



CAMBRIDGE
UNIVERSITY PRESS

Psychology

for Cambridge International AS & A Level

COURSEBOOK

Amy Papaconstantinou, Lizzie Gauntlett & Julia Russell



Second edition

Digital Access

 Cambridge Assessment
International Education

Endorsed for full syllabus coverage

› Chapter 1

Research methods

LEARNING INTENTIONS

In this chapter, you will learn how to:

- describe and use the concepts and terminology relating to research methods
- describe, evaluate and apply the six main research methods
- explain and use other concepts relevant to planning and conducting studies and analysing their results
- apply research methods concepts to novel situations
- consider the debate surrounding ethics in psychology

gettyimages[®]
Teera Konakan

Introduction

Psychological phenomena are explored using a 'research process' and the methods used to investigate questions in psychology are called 'research methods'. This chapter will help you to understand how those methods are used by psychologists to find out about human (and animal) cognition, emotions and behaviour.

The chapter is divided into several sections, covering the basic research methods that you need to understand: experiments, self-reports, case studies, observations, correlations and longitudinal studies. Once you have studied these sections, you should be able to apply your knowledge of research methods to novel research situations.

In addition, you will learn about the other features of the research process or 'research methodology'. These features are the steps a psychologist takes in developing and conducting studies to investigate a question or problem. This research process can be thought of as having several stages, the first of which is to develop an aim. Although developing an aim is one of many steps, it is useful to consider it here before moving on to the main part of the chapter.

The aim is the intention of the study, the idea it is trying to test, problem it is intended to solve or the question it is attempting to answer. Consider the idea of different ways to help students to study, perhaps using mind maps or revision apps. Imagine that a psychologist,

Dr Huang, asks a few of her psychology students which method they prefer, and finds that both are quite popular. Dr Huang wants to know which is most effective. This is Dr Huang's aim: *to investigate whether mind maps or revision apps are more effective at helping students to learn*. So, the aim tells you the purpose of the investigation: in this case, an experiment. It is generally expressed in terms of what the study intends to show. An aim can also express the intention to investigate a link or relationship between two variables, such as between the number of computer games a student plays and their final A Level grade: in this case, the study would be a correlation.

This chapter therefore covers the whole of this research process:

- development of an aim
- selection of a research method and the designs, formats or techniques to be used within that research method
- definition, manipulation, measurement, control of variables and variables in hypotheses
- selection of participants
- ethical considerations
- analysis of data, including the drawing of conclusions
- evaluation of research.

GETTING STARTED

Why do psychologists do research?

As students, you may be bombarded with 'facts' about how to improve your learning. Perhaps you have heard of different learning styles, or the benefits of repetition or mind maps to help you to revise. Each of these methods should have been tested to see if they actually work (although many haven't!). The process of research allows scientists such as psychologists to test ideas to discover whether there is evidence to support them. This is how we decide which drugs or therapies work best for mental illnesses, whether different displays or music help to sell products, and how we should organise factories to help workers to be efficient and healthy.

To be trustworthy, research needs to be planned well and conducted effectively. Imagine an investigation comparing a new classroom technique to an old one. If the researcher didn't know how hard the students worked and used the new technique on a lazy class, and the old one on a highly motivated class, this would produce false results. Consider a study into consumer psychology that compared how many goods were sold with and without music playing in the store. If the researcher only played music at the weekends and played no music on weekdays, would you believe the findings of the study?

CONTINUED



Figure 1.1: Where do we focus when we concentrate on a problem?

Watch or imagine someone thinking really hard, perhaps trying to remember a name or work out the answer to a question, where do they focus their eyes? It has been suggested that in such situations, we look upwards and to the left (Figure 1.1). Consider how you might test whether this is true. Would you wait for people to get confused and then look at what they do, or would you give them a puzzle to make them think? How would you decide where they are looking? What would you do to be sure that they aren't just looking around the room for clues? Being able to decide on the answers to questions such as these is the basis of designing experiments in psychology.

