it is drawn. Ultimately, psychologists are interested in entire populations, but it is often impossible to study an entire population. By selecting a smaller sample, psychologists can draw conclusions about the larger group without polling each and every person in it. The larger the sample taken from a population, the more representative of that population it is likely to be.

It is important that psychologists undertake proper sampling procedures so that results can be generalised to wider populations. 'A closer look: Is there a gender bias in psychological research?' outlines the importance of this.

A CLOSER LOOK

Is there a gender bias in psychological research?

Many doctors recommend that all adults – whether male or female – take an aspirin each day to help prevent heart attack. The problem? Not a single woman was included in the study sample on which this advice is based! Although females make up more than half of the population, they continue to be neglected in psychological and medical research (Denmark, 1994).

Without directly studying women, it is impossible to know how often such assumptions are wrong. A related problem occurs when researchers combine results from men and women, because doing this can hide important male–female differences.

Another problem is that unequal numbers of men and women may volunteer for some kinds of research; for example, more male university students volunteer to participate in studies of sexuality than do female university students (Wiederman, 1999).

Similar biases exist concerning race, ethnicity, age and sexual orientation in psychological research participants (Denmark, 1994). Far too many conclusions are based on small groups of people who are not representative of the enormous diversity in humanity. The solution to such problems is simple: where possible, most researchers now try to include a wider array of people in their studies.

The main types of sampling are shown in Figure 9.10. We will investigate these next.

PROBABILITY SAMPLING

Probability sampling involves all members of the population having an equal chance of being selected for a sample. In these samples, the mathematical probability that any one of the members of the population can be selected for the sample can be calculated. This means that the results of a probability sample are unbiased because every member of the population has the same chance of being selected. Probability sampling is used when the researcher needs the sample to be

Sampling procedures: Choosing participants

The process of choosing participants is called *participant selection*, or *sampling*. As previously mentioned, research participants are collectively called a *sample* – which is a subset of the *population*.

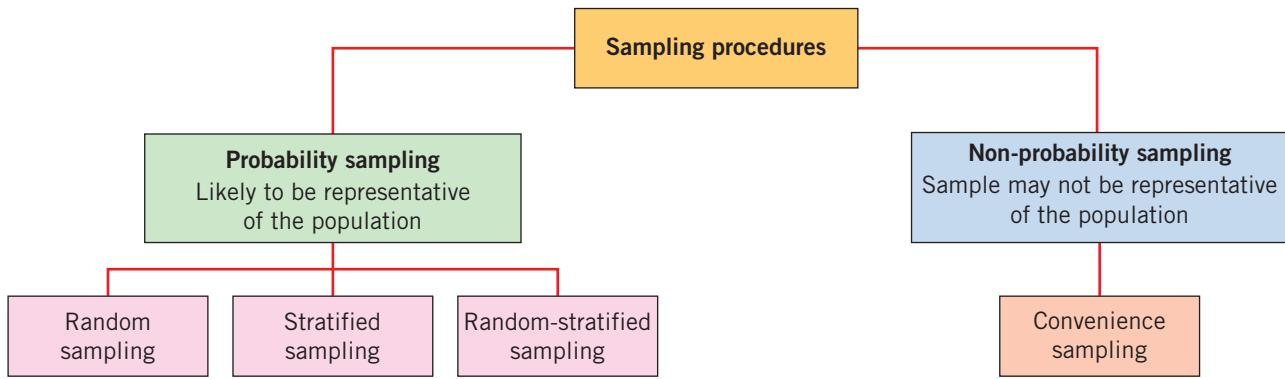


Figure 9.10 Types of samples: Probability samples involve all members of the population having a chance of being selected for a sample; therefore, they are more representative than non-probability samples.

representative of the population. If researchers want to generalise the conclusions of their study to the wider population, the results must be representative and unbiased.

Random sampling

Random sampling employs a carefully planned and systematic method of selecting participants for a study. Random sampling, a form of probability sampling, ensures that every member of a population has an equal chance of being selected for the sample. Further, the selection of one participant does not affect the chances of another participant being selected. Two common methods of random sampling are:

- putting the names of each member of the population of interest into a container and then pulling out a set number of participants for the sample
- generating a list of numbers (for example, student identification numbers) that are in no specific order and then choosing every fifth, 10th or 100th number (and hence participant) to be part of the sample.

Random sampling is suited to research involving a population that is homogenous (where members of the population are similar). It ensures that every member of the population has a chance of being selected; therefore, it increases the likelihood of the sample being representative of the population. Random sampling also improves the chances of making accurate inferences about the population based on the results gained from the sample because it involves a statistically-supported probability.

Random sampling does have limitations. Sometimes people are unwilling to participate in experiments, new therapies or research, so the sample becomes biased rather than representative. Furthermore, in some research a sample is sometimes difficult to obtain for a particular

population because the members of the population are not available to participate in the research. Random sampling may not always be possible because researchers may not be able to acquire a name list of the population if the population is very large. For example, if a researcher was studying sleeping habits of Australians over 60 years of age, it would be difficult to acquire a list of all people in this age group.

Stratified sampling

Stratified sampling is used in research that requires the sample to contain the same proportions of participants that are found in the population. Stratified sampling involves dividing the population into distinct subsets (or *strata*) that share at least one common characteristic of research interest, then selecting a separate sample from each group (or *stratum*) in the same proportion as occurs in the larger population.

Imagine that you want to investigate Australian's favourite foods in order to predict someone's favourite food based on characteristics such as age, ethnic background and religion. Say you predict that ethnic background is the characteristic likely to have the biggest impact on your research results. You establish that ethnic background in the target population is broken up as follows: 53 per cent Anglo-Saxon, 21 per cent Italian, 10 per cent Greek, 5 per cent Asian, 5 per cent Russian, 4 per cent Middle-Eastern and 2 per cent African. You divide

random sampling

A sampling technique ensuring that every member of the population of interest has an equal chance of being selected for the sample being used in a study

stratified sampling

A sampling technique that ensures the sample contains the same proportions of participants that are found in the population

your sample into these ethnic categories (strata). Your sample needs to contain 100 participants, so from the Anglo-Saxon ethnic group you randomly select 53 participants, then randomly select 21 participants from the Italian group, 10 from the Greek group, five from the Asian group and so on. Your sample now contains the same proportions of participants from different ethnic backgrounds as there are in the target population. You have created a stratified sample.

Stratified sampling is most commonly used when psychological characteristics or attitudes vary greatly among subgroups of the target population; therefore, it is most suited to populations that are dissimilar. Stratified sampling also reduces the possibility of sampling error and this, in turn, increases the precision of the results. The selection of a stratified sample does, however, require more time and effort than a random sample.

Random-stratified sampling

With **random-stratified sampling**, the population is divided into a number of strata according to some characteristic of interest related to the variable(s) being studied (for example, age, gender or income). Then, a list of all persons within each stratum is obtained. Simple random samples are then selected from each stratum to ensure that each member of the group has an equal chance of being selected for the sample. Random samples of proportionate size are drawn from within each stratum, making the sample a proportionately-stratified random sample.

Imagine you were a researcher conducting a survey of attitudes of homeowners in suburbs surrounding a proposed new shopping centre. It would be reasonable to suppose that individual property values would affect the homeowner's response. So the survey would use three strata from the population: those with houses in the bottom third of values, those with houses in the middle third of values and those with houses in the top third of values. Random samples would then be selected from each stratum. The same proportion will be selected within each stratum, making the sample a proportionately-stratified random sample.

NON-PROBABILITY SAMPLING

In **non-probability sampling**, participants are selected on the basis of their availability, so the sample may not actually be a representative sample. Also, the mathematical probability of selection in the sample cannot be calculated because some members of the population have no chance of being selected.

Convenience sampling

Convenience sampling (also known as opportunity sampling) chooses participants because they are readily available to the researcher. For example,

participants volunteer, or they are asked to participate. In other words, they are convenient.

Convenience samples, a form of non-probability samples, are used when researchers want an inexpensive approximation of the truth that will assist them to generate their hypothesis. They may also be working under time constraints or they may be unable to access the wider population. Before conducting the actual research experiment, researchers may conduct a pilot study using a convenience sample so they can identify any uncontrolled extraneous variables or weakness in method, and obtain a general estimate of the results.

If you answer a knock on your door and you agree to fill out a survey on the spot, you have become part of a convenience sample (see Figure 9.11). No random mechanism has been used to ensure that volunteers are representative of the research population, as only people who were at home at that time of the day and who were willing to participate were chosen. The researcher still doesn't know the views of other residents in the area. Other examples of convenience sampling you may be familiar with include newspaper, magazine, radio and phone polls.



Figure 9.11 Convenience samples consist of participants who are available and no mechanism has been used to ensure that participants are representative of the population.

Results from research using a convenience sample can be legitimately used provided the limitations associated with the research are clearly understood and stated. Although this is a quick and inexpensive way for a researcher to select a sample, the sample may not be representative of the population; some groups may be overrepresented while others are underrepresented. If the sample includes volunteers, it may also be biased. For these reasons, results of a convenience sample cannot be generalised (we will discuss generalisation of results on page 318).

'Try it yourself 9.2' contains a sampling activity.