1.1 Experiments

An **experiment** is an investigation that allows researchers to look for a cause-and-effect relationship. The researcher investigates the way one variable, called the **independent variable**, is responsible for the effect on another, the **dependent variable**. To test this, the researcher manipulates the independent variable (IV) to produce two or more 'levels' or conditions, such as creating 'bright' or 'dull' lighting or selecting 'early' and 'late' in the day. The effect of these conditions on the dependent variable (DV) is measured. For example, an IV of the brightness of lighting might affect attention, with people being better at paying attention when the light is bright. How well people pay attention would be the DV. If there is a big difference in the DV between the conditions, the researcher would conclude that the IV has caused the difference in the DV, i.e. that the brightness of light affects attention (Figure 1.2).

KEY WORDS

experiment: an investigation that allows researchers to look for a causal relationship; an independent variable is manipulated and is expected to be responsible for changes in the dependent variable

independent variable (IV): the factor under investigation in an experiment that is manipulated to create two or more conditions (levels) and is expected to be responsible for changes in the dependent variable

CONTINUED

dependent variable (DV): the factor in an experiment that is measured and is expected to change under the influence of the independent variable



Figure 1.2: An experiment can investigate whether the brightness of lighting affects how well we concentrate

To be more certain that the difference between the conditions is caused by the IV, the researcher needs to control any other variables that might affect the DV. For example, people might find it harder to be attentive if they have recently eaten, exercised or sat through a very dull class. The effect of such **uncontrolled variables** should therefore be minimised, e.g. by keeping all

factors except the IV the same in each condition (or 'level of the IV').

The levels of the IV being compared may be two or more **experimental conditions** (such as bright and dull artificial lights) or there may be one or more experimental conditions that are compared to a **control condition** (for example, artificial light compared to daylight). The control condition is simply the absence of the experimental variable. For example, in a comparison of the effect of eating chocolate on paying attention, we might compare either the effect of eating one bar or two bars (two experimental conditions) or the effect of eating one bar to no chocolate at all (one experimental and one control condition).

KEY WORDS

uncontrolled variable: a variable that either acts randomly, affecting the DV in all levels of the IV, or systematically, i.e. on one level of the IV (called a confounding variable) so can obscure the effect of the IV, making the results difficult to interpret, the effects of which have not or cannot be limited or eliminated

experimental condition: one or more of the situations in an experiment that represent different levels of the IV and are compared (or compared to a control condition)

control condition: a level of the IV in an experiment from which the IV itself is absent. It is compared to one or more experimental conditions

RESEARCH METHODS

A researcher might conduct an *experiment* to test the effect of the *independent variable* of time of day on the *dependent variable* of happiness of students. They might always use lessons immediately after a break or lunch so the students had always eaten recently since this might affect happiness and act as an *uncontrolled variable*. This would be a comparison between two *experimental conditions*.

ACTIVITY 1.1

Look at the Research Methods box. Can you suggest:

- two different times of day to use as the levels of the *independent variable*
- how the *dependent variable* might be measured
- one other *uncontrolled variable* that it would be important to control?

Experimental design

The way that participants are allocated to different levels of the IV is called the **experimental design**. They may participate in all, or only one, of the levels of the IV.

The three experimental designs are:

- independent measures design
- repeated measures design
- matched pairs design.

KEY WORDS

experimental design: the way in which participants are allocated to levels of the IV

independent measures design: an experimental design in which a different group of participants is used for each level of the IV (condition)

Independent measures design

In an **independent measures design**, a separate group of participants is used for each experimental condition or level of the IV. This means that the data for each level of the IV is 'independent' because it is not related to any other data: it has come from different people. Note that this is a different use of the word 'independent' from that in the 'independent variable'.

If we wanted to know whether seeing people behaving aggressively on television had long-term effects, we could (rather unethically) expose a group of young

people to aggressive television and then wait for them to grow older. However, it would much quicker to compare two groups of adults, one group who had been allowed to watch aggressive TV as children and one group who had not been allowed to. This second example uses an independent measures design.

An independent measures design is good because the participants only encounter the experimental setting once. They are therefore unlikely to notice or respond to clues that might tell them the aims of the experiment (**demand characteristics**). One problem is that there might be individual differences between participants in the groups for each level of the IV that could influence the findings. For example, in a study on the effect of noise on dreams, more people who normally remember their dreams well might end up in the 'no noise' group. If so, it could look as though noise prevented dream recall when in fact it had little effect. This risk can be reduced by the **random allocation** of participants to different conditions. This spreads possible differences between individuals across the levels of the IV. To randomly allocate participants, each person is given a number, and the numbers are then randomly divided into two groups. This can be done by giving each participant a number, putting numbers into a hat and drawing out two sets, or by using a random number generator (e.g. on a computer) to do the same thing. This and other strengths and weaknesses of independent measures designs are presented in Table 1.1.

a study looking at the effects of doodling on learning, we could count the number of words recalled in the same group of people when they did doodle and when they did not.

The main advantage of a repeated measures design is that each person acts as their own baseline. Any differences between participants that could influence their performance and therefore the DV will affect both levels of the IV in the same way. Individual differences are therefore unlikely to bias the findings. Imagine that in our experiment on doodling, one person was generally very quick to learn and another quite slow. In an independent measures design this might cause a problem if they were in different groups, but using a repeated measures design makes the differences between them less important, as both could show an improvement with doodling. Individual differences between participants that could affect the DV are called **participant variables**. These variables, such as age, gender, personality or intelligence could hide or exaggerate differences between levels of the IV, i.e. confound or confuse the results. Such variables can be referred to as '**confounding variables**'. It is therefore important to limit the possible effects of these variables in order to be sure that any change in the DV is due to the IV. A repeated measures design does so by comparing the same people in each condition of the IV. This and other strengths and weaknesses of repeated measures designs are presented in Table 1.1.

KEY WORDS

demand characteristics: features of the experimental situation which give away the aims. They can cause participants to try to change their behaviour, for example to match their beliefs about what is supposed to happen, which reduces the validity of the study

random allocation: a way to reduce the effect of confounding variables such as individual differences. Participants are put in each level of the IV such that each person has an equal chance of being in any condition

KEY WORDS

repeated measures design: an experimental design in which each participant performs in every level of the IV

participant variables: individual differences between participants (such as age, personality and intelligence) that could affect their behaviour in a study that would hide or exaggerate differences between levels of the IV

confounding variable: an uncontrolled variable that acts systematically on one level of the IV so could hide or exaggerate differences between levels and therefore 'confound' or confuse the results making it difficult to understand the effect of the IV on the DV

Repeated measures design

In a **repeated measures design**, the same group of people participate in every level of the IV. To help you to remember, think of the participants 'repeating' their performance under different conditions. For example, in

As each individual participates in every level of the IV, they will perform the same or similar tasks two or more times. This can lead to a problem called an **order effect**. Repeated performance can cause two different kinds of order effects. When participants' performance improves because they have encountered the task before this is a **practice effect**. This matters because participants who were tested on a condition second would perform better than those who did it first. Alternatively, repetition might make performance worse, perhaps if they were bored or tired: a **fatigue effect**. In addition, the participants see both levels of the IV and have more opportunity to work out what is being tested, so are more likely to respond to demand characteristics.

KEY WORDS

order effects: practice and fatigue effects are the consequences of participating in a study more than once, for example in a repeated measures design. They cause changes in performance between conditions that are not due to the IV, so can obscure the effect on the DV

practice effect: a situation where participants' performance improves because they experience the experimental task more than once, for example due to familiarity or learning the task

fatigue effect: a situation where participants' performance declines because they experience the experimental task more than once, e.g. due to physical tiredness or boredom with the task

Order effects can be solved in two ways: by randomisation or counterbalancing. Imagine an experiment with two conditions: learning while listening to music (M) and learning with no music (N). In **randomisation**, participants are randomly allocated to do either condition M followed by N, or vice versa. As some will do each order, any advantage of doing one of the conditions first will probably be evened out in the results. To be more certain that possible effects are evened out, **counterbalancing** can be used. Here, the group of participants is divided into two, so one half will do M followed by N and the other half N followed by M. If on the second test there was a risk of participants accidentally including items learned in the

first test, this would be a problem for exactly half the participants in the 'music' condition, and exactly half in the 'no music' condition. Alternatively, a different design could be used.