TRY IT YOURSELF 9.2

Sampling

You will need:

- Slips of paper
- A hat (or any container)

Each member of your class is a member of the population of research interest. Each person is to write his or her name on a slip of paper and to put it into the hat. The experimenter is then to pull out 10 names from the hat. These names are the people in the sample.

QUESTIONS

- 1 What type of sampling is this?
- 2 Identify the advantages of this type of sampling.
- 3 Identify the disadvantages of this type of sampling.

ASSIGNING PARTICIPANTS TO GROUPS

Once participants have been selected, they need to be allocated to either the experimental group or the control group within the experiment. As with sampling, the allocation (or assignment) of participants must be done in a systematic and carefully planned manner. This is to ensure that possible extraneous variables associated with participants' individual personality characteristics are evenly distributed among the groups.

Random allocation

Random allocation is the most common method of assigning participants to experimental and control groups. Random allocation is an experimental procedure in which each participant in the sample has an equal chance of being selected for the experimental group (and, therefore, exposure to the IV) and the control group (and, therefore, not exposed to the IV).

The main aim of random allocation is to ensure that participants in the experimental group are as similar as possible to participants in the control group, in terms of personality characteristics of interest. Thus, if the two groups think, behave or feel differently from each other at the end of the experiment, then the difference is more likely to have something to do with the treatment the experimental group has been exposed to – the IV – rather than any pre-existing difference.

Random allocation can be as simple as flipping a coin to determine which group each participant is allocated to. Other methods employed to ensure random allocation include (as for random sampling) pulling names out of a container or randomly selecting numbers (such as student-identification numbers) from a generated list.

CHECK YOUR UNDERSTANDING 9.5

1 Match each term with its definition.

- | | |
|----------------------------|--|
| a Convenience sample | i A form of stratified sampling involving random samples of each stratum being selected for the sample |
| b Random sample | ii Participants that are easily accessible to the researcher and usually volunteer |
| c Stratified sample | iii Gives everyone in the population the same chance of being selected for the sample |
| d Random-stratified sample | iv Contains the same proportion of participants as exists in the population |

2 Indicate whether the following statements are true (T) or false (F).

- Random sampling is a form of probability sampling.
- Stratified sampling chooses participants because they are readily available to the researcher.
- Convenience sampling is a form of probability sampling.
- Random samples represent sub-groups according to the proportions in which they exist in the population.

3 Fill in the gaps with the correct terms.

- If a researcher placed the names of all participants in a hat, pulled out half and assigned them to the experimental group, and assigned the other half the control group, they would be using _____ to assign their participants to an experimental condition.
- _____ represents groups in the sample according to the same proportions they exist in the population.
- _____ involves dividing the population into strata based on a characteristic of interest and then randomly sampling each stratum to obtain a sample.

Continued ▶

random-stratified sampling

A form of stratified sampling involving random samples of each stratum being selected

convenience sampling

A sampling technique involving the selection of participants because they are readily available to the researcher

random allocation

A procedure for assigning participants to either the experimental group or control group in an experiment, ensuring that all participants have an equal chance of being allocated to either group

- 4 Which of the following statements is correct about random allocation?
- A Random allocation occurs before random sampling.
 - B Random allocation occurs after random sampling.
 - C Random allocation involves selecting groups in the proportion they exist in the population.
 - D Random allocation provides the most representative sample.
- 5 Convenience sampling:
- A is a form of probability sampling.
 - B is a form of non-probability sampling.
 - C is a form of random sampling.
 - D is a form of stratified sampling.

CASE STUDIES

A **case study** is an in-depth, detailed study of an aspect(s) of a single participant, group or event, usually undertaken to gain insight into a particular psychological phenomenon (for example, creative genius). Case studies are a non-experimental form of data-collection that may arise when accidents or other natural events provide psychological data. Case studies of individuals who have experienced gunshot wounds, brain tumours, accidental poisonings and other adversities resulting in brain damage have provided much information about the structure and function of the human brain. Case studies can also take the form of clinical tests, interviews or observational studies – or a combination of these. When conducting a case study, the researcher records as much relevant information as possible about a variety of factors, such as the person's thoughts, feelings, life experiences, relationships and behaviours.

Many experiments that might be revealing are impractical, unethical or impossible to perform. Case studies, however, allow researchers to gather information about a variety of psychological phenomena that cannot be obtained by any other means (Edwards, 1998). For example, if a psychologist wanted to research the origins of language development in infants, it would be unethical to deprive a group of infants of any exposure to language for 10 years in order to conduct an experiment. However, case studies of individual children who have unfortunately been deprived of normal human contact for years after birth (because of mistreatment) do provide an insight into language development in children. Insights such as these allow psychologists to establish psychological principles of human behaviour.

One limitation of case studies is that they are uncontrolled studies. Case studies lack formal control groups, and this limits the conclusions that can be drawn from them. In psychological research, case studies usually involve atypical cases of behaviour, often as a result of brain damage. Individuals who suffer similar brain injuries may demonstrate very different reactions to the injury, so this is why psychologists prefer controlled experiments and often use laboratory animals for studies of the brain.

Case studies allow researchers to describe behaviour, but because behaviour is determined by a number of variables, a case study cannot determine what *caused* the behaviour. Therefore, case studies provide useful insights into human behaviour but their results cannot be generalised. Analysing and reporting on the amount of detailed information yielded by the case study method is also a very time-consuming process.

OBSERVATIONAL STUDIES

Observational studies involve one person watching and recording the behaviour of another person(s) or animal(s) within a specific environment and drawing conclusions based on the recorded observations. When this non-experimental method of data collection

is used, the behaviour of interest must be *overt*, or able to be observed by others. For example, we can observe someone building a sandcastle, but we cannot observe how they are feeling as they are building it. In psychology, observational studies can be applied to a range of settings such as watching chimpanzee societies in the jungle (see Figure 9.12), watching parent-child interactions in different cultures, or recording students' self-seating patterns in the cafeterias of multiracial schools.



Figure 9.12 Naturalistic observation allows psychologists to study natural behaviour such as a chimpanzee using a grass stem to obtain termites from inside a nest.

Observational studies are an indirect means of gathering data and, similar to case studies, observational studies describe behaviour, but they do not explain behaviour.

In observational studies, the accuracy of the data gathered may be influenced by whether or not the person(s) being observed know they are being observed. If observational studies are carried out in situations where participants are observed in their natural environment and do not know they are being observed, this is an effective way of reducing the extraneous variable of artificiality. If, however, the observed person/animal is aware that their behaviour is being observed, they may alter their behaviour and thus bias the results. This is known as the *observer effect*. For example, if you were interested in the differences between aggressive and non-aggressive

schoolchildren, you could not simply stroll onto a playground and start taking notes. As a stranger, your presence would probably change students' behaviour. So, when possible, observers must be as unobtrusive as possible.

Observer bias is another limitation of observational studies. Observer bias occurs when the observer sees what he or she expects to see, or records only selected details – that is, such bias occurs when an observer's expectations, past experience, motives or other personal factors interfere with the accuracy of his or her observations.

SELF-REPORTS

In many situations, it is not possible to collect data about people's attitudes, beliefs and behaviours by observation only. For example, it would not be ethical to gather data about punishment methods employed by parents by observing them hitting their children. An alternative data collection method would be to ask people to report specific information about themselves.

In a *self-report*, individuals are simply asked to freely express their thoughts by answering questions (verbally or in writing) about a particular object, person, issue or experience. Self-reports are a form of *subjective data* because information given cannot be applied to other individuals.

Self-reports are a means of collecting both qualitative and quantitative data and questions involved may take a variety of formats: open-ended, fixed-response or indirect questions, or ratings on a multi-point scale such as a Likert scale.

qualitative data

Data that describe the changes in the quality of a behaviour; often accounts of personal attitudes or experiences, or descriptions of feelings

quantitative data

Data collected through systematic and controlled methodology and presented in numerical form

case study

An in-depth, detailed study of all aspects of a single participant, group or event, usually undertaken to gain insight into a particular psychological phenomenon

observational studies

A method of data-collection that involves watching and recording the behaviour of another person(s) or animal(s) within a specific environment and drawing conclusions based on the recorded observations

observer effect

Changes in the behaviour of a person being observed caused by their awareness of the presence of an observer

observer bias

Bias in results of an observational study that occurs when an observer sees what he or she expects to see, or records only selected details of an observed behaviour

self-report

A data-collection technique in which individuals are asked to freely express their attitudes (verbally or in writing) by answering questions

Open-ended questions ask individuals to comment freely and without limitation on their attitude towards a particular issue such as, ‘What are your thoughts about a mother breastfeeding her baby in public places?’ Open-ended questions often produce a large amount of descriptive data for analysis, which can give very specific, interesting and useful information about people’s attitudes; but sifting through such a large amount of information can be time-consuming. Also, people are free to express themselves in any way they wish, and some people might find it difficult to explain their thoughts and feelings. It can be difficult to analyse and group this data and determine what people truly mean or which answers are the most significant.

Closed questions, however, do not involve large amounts of information, nor do they require the participant to search for appropriate words to express their thoughts. Closed questions restrict responses to a limited choice of answers. For example, if a person was asked, ‘Do you believe teachers should be able to strike at any time for a pay rise?’, participants may only have the response options of ‘Yes’, ‘No’ or ‘Undecided’.