Matched pairs design

The problems associated with both independent measures and repeated measures designs are overcome in a **matched pairs design**. For each participant, another person is found who is similar in ways that are important to the experiment, such as age, gender, intelligence or personality, to make a pair. This matching is done on variables relevant to the study, so in a study on the effects of playing a violent computer game, participants might be matched on their existing level of aggression. Identical twins make ideal matched pairs as they are both genetically the same and are likely to have had very similar experiences (Figure 1.3). Different groups of participants are then used for each level of the IV, with one participant from each pair being in each level of the IV. By using different participants in each group order effects are avoided and the matching of participants minimises the influence of individual differences. This and other strengths and weaknesses of matched pair designs are presented in Table 1.1.

KEY WORDS

randomisation: a way to overcome order effects in a repeated measures design. Each participant is allocated to perform in the different levels of the IV in a way that ensures they have an equal chance of participating in the different levels in any order

counterbalancing: a way to overcome order effects in a repeated measures design. Each possible order of levels of the IV is performed by a different sub-group of participants. This can be described as an ABBA design, as half the participants do condition A then B, and half do B then A

matched pairs design: an experimental design in which participants are arranged into pairs. Each pair is similar in ways that are important to the study and one member of each pair performs in a different level of the IV



Figure 1.3: Identical twins are perfect participants for a matched pairs design

Experimental design			
	Independent measures	Repeated measures	Matched pairs
Strengths	Different participants are used in each level of the IV so there are no order effects. Participants see only one level of the IV, reducing the effect of demand characteristics. Random allocation to levels of the IV can reduce the effects of individual differences.	Participant variables are unlikely to distort the effect of the IV, as each participant does all levels. Counterbalancing reduces order effects. Uses fewer participants than repeated measures so this is good when participants are hard to find or if participants are at risk.	Participants see only one level of the IV, reducing the effect of demand characteristics. Participant variables are less likely to distort the effect of the IV than in an independent measures design as individual differences are matched. Different participants are used in each level of the IV so there are no order effects.
Weaknesses	Participant variables can distort results if there are important individual differences between participants in different levels of the IV. More participants are needed than in a repeated measures design so the study may be less ethical if participants are harmed and less effective if there is a small sample because participants are hard to find.	Order effects could distort the results. As participants see the experimental task more than once, they have greater exposure to demand characteristics.	The similarity between pairs is limited by the matching process, so the right matching criteria must be chosen in advance for this to be effective. Availability of matching pairs may be limited, making the sample size small (although some studies conducted on twins use very large numbers of pairs).

Table 1.1: Strengths and weaknesses of experimental designs

RESEARCH METHODS

A child psychologist conducted an experiment to look at the effect of violent computer games (Figure 1.4). There were two experimental conditions (violent and non-violent). The *dependent variable* was the children's subsequent violent behaviour. The *experimental design* chosen was an *independent measures design*, with different children in each of the experimental conditions. If a *repeated measures design* had been used, in which the same children played each type of game, there could be *order effects*. For example, aggression caused by playing the violent game could still affect children in the non-violent game condition if they did this second. If this were the case, the problem could be reduced by using *counterbalancing*.

However, the use of an independent measures design risks *participant variables*, such as the original level of violence of each child, affecting the results. This could be reduced by either using *random allocation* of participants to each condition or by using a *matched pairs design*. In this case, children with similar aggression levels would be put in the different conditions. To avoid *demand characteristics*, the children would ideally be unaware that they are in an experiment, perhaps by telling them that they are in a computer games competition.



Figure 1.4: Are children more violent after they have played a violent computer game than before?

ACTIVITY 1.2

Look at the Research Methods box. Think about the following:

- The study could have had another level of the IV that did not use a computer game but did use a computer, such as looking at non-violent pictures. Would this have been a control condition or another experimental condition?
- An order effect that could arise if a repeated measures design was used for this study is that the children might get fed up with playing computer games by the second condition. Is this a practice effect or a fatigue effect?
- Suggest a participant variable, other than initial level of violence, that could affect the results of this study.

Types of experiment

Laboratory experiments

Many experiments in psychology are conducted in artificial surroundings. Experiments conducted in this way are called **laboratory experiments**; the participants are not in their usual environment for the behaviour they are performing. The artificial situation means that the research can try to limit the effect of variables that might affect the DV. For example, a laboratory experiment on the attention of schoolchildren in high and low light levels could be conducted. It might be investigated by testing the children on a computerised attention task conducted in a psychology room in a university. In this situation, the research could choose to control variables such as noise, that could also affect the children's concentration.

KEY WORDS

laboratory experiment: a research method in which there is an IV, a DV and strict controls. It looks for a causal relationship and is conducted in a setting that is not in the usual environment for the participants with regard to the behaviour they are performing

controls: ways to keep potential confounding variables constant, for example between levels of the IV, to ensure measured differences in the DV are likely to be due to the IV, raising validity

CONTINUED

standardisation: keeping the procedure for each participant in a study (e.g. an experiment or interview) exactly the same to ensure that any differences between participants or conditions are due to the variables under investigation rather than differences in the way they were treated

Evaluating laboratory experiments

Laboratory experiments have a range of strengths and weaknesses, these are summarised in Table 1.2. Experiments in general, and laboratory experiments in particular, use many **controls**. In addition, researchers in laboratory experiments can use **standardisation**, which means that the procedure for each participant can be kept exactly the same. Both controls and standardisation help to make the findings of the experiment **reliable**, that is, the researchers would be more certain that the procedures and measures they are using are consistent. Controlling variables also improves **validity**, that is how certain the researcher can be that they are testing what they claim to be testing. By keeping the situation the same, the researcher can be more certain that any differences in the DV really are due to the differences between levels of the IV rather than due to any uncontrolled variables.

One way to identify uncontrolled variables is by conducting a **pilot study**. This is a preliminary test of the procedure and can lead to changes in the procedure or materials to improve the validity and reliability. For example, in the experiment of children's attention (previously), it might be found that hunger also affects attention. A control could then be introduced to give all the children a snack before the study begins. A pilot study should not be used to test whether a study is ethical; this is the responsibility of the researcher and ethical committees.

KEY WORDS

reliability: the extent to which a procedure, task or measure is consistent, for example, that it would produce the same results with the same people on each occasion

validity: the extent to which the researcher is testing what they claim to be testing

CONTINUED

pilot study: a small-scale test of the procedure of a study before the main study is conducted. It aims to ensure that the procedure and materials are valid and reliable, so that they can be adapted if not

Another strength of laboratory experiments is that they are relatively easy for a researcher to repeat using exactly the same procedure and materials. This is called **replication**. The level of standardisation and controls in laboratory experiments help with replicability, as does the **operationalisation** of the independent and dependent variables, that is the clear and precise way that they are defined. It is important to be able to replicate studies so that they can be checked by other researchers and so that other comparisons can be made. For example, if a study shows that a new drug is better than no drug for treating a mental health condition, other researchers might want to verify this, to check it against a **placebo**—a pill or procedure that appears to be a treatment but in fact has no active 'ingredient'—or against other drugs.

As participants are usually aware that they are involved in a laboratory experiment they can be told what to expect to happen and can be given a choice about whether to participate. They can also be given a full explanation at the end of the study. This may be necessary as experimenters sometimes need to actively lie to or mislead participants to reduce the risk of demand characteristics affecting participants.

KEY WORDS

replication: keeping the procedure and materials exactly the same between studies when attempting, for example, to verify results or to enable other studies to use exactly the same techniques to answer related questions

operational: the clear description of a variable such that it can be accurately manipulated, measured or quantified, and the study can be replicated. This includes the way that the IV and DV in experiments and the co-variables in correlations are described

placebo: a pill or procedure given to a patient who believes it to be a real treatment that in fact has no active 'ingredient', i.e. no active drug in the case of a pill or no therapeutic value in the case of an intervention



RESEARCH METHODS

Dr Singh is conducting a laboratory experiment to test whether older adults detect emotions as quickly as younger people. He has two groups of participants, older and younger ones. This is his IV. He tests them by comparing how quickly they press a button to say that they have recognised the emotion on a face of a screen. This is the DV. Each participant sits at the same distance from the screen. This is one aspect of the *standardisation* of the procedure. In a *pilot study*, Dr Singh had shown the participants pictures and used a stopwatch to time their reactions himself, but he found he was not very consistent in his ability to stop timing exactly when the participant responded. He therefore changed to the computerised system to improve *reliability*. The pictures of faces included both younger and older people to ensure it was a *valid* test. The controls and standardisation mean that it would be possible for Dr Singh's colleague, Dr McArthur, to *replicate* the study exactly.