Although self-reports are a useful means of gathering data, they do have their limitations. Participants may misunderstand some questions or they may have a tendency to give socially acceptable answers rather than honest answers. Thus, it may be difficult to determine the accuracy of the data and the conclusions drawn from it.

Self-report questions can be asked and undertaken in the form of questionnaires or interviews.

Questionnaires

Self-report data can be gathered using a *questionnaire* (see Figure 9.13). A questionnaire is a written set of standardised questions that can be administered face-



Figure 9.13 Self-reports, in the form of a questionnaire, can provide both qualitative and quantitative data.

to-face, by mail, by telephone or via the Internet, and often take the form of a survey. Questionnaire types can range from factual to opinion-based, or from tick-the-boxes to free-response answers.

Questionnaires allow participants to remain anonymous, so they are more likely to be honest in their responses. However, the validity of results relies on participants’ honesty, and this cannot be controlled by the researcher. Another advantage is that responses are limited to a list of predetermined, standardised questions that cannot be varied from participant to participant. This may eliminate any extraneous variables that may be associated with non-standardised instructions or procedures. Also, because they can be administered by mail, telephone or the Internet, questionnaires are a relatively inexpensive way of gathering data when a large sample is required. They can also be administered to a large group simultaneously, which allows researchers to include a large geographical area in their study.

Questionnaires can provide both qualitative and quantitative data on attitudes, beliefs and behaviours; however, by their very nature, qualitative questions must often be more exact than quantitative questions. Qualitative questions assume that people will use words or understand language in the same way. This is not always the case, and there is the chance that some people may misunderstand some questions or that different people will interpret questions differently. Therefore, it is very important to avoid ambiguity in phrasing when constructing a questionnaire. Qualitative questions also require more thought and effort on the participant’s behalf. They may become tired or bored, so responses may not be an accurate reflection of their attitudes.

Interviews

Interviews involve person-to-person questions and answers, in the form of face-to-face contact or voice-to-voice contact (i.e. phone interviews). In an *unstructured interview*, conversation between the interviewer and the person being interviewed is informal, and topics are taken up freely as they arise. In a *structured interview*, the interviewer obtains information by asking a planned series of questions and tries not to deviate from the plan. The data collected is qualitative.

In addition to providing information, interviews make it possible to observe a person’s tone of voice, facial expression or body language – cues that can radically alter the interpretation of a message and cues that are absent from answers provided in written form. For example, a person may claim to be ‘completely calm’ while their body trembles. Interviews are also suited to gathering information from people who cannot write their responses due to physical deficiency or poor literacy levels.

Interviews do have limitations as a data-collection method. Deviations from set questions can make it difficult to summarise, organise and analyse the

Table 9.1 Comparison of questionnaires and interviews as data-collection methods

	QUESTIONNAIRES	INTERVIEWS
Advantages	<ul style="list-style-type: none">• Inexpensive in terms of time and money• Participants can remain anonymous and are likely to be honest in their responses• The format is standard for all participants and is not dependent on the mood of the interviewer• Information can be gathered from large samples, covering large geographic areas	<ul style="list-style-type: none">• Participants do not need to be able to read or write• Data may be obtained on any subject as questions are not limited to predetermined questions• Questions can be clarified if participants are not sure of their meaning
Disadvantages	<ul style="list-style-type: none">• Questions may be misunderstood	<ul style="list-style-type: none">• Participants cannot remain anonymous so may give dishonest answers to make themselves appear more socially acceptable• Not practical for large samples

results, and this puts the validity of the results in question. Also, interviewers can be swayed by their own preconceptions of the interviewee if they have stereotyped them. For example, if a person has been identified as 'a housewife', 'a heroin addict' or 'unemployed', they may be misjudged because of the interviewer's bias towards a particular lifestyle.

Additionally, participants cannot remain anonymous, and because responses are given in the presence of the interviewer (when in a face-to-face interview), interviewees may not answer honestly. Participants may try to deceive interviewers if they feel that their honest answer be regarded as socially unacceptable or incriminate them in some way. In addition, the interviewer's personality or behaviour may influence the interviewee's behaviour. When this occurs, it can accentuate or distort the person's apparent traits (Pollner, 1998).

Table 9.1 compares the questionnaires and interviews as data-collection methods.

BRAIN IMAGING AND RECORDING TECHNOLOGIES

Much information about the brain has been gathered using the methods just described, plus via autopsy or investigations carried out while a person undergoes invasive and dangerous brain surgery. Fortunately, psychologists are no longer limited to these methods of data collection. Modern technology has provided new 'windows' into the brain, making it possible to gather data about the living brain's structure and function. A variety of devices now exists that allow researchers to compile and interpret large amounts of information by recording or imaging the brain (see Figure 9.14).

As we discovered in chapter 2, tasks and tests can now be devised to determine the processes and locations of sensation, perception, cognition, motor responses and emotional responses in the brain. Clear images of damaged brains and recordings of their activity levels can now be compared with images and recordings of healthy, intact brains as

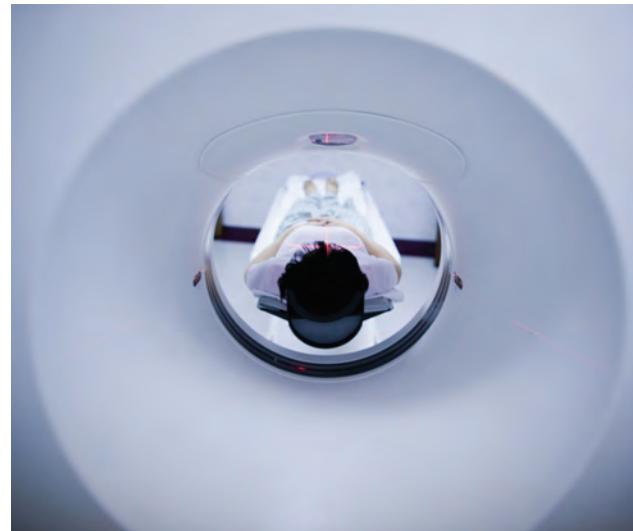


Figure 9.14 Modern technologies, such as taking a CT scan, provide detailed information about brain structure.

they function in real time. This information has added enormously to our understanding of the processes of sensation, perception, cognition, motor and emotional responses. Brain imaging techniques have enhanced psychologists' abilities to determine problem sites and diagnose causes of cognitive or behavioural losses as well as assist them with formulating treatment options. Modern brain imaging techniques have revolutionised the study of the brain and more sophisticated technology is being invented every day.

questionnaire

A written set of standardised questions that can be administered face-to-face, by mail, by telephone or via the Internet

interview

A form of qualitative data-collection where individuals are asked to comment on their attitude towards a particular issue(s)

VIDEO

Brain imaging

View 'Videolink: Brain imaging' for a look at brain imaging technology in action.

CHECK YOUR UNDERSTANDING 9.6

- Indicate whether the following statements are true (T) or false (F).
 - A person studying elephant behaviour in the elephant's natural surroundings is an example of an observational study.
 - In observational studies, concealing the observer will limit observer bias.
 - A tendency to give socially desirable answers can lower the accuracy of data gathered via interview.
- Fill in the gaps below with the correct terms.
 - _____ data are presented in numerical form.
 - Self-reports are a form of _____ data.
 - _____ do not allow participants to remain anonymous.
 - The _____ occurs when the presence of an observer influences participant behaviour.
- Which data-collection method would yield the most detailed information about the effects of tumours in the frontal lobe of the brain?
 - Brain recording techniques
 - Brain imaging techniques
 - A case study
 - Questionnaires
- An advantage of a questionnaire is that:
 - because participants remain anonymous, they are more likely to give honest answers.
 - the pre-determined questions don't vary between participants, so they eliminate extraneous variables associated with non-standardised procedures.
 - they don't have to be administered face-to-face, so they can be used simultaneously with large groups.
 - All of the above
- Match each term with its definition.

a Qualitative data	i Information usually presented in numerical (statistical) form
b Quantitative data	ii A list of standardised questions about a participant's thoughts
c Questionnaire	iii Describes the changes in the quality of a behaviour, usually written in words
d Interview	iv When another person asks an individual to comment on their attitude to an issue

Statistics: Analysing and interpreting data

Researchers use two forms of statistics to help them analyse and interpret data: inferential and descriptive statistics.

Inferential statistics (discussed earlier in this chapter) are used to interpret results so researchers can decide what the results mean. They allow researchers to determine the significance of the results. Researchers use inferential statistics to make inferences (draw conclusions) about whether the results support or do not support the research hypothesis and decide whether the conclusions can be generalised to the population of research interest.

Descriptive statistics are used to describe, summarise, organise and analyse data so that it can be more easily interpreted and explained to others. Examples of descriptive statistics include graphs, tables, measures of central tendency and frequency distributions. Descriptive statistics do not determine the meaningfulness of the results; that is, they do not determine whether or not the results support the research hypothesis.

MEASURES OF CENTRAL TENDENCY: MEAN, MEDIAN AND MODE

One form of descriptive statistics is the measure of central tendency – a number describing a 'typical score' around which other scores fall; that is, the tendency for a majority of scores to fall in the mid-range of possible values.

A familiar measure of central tendency is the mean, or 'average', but other measures of central tendency include the median and the mode. To illustrate each measure, we will use an example: Table 9.2 shows

Table 9.2 Raw scores on a memory test for participants taking Rememberine or a placebo

PARTICIPANT NUMBER	GROUP 1 (REMEMBERINE)	GROUP 2 (PLACEBO)
1	65	54
2	67	60
3	73	63
4	65	33
5	58	56
6	55	60
7	70	60
8	69	31
9	60	62
10	68	61
Sum of participants	650	540
Mean	65	54
Median	66	60
Mode	65	60

the raw data for an imaginary experiment into the effects of a potentially memory-enhancing drug (let's call the drug Rememberine). Group 1 was given the drug. Group 2 received a placebo. Both groups then undertook a test of memory and their scores recorded.

In order to tell whether there is a difference in memory scores between the two groups, we need to compute an average score for each group: the *mean*. As one type of 'average', the mean is calculated by adding all the scores for a given group, then dividing the total by the number of scores in the group.

So, the mean for Group 1 is our experiment = $65 + 67 + 73 + 65 + 58 + 55 + 70 + 69 + 60 + 68 = 650$.

There are 10 participants in the group, so we divide the total (650) by 10, which gives a mean of 65.

The same is done for Group 2. As you can see in Table 9.2, the group means reveal a difference between the two groups.

Let us look at another set of data, with the following scores: 23, 14, 26, 45, 33, 23, 11, 20, 29, 50, 13, 17, 22.

To determine the mean for this data set, we must first add the scores together: $23 + 14 + 26 + 45 + 33 + 23 + 11 + 20 + 29 + 50 + 13 + 17 + 22 = 326$

Then you divide the total (326) by the number of scores (13), which yields the mean for this set of scores: $326 \div 13 = 25$

The mean is sensitive to very high or very low scores in a distribution. For this reason it is not always the best measure of central tendency. (Imagine how distorted your results would be if you were to calculate average yearly income using a small sample of people that happened to include a multimillionaire!). In such cases, the middle score in a group of scores is used; this is known as the *median*.

The median is found by arranging scores from the highest to the lowest and selecting the score that falls in the middle. In other words, half the

values in a group of scores fall below the median and half fall above.

When we arrange the scores for Group 1 in Table 9.2 in order, we have: 55, 58, 60, 65, 65, 67, 68, 69, 70, 73. Because there are 10 scores, an even number, there is no actual 'middle score'. This problem is overcome by taking the mean of the two middle scores – in this case 65 and 67 – yielding a single number to serve as the median for that data set. The median for Group 1 is, therefore, 66.

The final measure of central tendency is the *mode*, which is the most frequently occurring score in a group of scores. If you were to count the scores in Table 9.2, you would find that Group 1's mode is 65, and Group 2's mode is 60. The mode is usually easy to obtain; however, it can be an unreliable measure, especially in a small group of scores. The mode's advantage is that it reveals the score actually obtained by the greatest number of people.

See Figure 9.15 for an overview of the measures of central tendency.

descriptive statistics

Statistics used to describe, summarise, organise and analyse data

measure of central tendency

A measure of the tendency for a majority of scores to fall in the mid-range of possible values

mean

A measure of central tendency found by adding up all the values and dividing the total by the number of values

median

A measure of central tendency found by arranging scores from the highest to the lowest, and selecting the score that falls in the middle

mode

A measure of central tendency found by selecting the most frequently occurring score in a group of scores

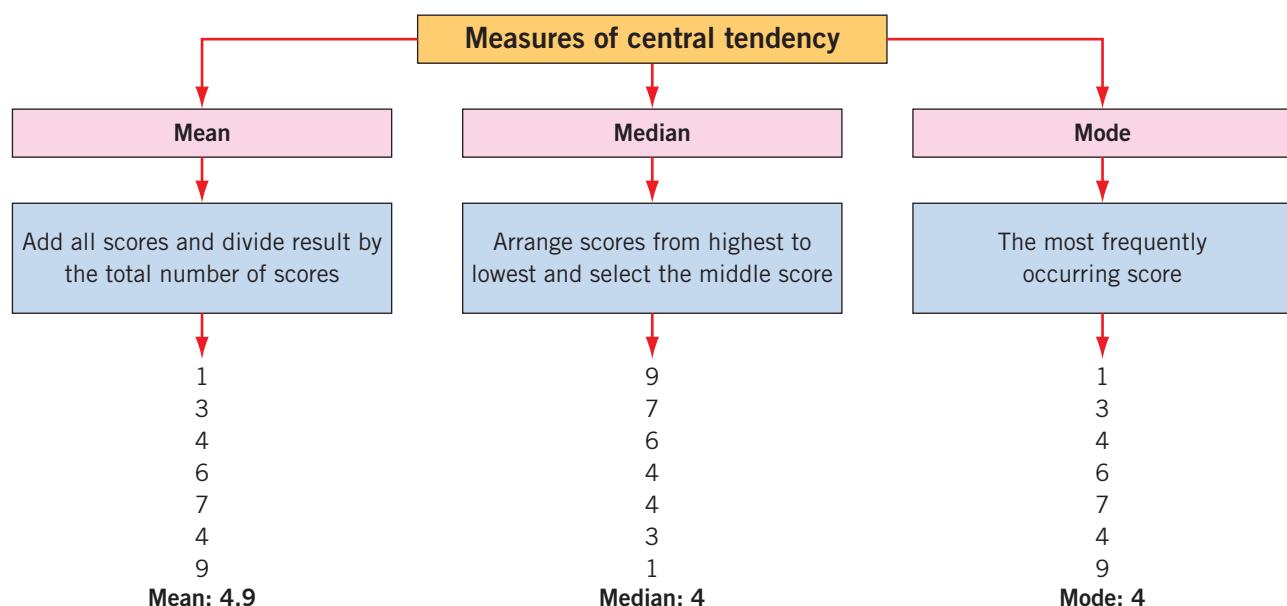


Figure 9.15 A measure of central tendency provides a number that describes a 'typical' score around which other scores fall.

PROBABILITY IN PSYCHOLOGY: P-VALUES AND CONCLUSIONS

Probability refers to the likelihood of an event occurring. In the imaginary drug experiment used as an example in Table 9.2, we found that the average memory score was higher for the group given the drug than it was for the group given the placebo. Certainly, this result is interesting, but might it not have occurred by chance? We can assume that if two groups were tested repeatedly, with neither receiving any drug, then their average memory scores would sometimes differ. But how much would the two means have to differ before we could consider the difference *significant*, or not wholly due to chance?

Tests of *statistical significance* provide an estimate of how often experimental results might have occurred by chance alone. The results of a significance test are stated as a probability, or *p-value*. This probability gives the odds that an observed difference is due to chance alone. In psychology, an experimental result that could have occurred by chance five times (or fewer) out of 100 (in other words, a less than 5% probability, or $p<0.05$) is considered statistically significant (see Figure 9.16).

Some researchers use a more conservative estimate of statistical significance, depending on the research they are conducting. In a conservative experiment (a drug trial, for instance) where it is important that only results due to the IV are reported, the significance levels are set very high (for example, at 0.01 or 0.001). In such cases, a statistically significant result is considered to be one that could have occurred by chance alone one time (or fewer) out of 100 (in other words, an equal or less than 1% probability, or $p<0.01$). In other exploratory studies where researchers are examining general trends, much lower significance levels can be set (for example, at 0.1). In our

imaginary memory experiment, the probability is less than 0.05 ($p<0.05$) that the group means would differ by chance alone. This allows us to conclude with reasonable certainty that the drug actually did improve memory scores.

If you conduct a study with a significance level set at $p<0.05$, then the following will be true:

- If p is less than or equal to 0.05 (for example, 0.04), then the difference between the experimental group's results and the control group's results is said to be statistically significant (that is, not due to chance alone), and likely due to the effect of the IV. The experimental hypothesis is therefore supported.
- If p is greater than 0.05 (for example, 0.08), then the difference between the experimental group's results and the control group's results is said not to be statistically significant (that is, it is likely due to chance alone). The experimental hypothesis is rejected.

Table 9.3 summarises *p*-values.