ACTIVITY 1.3

Look at the Research Methods box.

- Write an operational definition for the independent variable.
- Write an operational definition for the dependent variable.
- Name and explain the experimental design being used in this study.
- Suggest why the use of faces of a range of ages would have improved validity.

Field experiments

Returning to the idea at the beginning of this section, about the effect of light levels on attention, the school children could alternatively be tested by altering the number of lights turned on in their normal classroom. Light level would still be the IV and the levels of the IV could be 'all the lights on' and 'half the lights on'. The DV of attention could then be measured by looking at the children's scores on a topic test they were due to take that day. This is still an experiment because it has an IV and a DV (and there will still be some controls, such as the amount of time they spend studying for the test). However,

it would be a **field experiment** because the children are being tested on a usual behaviour (performance in a topic test) in their normal environment (the classroom).

Evaluating field experiments

Field experiments, like laboratory experiments, have a range of strengths and weaknesses, these are summarised in Table 1.2. It is a little harder to control variables and standardise procedures in a field experiment than a laboratory experiment, although some controls can still be used. Reliability and validity may therefore be lower. However, validity may benefit as the participants are performing a task that seems normal, in a familiar environment. School students taken into a university laboratory might concentrate really hard because they are nervous or interested, which might cover up any differences between the different light level conditions. This means the findings from the laboratory would not **generalise** to other settings as well as findings from the classroom. This is a problem of **ecological validity**, and field experiments often have better ecological validity than laboratory experiments (but not always).

Another advantage arises if the participants are unaware that they are in an experiment. If so, there may be fewer demand characteristics than there would be in a laboratory experiment. These are any features of the experiment that give away the aims and cause participants' behaviour to change, for example to try to 'make the experiment work'. However, if participants are unaware that they have been in a study they have not been given a choice about participating and may not have wanted to.

KEY WORDS

field experiment: an investigation looking for a causal relationship in which an independent variable is manipulated and is expected to be responsible for changes in the dependent variable. It is conducted in the normal environment for the participants for the behaviour being investigated and some control of variables is possible

generalise: to apply the findings of a study more widely, e.g. to other settings and populations

ecological validity: the extent to which the findings of research in one situation would generalise to other situations. This is influenced by whether the situation (e.g. a laboratory) represents the real world effectively and whether the task is relevant to real life (has mundane realism)

	Type of experiment	
	Laboratory experiment	Field experiment
Strengths	Good control of variables, raising validity. Causal relationships can be determined as only the IV should be affecting the DV. Standardised procedures raise reliability and allow replication.	As participants are in their normal situation for the activity being studied, they are likely to behave naturally, making the results representative. If participants are unaware that they are in a study, the problem of demand characteristics is less than in laboratory experiments.
Weaknesses	The artificial situation could make participants' behaviour unrepresentative, lowering ecological validity. Participants could respond to demand characteristics and alter their behaviour.	Control of variables is harder than in laboratory experiments, lowering reliability and making replication difficult. The researcher will be less sure that changes in the DV have been caused by changes in the IV than in a laboratory experiment. Participants may be unaware that they are in a study, raising ethical issues.

Table 1.2: Strengths and weaknesses of experimental methods

RESEARCH METHODS

A research team is deciding how to test the effect of watching television on children's pro-social behaviour, that is, how kind children are to others. They will measure pro-social behaviour by observing how often the children help a doll that cries. They are considering two experimental methods, both using an independent measures design. One is a *laboratory experiment*, in which each child is either shown a cartoon in which the characters are very helpful or a cartoon with characters who behave in a neutral way. Each child is then given a doll for 10 minutes and observed to count how many times they help the crying doll. Alternatively, the children could be observed in a *field experiment*. In this study, half of the children's parents would show their child the pro-social cartoon on the family

television. The other half would show their child the neutral cartoon. Each child would be videoed with doll for 10 minutes and the researchers would view the video to count how many times the child helped the crying doll. The field experiment would have more *ecological validity* than the laboratory experiment because the children would be in a familiar environment, in which they would normally play, so would be less likely to be nervous than if they were in a laboratory. For this reason, the results of this study may *generalise* to other situations, such as performing pro-social behaviours in school or at the park. However, a disadvantage of the field experiment, is that there may be distractions in the home that prevent them from paying attention to cartoon.