Table 9.3 Summary of *p*-values

IF p IS LESS THAN 0.05 ($p<0.05$)	IF p IS GREATER THAN 0.05 ($p>0.05$)
The difference between control and experimental groups <i>is</i> statistically significant	The difference between control and experimental groups <i>is not</i> statistically significant
The difference is <i>unlikely</i> to be due to chance alone	The difference between the groups is <i>likely</i> to be due to chance alone
The difference between the groups is <i>likely</i> to be due to the IV	The difference is <i>unlikely</i> to be due to the IV
The experimental hypothesis is <i>supported</i>	The experimental hypothesis is <i>rejected</i>

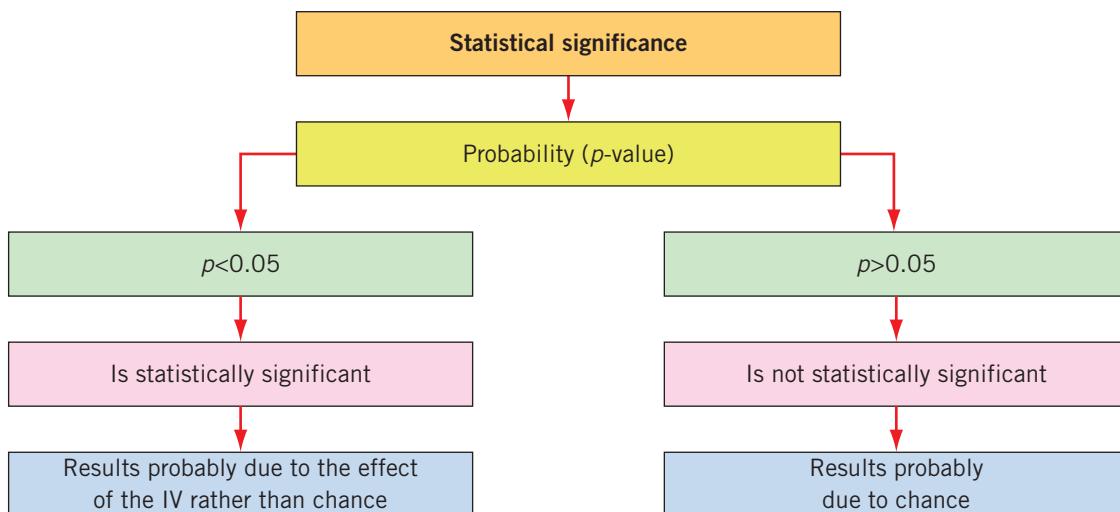


Figure 9.16 Tests of statistical significance provide an estimate of how often experimental results might have occurred by chance alone.

CHECK YOUR UNDERSTANDING 9.7

- 1 Match each term with its definition.
 - a Statistical significance i The score that falls in the middle of a score set when scores are arranged from highest to lowest
 - b Median ii Statistics that are used to describe, summarise, organise and analyse data
 - c Descriptive statistics iii Statistics that are used to interpret results
 - d Inferential statistics iv A statistic that provides an estimate of how often experimental results could have occurred by chance alone
- 2 Indicate whether the following statements are true (T) or false (F).
 - a If scores are placed in order from smallest to largest, the median is defined as the middle score.
 - b The mode calculates the average score in a set of scores.
 - c The mean in a set of scores is the score that occurs most often.
 - d Descriptive statistics describe, summarise and organise data but they do not interpret data.
- 3 Fill in the gaps below with the correct terms.
 - a _____ statistics summarise raw data so it has meaning and is easier to communicate.
 - b _____ statistics are used for decision-making, generalising or drawing conclusions.
- 4 Which of the following statements is incorrect about tests of statistical significance?
 - A Tests of statistical significance describe, organise, summarise and analyse data.
 - B Tests of statistical significance involve probability.
 - C Tests of statistical significance give an estimate of how often research results could have occurred by chance alone.
 - D Tests of statistical significance suggest a result is significant if $p < 0.05$.
- 5 Which of the following p -values would be considered the most statistically significant?
 - A $p < 0.01$
 - B $p < 0.05$
 - C $p < 0.28$
 - D $p < 0.80$

VALIDITY AND RELIABILITY

Before an assessment instrument such as psychological test, questionnaire or rating scale is accepted as a credible measure of behaviour, attitudes or personality characteristics, it must be shown to have **validity** and **reliability** (see Figure 9.17).

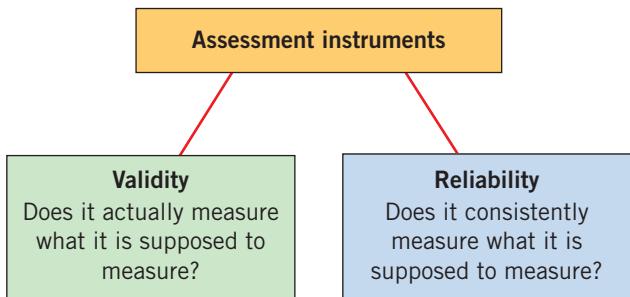


Figure 9.17 To be accepted as a credible measure of behaviour, attitudes or personality characteristics, assessment instruments must be shown to have validity and reliability.

Validity

Validity refers to whether an assessment instrument actually measures what it is supposed to measure. For example, a questionnaire about depression would be expected to measure depression – not anxiety or irritability (that is, the test is *valid* for its purpose).

There are various methods of determining validity:

- **Content validity** – This refers to whether the test measures what it claims to measure by looking at the types of questions it contains. For example, a content-valid intelligence test must include questions that measure an individual's intelligence, not their reading skills or experience sitting intelligence tests.
- **Construct validity** – This refers to whether scores on a questionnaire are consistent with the trait (or construct) being measured. For example, as IQs are supposed to increase with age, an older child will score higher on an IQ test than a younger child if the test has construct validity. Researchers always attempt to produce valid research. This allows them to make accurate conclusions from their results that can be generalised. Sometimes, however, a research finding may be entirely valid in one setting, but not in another. For example, if a new sleep drug improved the quantity of sleep in a sample of 80-year-old residents in a particular nursing home in Mildura, it is not necessarily the case that it will increase the quantity of sleep in *all* 80-year-olds.

statistical significance

A number obtained from inferential statistics that provides an estimate of how often experimental results could have occurred by chance alone; expressed as ' p -value'

validity

The extent to which an assessment instrument actually measures what it is supposed to measure

reliability

The extent to which an assessment instrument consistently measures what it is supposed to measure

content validity

Whether the content (types of questions) of a test measures what it claims to measure

construct validity

Whether scores on a test are consistent with the trait (or construct) being measured

If a test or research finding cannot be applied to a wider group of people, it does not have *external validity*. External validity is the extent to which a study is valid for a range of people and a range of times. If the study described above proved valid for all 80-year-olds in all settings (not just those in the Mildura nursing home), it could be said to have external validity. This is why experimenters must carefully choose their sample group – so that the sample is representative of the wider population and not just a small subset.

Reliability

Reliability refers not only to whether an assessment instrument measures what it is supposed to measure, but also to whether the assessment instrument *consistently* measures what it is supposed to measure. For example, an intelligence test is said to be reliable if the same person takes it twice, with testings one month apart, and scores the same both times (that is, the test is reliable over time).

If a study is based on a test, self-report or rating scale, reliability can be measured in two ways:

- 1 *Test-retest* – The test-retest method occurs when participants are given a test or questionnaire at two different times, with a period of time in between each test. (That is, they are tested and retested.) If the test is reliable, there should be a significant positive correlation between the scores obtained the first and second time the person took the test.
- 2 *Parallel forms* – Parallel forms of the same test can also measure the reliability of an assessment instrument. One way of achieving this is to give the test to the participants and then correlate the odd and even answers. Again, if the test is reliable, there should be a significant positive correlation.

If a study relies on observations rather than self-reports, and if those observations are being made by more than one rater (observer), the study must be tested for *inter-rater reliability*. Inter-rater reliability is high when the raters are consistent in their observations. There are two ways to actually estimate inter-rater reliability:

- 1 If the study consists of categories, the raters check off which category each observation falls in, and the percentage of agreement between the raters can be calculated. For example, consider that 100 observations are being rated by two raters, and for each of the 100 observations each rater can allocate it to one of three categories. If the raters allocated 86 of the 100 observations to the same category, the percentage of agreement would be 86 per cent.
- 2 If the study consists of continuous data, then to establish inter-rater reliability a correlation between the ratings of the observers is calculated. For example, two observers might rate the overall level of activity in a classroom on a 1-to-7 scale at regular time intervals (for example, every 30 seconds). The correlation between the ratings of the two observers would give an estimate of the reliability or consistency between the raters.

For a test to be reliable, it must also have *internal*

consistency. A test is said to have internal consistency if it measures the same variable (such as intelligence) to the same extent every time.

GENERALISATION OF FINDINGS

A researcher studies the effects of a new therapy on a small group of depressed people. Is the researcher interested only in these particular people? Usually not. Except in rare instances, psychologists seek to discover general laws of behaviour that they can apply widely to humans and animals. Undoubtedly, the researcher in this case would like to know if the therapy holds any promise for *all* depressed people.

A *generalisation* is a decision or judgement about whether findings obtained from a sample are representative of the relevant population. The extent to which research findings can be generalised depends on the topic being researched. For example, phenomena where there is little individual variation – such as sensation, attention and memory – can be easily generalised to the population. On the other hand, characteristics where there are large individual differences – such as those of personality, temperament and intelligence – often make it far more difficult for researchers to generalise sample results to the population of research interest.

If research findings can indeed be generalised to the population of research interest, then we would expect that the results would be similar if the study was replicated using a different sample from the same population.

CHECK YOUR UNDERSTANDING 9.8

- 1 Match each term with its definition.

<p>a Reliability</p> <p>b Validity</p> <p>c Conclusion</p> <p>d Generalisation</p>	<p>i A judgement about the meaningfulness of the results</p> <p>ii The extent to which tests consistently measure what they claim to measure</p> <p>iii A judgement about how representative of the population the results are</p> <p>iv The extent to which tests actually measure what they claim to measure</p>
--	--
- 2 Indicate whether the following statements are true (T) or false (F).
 - a Reliability is concerned with whether a test measures what it is supposed to measure on more than one occasion.
 - b Validity refers to whether or not the participant sample is representative of the population of interest.
 - c Internal consistency relates to whether a research study's results are valid.
 - d If results of a study are reliable and valid they can be generalised to the population of interest.
- 3 Fill in the gaps with the correct terms.
 - a A judgement about whether the results of a sample represent the relevant population is called a _____.

- b** A _____ is based on a research study's conclusion.
- c** When a test measures the same variable to the same extent every time, it has _____ consistency.
- d** When a study is valid for a range of people and a range of times it is said to have _____ validity.
- 4** Which of the following statements is incorrect about reliability in a research study?
- A** Reliability refers to whether an assessment instrument consistently measures what it says it measures.
 - B** Reliable tests should have a significant positive correlation.
 - C** Reliability refers to whether an assessment instrument actually measures what it says it measures.
 - D** Reliable tests should have a significant negative correlation.
- 5** When the findings of a study based on the results of the sample are applied to the population from which the sample was taken, _____ has occurred.
- A** random allocation
 - B** stratified allocation
 - C** generalisation
 - D** artificiality

Ethical principles and professional conduct

Ethical principles apply to all types of psychologists. The term *ethics* refers to moral principles and codes of behaviour that must be followed. The Australian Psychological Society (APS) has developed a *Code of Ethics* (1997) – as well as a complementary publication called *Ethical Guidelines* (2002) – that all psychologists must follow when conducting research and treating clients (see Figure 9.18). The code covers rules, regulations and ethics that psychologists are bound to observe, and the principles apply to all psychologists, whether they are conducting research using humans or animals. These guidelines ensure that participants' physical and psychological welfare is protected during and after an experiment is conducted.

external validity

The extent to which a study is valid for a range of people and a range of times, not just a small subset of a population

internal consistency

The degree to which a test measures the same variable to the same extent every time

generalisation

A decision or judgement about whether results obtained from a sample are representative of the relevant population

ethics

Moral principles and codes of behaviour

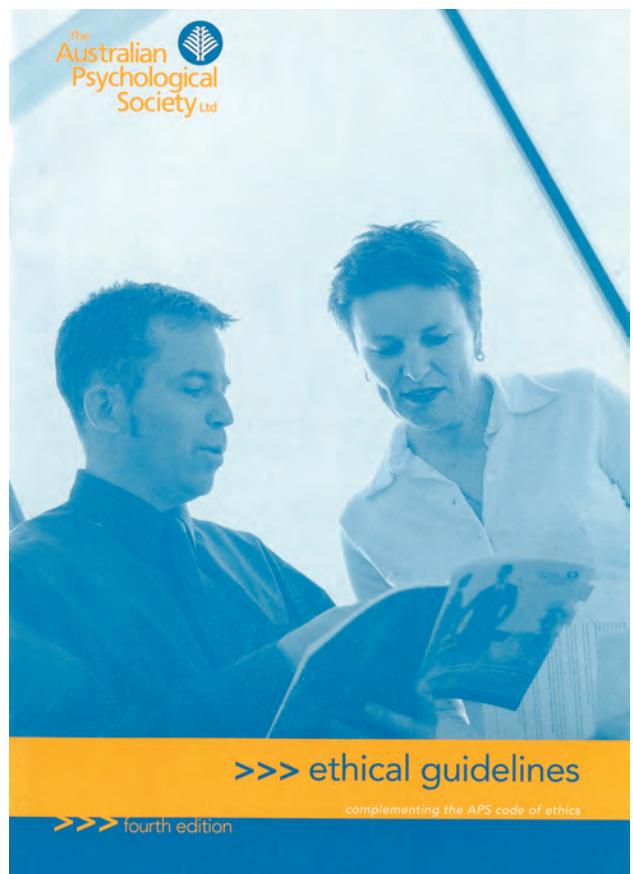
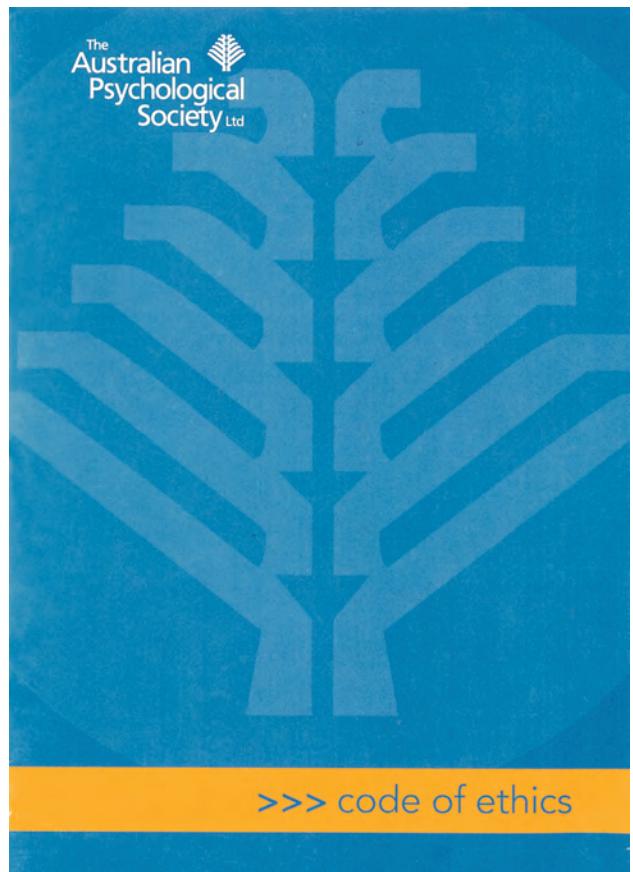


Figure 9.18 All psychologists must follow the Australian Psychological Society's *Code of Ethics* and its complement, *Ethical Guidelines*.

EXPERIMENTER'S ROLE

It is the experimenter's responsibility to protect participants' physical and psychological welfare. Participants' welfare is the most important aspect of a study – even more important than the research findings. At no time must an experiment be conducted that causes severe distress to participants. If a participant does encounter unexpected distress, the experiment must immediately stop and the experimenter must provide the participant with access to follow-up counselling or therapy.

At all times during research or counselling, psychologists must behave in a professional manner. They must not bring disrepute to the profession of psychology or to scientific research. If a psychologist is conducting a study with another professional who is not a psychologist – and who is therefore not bound by the APS's Code of Ethics – the researcher must ensure that the other professional(s) agree to follow the code and any other relevant guidelines prior to commencement of the research.

PARTICIPANTS' RIGHTS

All research in psychology must respect the **participants' rights**; therefore, all relevant ethical guidelines must be adhered to, including not only the APS guidelines but also the Australian National Health and Medical Research Council (NHMRC) guidelines. These guidelines are designed to ensure that participants suffer no physical or psychological harm during or as a result of the experiment. If you are a research participant, you should expect that all guidelines will be adhered to and all your rights upheld. Types of participants' rights are discussed next.

'Focus on research: The Monster Study' demonstrates why participants' rights should be protected at all times.

FOCUS ON RESEARCH

The Monster Study

In 1939, psychologist Wendell Johnson, of the University of Iowa, was interested in how children acquired language. Specifically, Wendell wanted to know whether reinforcement could be used to teach children to stutter.

He used a group of orphaned children of normal ability as his participants. He separated the children into two groups. Group 1 received positive reinforcement, in the form of praise and encouragement, for their speech. Group 2 received negative consequences for their speech as every mistake they made was criticised and they were labelled as *stutterers*. Children who received negative consequences showed negative psychological effects, such as a deflated sense of self, and many were so damaged that they retained speech problems throughout their lives.

The use of orphaned children in this way was thought so shocking that Johnson's results were originally hidden, for

fear of this experiment damaging his reputation. They were, however, eventually released, and in 2001, 62 years after the experiment, the University of Iowa publicly apologised for the study.

Source: *Null Hypothesis – The Journal of Unlikely Science online* (2007) The Monster Study.

QUESTIONS

- 1 Write a research hypothesis for this experiment.
- 2 Identify the IV and the DV.
- 3 Identify the experimental group and the control group.
- 4 Identify three modern ethical principles that were breached by this experiment.

Confidentiality

Confidentiality refers to the participant's right to privacy in terms of access, storage and disposal of information related to the research. A participant's involvement in and results from an experiment cannot be disclosed to anyone else unless written consent has been obtained. If a participant is under the age of 18 years, their parent or guardian must provide this written consent.

Voluntary participation

A participant must decide to participate in an experiment of their own free will; they must demonstrate **voluntary participation**. Participants must not experience any pressure to participate (for example, coercion or bribery) or any negative consequences (for example, threats) if they decide not to participate in the experiment. Also, they must be free to withdraw from the experiment or have their results withdrawn at any time without pressure or negative consequences.

Withdrawal rights

The participant's **withdrawal rights** should be disclosed during the process of obtaining informed consent. A participant is entitled to withdraw from a study at any time, or have their results withdrawn, without experiencing any pressure or negative consequences. A researcher cannot withhold a participant's withdrawal rights, even if their withdrawal is detrimental to the research.

Informed consent procedures

Before a study commences, a researcher must follow certain procedures when recruiting participants – they must obtain **informed consent**. Where appropriate, before an experiment commences, the researcher must fully inform the participants of the true nature and purpose of the experiment, and obtain their written consent to participate. If a participant is under the age of 18 years or is legally unable to give consent, their parent, guardian or power of attorney should complete the informed consent form.