ACTIVITY 1.4

Look at the Research Methods box. Which of the following can you identify?

- Independent variable
- Dependent variable.

Does the field experiment have a *control condition*?

Can you suggest a *variable* that it would be important to *control*?

What effect might *demand characteristics* have in this study?

What other details would you need to be able to *replicate* the laboratory experiment?

feels unconsciously positive or negative—both would fit with his explanations.

The main hypothesis in a study (sometimes called the **alternative hypothesis**) can be written in several different ways. They differ in terms of the nature of the prediction they make about the results of an investigation.

KEY WORDS

hypothesis (plural hypotheses): a testable statement based on the aims of an investigation

alternative hypothesis: the testable statement predicting a difference between levels of the independent variable in an experiment (or a relationship between variables in a correlation)

Hypotheses in experimental studies

At the very beginning of this chapter, we considered a study that Dr Huang was conducting, with the aim to investigate whether mind maps or revision apps are more effective at helping students to learn (Figure 1.5). Dr Huang would also need to write a hypothesis based on this aim. A **hypothesis** is a testable statement. It provides a little more detail about the variables being investigated than the aim.

Importantly, a hypothesis should be falsifiable, that is it should be possible for it to be shown to be wrong. This is essential, otherwise an alternative hypothesis would always be accepted! At A2 you may learn about the psychodynamic approach and the work of Sigmund Freud, a famous early psychologist. Aspects of his work was contentious, partly because some of his ideas were non-falsifiable, such as his ideas about our behaviour being the consequence of unconscious motives. It has been argued that one problem with unconscious motives is that they are impossible to test in an objective way. Here is a basic example: a young woman is slightly rude to a young man. Freud could use two different explanations. One is that the young woman unconsciously hates the young man, and she is avoiding being as rude as she unconsciously wants to be. The other is that the young woman unconsciously loves the young man, and she is avoiding showing her unconscious desire so is behaving in a contrary way. The problem here is that Freud cannot be wrong—either she

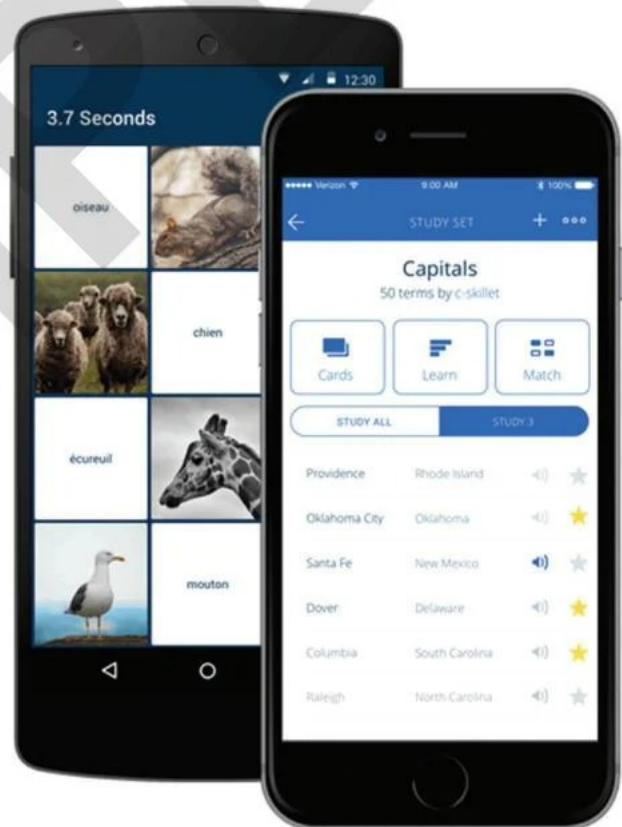


Figure 1.5: Revision apps such as Quizlet and Gojimo may, or may not, help students to learn

Hypotheses should have operationalised variables. For example, in the hypothesis 'Students using revision apps will learn better than students using mind maps',

we do not know how ‘better learning’ will be measured, or which apps are being used because the variables are not operationalised. To be complete, the hypothesis needs more detail, such as ‘Students using the Gojimo revision app will gain better test marks than students using mind maps’.