Consent form

Dear participant,

You are invited to participate in a project being conducted by the University of Springfield entitled 'Effects of alcohol consumption on driving ability'. The study will be conducted on Friday 13 October.

The purpose of this study is to determine whether the consumption of alcohol affects the number of errors made while driving, with the view to developing educational programs that will dissuade young people from drink-driving. In this important phase of research, we are examining how alcohol consumption affects the ability to control a motor vehicle.

On the day of information collection, you will be asked to consume a series of drinks that may or may not contain alcohol or caffeine. Your driving ability will then be tested on a driving simulator. If you do not want to consume alcohol or caffeinate drinks, please inform the researcher at the start of the test.

The experiment will take approximately two hours. During this time, your vision, ability to detect colour, auditory processes and reflexes will also be tested.

All results will remain completely confidential. You will be assigned a code so that your name will not be identified with or appear on any records. The results will be presented as group results only, and you will not be identified in any reports resulting from this work.

You are free to withdraw your participation at any time. If you would like to discuss any issues of concern after the experiment, you may contact one of the chief investigators. All questions will be treated confidentially.

We thank you for your assistance.

Project title: Effects of alcohol consumption on driving ability

Participant consent:

I, _____ (participant's name), have read and understood the information above, and any questions I have asked have been answered to my satisfaction. I agree to participate in this research, aware that I may withdraw my consent at any time. I agree that research data collected for the study may be published or provided to other researchers on the condition that my name is not used.

Name of participant: _____

Signature of participant: _____

Date: _____

Figure 9.19 An example of a consent form

participants' rights

The individual rights of all participants that must be respected by the researcher, as outlined in ethical guidelines relating to psychological research

confidentiality

A participant's right to privacy in terms of access, storage and disposal of information related to a research study in which they participated

voluntary participation

Participation whereby participants agree to take part in an experiment free from pressure or fear of negative consequences

withdrawal rights

A participant's right to withdraw from a study or research at any time without experiencing any negative consequences

informed consent

Where a participant gives their written consent to participate in a study after being fully informed of the true nature and purpose of the experiment (where appropriate), any foreseeable risks and their rights before an experiment commences

Researchers must also outline any reasonable foreseeable risks to the participant and inform participants of their rights, including their right to withdraw. The consent form must identify any possible physical or psychological stress that may be encountered during the experiment (see Figure 9.19). The researcher must ensure that any psychologically or physiologically vulnerable person does not participate in the study. Where possible, participants must also be informed about the research procedures employed in the study. In situations where it is considered inappropriate to gain informed consent from participants, the researcher must gain the permission of an ethics committee to withhold information.

Deception

Deception refers to withholding information from the participant about a study's true purpose, before the experiment begins. Deception is used in cases where giving participants information about an experiment beforehand might influence their behaviours during the study and thus affect the accuracy of results. Approval for the use of deception must be given by an ethics committee prior to the commencement of a study.

An example of deception is where a researcher studying guilt led subjects to believe they had broken an expensive piece of machinery. During the experiment, a machine suddenly popped loudly, released a plume of smoke and sputtered to a stop. As the embarrassed participants were about to leave, the experimenter asked them to sign a petition he was circulating. The petition called for doubling tuition fees at the school. Almost all control participants had previously refused to sign the petition; however, because of their guilt after 'breaking' the machine, more than 50 per cent of the experimental participants signed (Rubin, 1970). In this case, if the experimenter had informed participants of the true nature of the study, they would have known the machine was not really broken, and their responses may have differed.

Perhaps guilt could have been studied in some other way; nevertheless, some questions simply cannot be answered without using deception. When this is the case, researchers must deceive participants as little as possible, and offer debriefing after the study.

Debriefing

When deception is used, the researcher must ensure that participants do not suffer any psychological or physical stress as a result of the deception. Therefore, participants must be debriefed. **Debriefing** is where participants are informed of the study's true purpose once the experiment has ended. During debriefing, a researcher must correct any mistaken attitudes or beliefs that have been caused by or that relate to the experiment. They must provide an opportunity for the participants to gain access to information about the study, including procedures, results and conclusions. Participants must also be given information on available services they may want to use if they experience any distress as a result of their participation.



Figure 9.20 Not all people support the use of animals in experiments.

USE OF NON-HUMAN ANIMALS IN RESEARCH

Research using non-human animals can raise difficult ethical questions and debate continues about the advantages and limitations of using animals in research (see Figure 9.20).

Much psychological research is devoted to discovering the natural laws that govern human behaviour. In many instances, animal models are used to discover the principles that apply to humans. Using animals can be a starting point for human research because some animal body systems are very similar to those of humans. Studying animals has added greatly to our understanding of obesity, memory, stress, learning, psychosis, therapy, ageing and many other topics. However, researchers must be cautious when generalising the findings of animal studies to humans. Although the body systems of the two species are similar in some cases, in most cases they are completely different.

All researchers, whether they use human or non-human animal participants, must follow ethical guidelines and adhere to ethical principles. It is also expected that researchers using animal participants will protect their rights and welfare.

ROLE OF ETHICS COMMITTEES

Prior to conducting research, an experimenter must submit a research plan to an ethics committee for approval. The plan must explain the nature and purpose of the study and how society will benefit from it. Ethics committees are usually composed of a range of health-care professionals and, at times, some non-medical members. Ethics committees weigh the potential benefits to society of the proposed research against any foreseeable risks or discomfort to participants.

Ethics committees exist to protect participants' rights and well-being and ensure that the ethical standards relevant to psychology and the integrity of the profession are maintained. They also ensure that all aspects of the proposed research conform to relevant legislation and/or government guidelines. Following the review of a research proposal, the committee can approve the proposal as stated, approve it with changes, or reject the proposal.

Ethics committees also investigate any claims of breaches of standards. Those found guilty may be excluded from membership of professional associations, or they may be suspended or de-registered from practice.

CHECK YOUR UNDERSTANDING 9.9

1 Match each term with its definition.

- | | |
|---------------------------|--|
| a Voluntary participation | i Telling participants of the true aim and nature of the experiment and informing them of their rights before they agree in writing to participate |
| b Debriefing | ii Withholding the true purpose of the experiment from participants |
| c Deception | iii At the end of the experiment, informing participants of the true purpose of the experiment if deception was used, correcting any mistaken beliefs about the experiment, and advising participants of counselling services available if they are distressed |
| d Informed consent | iv Individuals agree to take part in the experiment of their own free will |

2 Indicate whether the following statements are true (T) or false (F).

- a Ethics committees exist to protect the participants welfare and the integrity of the profession of psychology.

b Ethics committees cannot tell a researcher not to carry out research.

c Ethics committees scrutinise proposals of all planned psychological research.

3 Fill in the gaps below with the correct terms.

a Participant consent must be in _____.

b Deception is agreed to before an experiment begins, but _____ occurs after it has finished.

c The ethical principle of _____ guarantees to uphold a participant's privacy rights.

d Informing participants of their right to withdrawal at any time is part of procedures relating to _____.

4 Which of the following statements is incorrect?

A When researchers say they will reimburse participants for any costs incurred as a result of participation in the study, this is not a violation of voluntary participation.

B When researchers say they will reimburse participants for any costs incurred as a result of participation in the study, this is a violation of voluntary participation.

5 To ensure that the ethical principle of voluntary participation is upheld, researchers must:

A inform participants of the nature of the research prior to commencing the research.

B inform participants of all their rights (including confidentiality rights and withdrawal rights) prior to commencing the research.

C not offer participants a bribe or threaten or pressure them to participate in any way.

D All of the above

deception

When information about the true purpose of a study is not given to participants before a study begins

debriefing

Informing participants of the true purpose of an experiment once it has ended; correcting mistaken attitudes or beliefs; providing the opportunity to gain information about the study; providing information about services to help with distress resulting from participation

Chapter summary

WORDCHECK

TEST
YOURSELF

Experimental research:

- The steps in experimental research: a research problem is identified, a hypothesis is formulated, a method is designed, data is collected and analysed, results are interpreted and findings are reported.
- A research hypothesis identifies the IV and the DV, states how these will be measured and predicts the relationship between them.
- Researchers must ensure that all extraneous variables and possible confounding variables are identified and eliminated, or at least controlled, otherwise they will not be able to determine whether the DV was caused by the IV or whether the DV was the result of some unwanted variable.
- Extraneous and confounding variables can be linked to individual participant differences, placebo effect, demand characteristics, the experimenter effect, order effects, artificiality or non-standardised instructions and procedures.
- A number of methods exist for controlling or eliminating unwanted variables, such as type of experiment, counterbalancing, single- and double-blind procedures, placebos, type of sampling procedures, standardised instructions and procedures, and the type of experimental research design chosen.

Sampling procedures and allocation to groups:

- Participants from the population of research interest can be selected for the sample using a variety of methods, including random sampling, stratified sampling, random-stratified sampling and convenience sampling.
- Once selected for the sample, participants are usually randomly allocated to either the experimental group (exposed to the IV) or the control group (not exposed to the IV).

Types of qualitative and quantitative data:

- Data can be collected in a variety of ways including case studies, observational studies, self-reports, questionnaires, interviews or via brain imaging and recording technologies.
- Each type of data has advantages and limitations. For example, qualitative data (data in words) is more detailed, but is sometimes difficult to analyse and interpret, whereas quantitative data (data in numbers) is easy to analyse and interpret, but often does not include a lot of detail.
- Qualitative data involves statistics such as measures of central tendency (mean, median and mode), probability values and their interpretation.
- Conclusions are based on the statistical significance of the results. If the results of the sample are to be generalised to the wider population, then it is crucial that the results are both reliable and valid.

Ethical principles and guidelines for professional conduct:

- Psychologists must follow a code of conduct and ethics when conducting research and treating clients and patients. This ensures that the participants' rights are upheld and protected at all times.
- The ethical principles that psychologists are bound to follow relate to confidentiality, voluntary participation, withdrawal rights, informed consent procedures, use of deception and debriefing.
- Ethics committees monitor the conduct of psychologists and scrutinise all proposed research study's to ensure that all ethical principles and appropriate guidelines are followed.
- When animals are used as research participants, psychologists must treat them in an ethical and humane manner.
- Findings from animal studies have proved invaluable to psychologists in understanding human behaviour but caution must be exercised when generalising the findings of animal studies to human populations.

Apply your knowledge and skills

SECTION A: MULTIPLE-CHOICE QUESTIONS

- In an experimental study, an experimenter effect is said to have occurred if:
 - the researcher's characteristics or expectations influence the results.
 - statistical analysis shows that the researcher's hypothesis is supported.
 - there are no extraneous variables.
 - statistical analysis shows that the researcher's hypothesis is not supported.
- An operational hypothesis is a statement that describes:
 - the size of the study sample.
 - the methods used to address the research questions.
 - how the participants in the study will be recruited.
 - how the study will be statistically analysed.
- The primary reason for using random sampling is to ensure that:
 - the participants are less likely to provide biased responses.
 - there is no experimenter effect.
 - the participants are representative of the population of research interest.
 - there are no confounding variables.

- 4** One purpose of using inferential statistics is to:
- A** identify extraneous variables.
 - B** generalise the results of a study to a sample.
 - C** identify whether an experimenter effect has occurred.
 - D** generalise the results of a study to a population.
- 5** A researcher is conducting a study that requires participants to name the colour of written words as quickly as possible. In the first condition, the colour of the words matches the word that is written; for example, 'red' is written in red. In the second condition, the colour does not match the word written; for example, 'red' is written in blue. The researcher records the number of errors made and the time taken to say the colour of the word. The IV in this study is:
- A** the actual colour of the word.
 - B** the number of errors made.
 - C** the time taken to say the colour of the word.
 - D** whether or not the colour matches the word.
- 6** A researcher wanted to test the effects of air pollution on problem-solving ability. Participants were placed in an unpolluted room and given a set of problems to solve. The time this took was recorded. Participants were then given a different list of problems of equal difficulty and asked to solve them in a polluted room. The time it took to do so was recorded. To what experimental condition were participants exposed in the unpolluted room?
- A** The experimental condition
 - B** The control condition
 - C** The independent condition
 - D** The dependent condition
- 7** A sample is selected to match the distribution of age and religion that occurs in the population. This type of sample is called a:
- A** matched sample.
 - B** stratified sample.
 - C** random sample.
 - D** convenience sample.
- 8** If a researcher wants to give all members of a research population an equal chance of being part of a sample, then they would use a:
- A** random sample.
 - B** stratified sample.
 - C** random-stratified sample.
 - D** convenience sample.
- 9** In an experimental study, a placebo effect is said to have occurred if:
- A** the participant's expectations influence the study results.
 - B** the results for the experimental and control groups are very similar.
- C** the researcher is not aware of which subjects are in the experimental group and which subjects are in the control group.
- D** the researcher's expectations influence the study results.
- 10** In generalising from a sample to the population, it is important that:
- A** the sample is representative.
 - B** the sample has been standardised.
 - C** the sample is not too large.
 - D** All of the above
- 11** What is the mode of the following distribution of scores: 2, 2, 4, 4, 4, 14?
- A** 2
 - B** 5
 - C** 6
 - D** 4
- 12** Which of the following would be the measure of central tendency that would most likely be affected by a few extreme scores?
- A** Mean
 - B** Median
 - C** Mode
 - D** None of the above
- 13** If a difference between two samples is not statistically significant, which of the following can be concluded?
- A** The difference is probably due to the effect of the IV.
 - B** The difference is probably due to the effect of the DV.
 - C** The difference could be due to chance.
 - D** The difference could not be due to chance.
- 14** If a researcher uses deception in their study, they must:
- A** inform participants of this before participants agree to participate.
 - B** allow participants to withdraw at any stage.
 - C** debrief participants at the end of the study.
 - D** debrief participants at the beginning of the study.
- 15** Which of the following is a feature of confidentiality?
- A** participant awareness of the study's purpose.
 - B** no lasting harm resulting from participating in a study.
 - C** no identifying information being revealed about a study participant.
 - D** the right to withdraw from the study at any time.

SECTION B: SHORT-ANSWER QUESTIONS

- 1 A researcher believes that banning smoking at work will result in a reduction in the number of cigarettes smoked each day. She identifies companies planning to ban smoking at work, and surveys the smokers at work before and after the ban. On both occasions, smokers are asked how many cigarettes they smoke each day.
Write an appropriate research hypothesis for this study.
- 2 You are going to conduct a study on the amount of sleep obtained by Year 12 students one week before end-of-year exams. Explain how you might obtain a random sample for your study.
- 3 Identify one advantage and one limitation of the case study method of data collection.
- 4 Identify one similarity and one difference between an extraneous variable and a confounding variable.
- 5 Explain why a random sample is considered to be representative of the population of research interest.
- 6 Identify the main characteristic and purpose of a control group in an experiment.
- 7 Explain why reliability and validity are important to any research study.
- 8 Under what circumstances can the results of a study using a convenience sample be considered legitimate?
- 9 What is the role of an ethics committee?
- 10 Explain the difference between the ethical principles of informed consent and voluntary participation.

SECTION C: EXTENDED-RESPONSE QUESTION

In reference to experimental research, explain how the independent-groups design, the matched-participants design, and the repeated-measures design can be used to minimise the effect of extraneous variables associated with participant characteristics.
This question is worth 10 marks.

SECTION D: ASSESSMENT TASK

Empirical research activity: Investigating people's views on animal experimentation

You (and your psychology class) are to write an ERA on the issue of animals being used in experiments. In order to write the ERA, you will need to collect data about people's views on using animals in experiments.

DIRECTIONS

Do the following as a class:

- 1 Formulate an experimental hypothesis for this study.

- 2 Formulate an operationalised research hypothesis for this study.
- 3 Create an informed consent form for the participants to read and sign.
- 4 Each class member must ask five people about their views on animal experimentation using photocopies of the questionnaire on page 327.
- 5 All class members are to bring the raw data to class. (The ERA will be based on the class data, not on your five participants.)
- 6 Collate and analyse the class data.
- 7 Write up your ERA following the reporting conventions outlined on pages 293–5 of this chapter.

THE QUESTIONNAIRE

HOW TO SCORE THE QUESTIONNAIRE

- Add up answers from questions 1, 2, 5, 6 and 10.
- For questions 3, 4, 7, 8 and 9 you need to reverse the scoring. That is, if the participant scored 1 on these questions, change it to 5; if the participant scored 2 on these questions, change it to 4, and so on. (If you have trouble keeping track of this, mark these 'reverse' questions with * or some other symbol before you begin scoring.)
- Add up the changed scores from questions 3, 4, 7, 8 and 9 and total them with the score you added up for questions 1, 2, 5, 6 and 10 to obtain one final overall score for the questionnaire.

WHAT THE SCORES MEAN

A score of 10–26:

Category 1 – Animals should *never* be used in experiments.

The respondent is against using animals in experiments at any time or in any place. If researchers want to know about drugs for humans, they should test them on humans – not animals.

A score of 27–39:

Category 2 – Animals should *sometimes* be used in experiments.

The respondent believes that there is a time and a place for using animals in experiments – when it is necessary. However, the respondent also feels that unnecessary or very painful experiments should not be conducted.

A score of 40–50:

Category 3 – Animals should *always* be used in experiments.

The respondent believes that animal species are inferior to humans and should be used for all experiments – even painful, unnecessary ones.

Questionnaire

Age: _____

Gender: _____

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Circle your response to each of the following questions, based on the strength of response shown above.

- 1** Animals should be used in all experiments. 1 2 3 4 5
- 2** Animals should only be used in medical research. 1 2 3 4 5
- 3** Animals should not be used to test cosmetics. 1 2 3 4 5
- 4** It is inhumane to use animals in experiments. 1 2 3 4 5
- 5** Many drugs and vaccines could not have been developed without experiments on animals. 1 2 3 4 5
- 6** Only rats and mice should be used in experiments. 1 2 3 4 5
- 7** Monkeys should not be used in experiments. 1 2 3 4 5
- 8** It is cruel to use animals in experiments. 1 2 3 4 5
- 9** Experiments should be done on humans instead of animals. 1 2 3 4 5
- 10** It is OK to inflict pain or harm on animals but not humans. 1 2 3 4 5

For researcher's use only

Score: _____

Category: _____