Non-directional hypotheses

A **non-directional (two-tailed) hypothesis** predicts that there will be an effect, but not the direction of that effect (Figure 1.6). In an experiment, this means that the hypothesis will suggest that the IV will change the DV but not whether the effect will be an increase or a decrease. This type of hypothesis is chosen if the effect of the variable is being tested for the first time, so there are no previous results to suggest what the results might be. For example, Dr Huang’s hypothesis could be: *There is a difference between the effectiveness of mind maps and revision apps in helping students to learn*. Note that it is predicting a difference, but not which condition will be better at helping with learning.



Figure 1.6: Unlike with a one-tailed hypothofish, you can’t see which way a two-tailed hypothofish will swim

Directional hypotheses

When most previous research or other evidence suggests the nature or ‘direction’ of an effect we can use a **directional (one-tailed) hypothesis**. In an experiment, this means saying which condition will be ‘best’ i.e. produce the ‘highest’ scores (and in a correlational study, whether there will be a positive or negative correlation).

Returning to Dr Huang’s study, there might be evidence that revision apps are better than mind maps, perhaps because they are more ‘active’ and being actively engaged helps memory. This is a directional prediction so the hypothesis might be: *Students using revision apps will learn better than students using mind maps*. Note that the opposite prediction could also be expressed as a directional hypothesis. This would be: *Students using mind maps will learn better than students using revision apps*. We might make this prediction if we believed that writing a mind map yourself was more effective than just re-using ready-made materials on revision apps.

Null hypotheses

The alternative hypothesis is an alternative to the **null hypothesis**. In an experiment, the null hypothesis states that any difference in the DV between levels of the IV is so small that it is likely to have arisen by chance. The concept of chance becomes important in ‘inferential statistics’. These are mathematical processes used to find out how likely it is that the pattern of results (a difference or a correlation) found from the sample of participants could have arisen by chance. This information helps us to *infer* whether the findings would apply to the wider population. When you read about the core studies, the most important results are the ‘significant’ ones. This word has a particular meaning, that there is a mathematically significant probability that the pattern in the results could not have arisen by chance. So, when inferential statistics are used, a ‘significant’ pattern enables the researchers to reject their null hypothesis and accept their alternative hypothesis. If, on the other hand, the pattern in the results is non-significant, the researchers must accept their null hypothesis: that the difference or correlation could have arisen by chance.

There are two ways to write a null hypothesis for an experiment. In the case of Dr Huang’s study, the null hypothesis could be written either as: *There will be no difference between the effectiveness of mind maps and revision apps in helping students to learn* or *Any difference in effectiveness of mind maps and revision apps in helping students to learn is due to chance*.

KEY WORDS

non-directional (two-tailed) hypothesis: a statement predicting only that one variable will be related to another, for example that there will be a difference in the DV between levels of the IV in an experiment (or that there will be a relationship between the measured variables in a correlation)

directional (one-tailed) hypothesis: a statement predicting the direction of a relationship between variables, for example in an experiment whether the levels of the IV will produce an increase or a decrease in the DV (or in a correlation whether an increase in one variable will be linked to an increase or a decrease in another variable)

null hypothesis: a testable statement saying that any difference or correlation in the results is due to chance, that is, that no pattern in the results has arisen because of the variables being studied

Ethics in experiments

The role of ethics in psychology is discussed in Section 1.10. Here we will briefly consider ethics in experiments (Figure 1.7). A participant in a laboratory experiment is likely to know that they are participating in a study and can readily be asked for their **informed consent**. However, for the sake of validity, it may be necessary to deceive them to avoid them working out the aim of the study and altering their behaviour, i.e. to reduce demand characteristics. There is therefore a need to balance good ethics and good science, such as high validity. In field experiments, in contrast, it may not be possible to gain consent as the participants may be unaware that they are even in a study. This is an ethical problem because participants should have the right to know what they are entering into and to agree to participate or not. They should also have the **right to withdraw**, which they cannot do if they do not even know that they are in a study. In order to achieve a balance between good science and good ethics, it is therefore very important that participants in a field experiment are effectively **protected from harm**.

In experiments, whether they are in a laboratory or the field, experimenters may sometimes need to actively lie to or mislead participants. This is called **deception** and may be necessary to reduce the risk of demand characteristics affecting participants. Deception should be avoided and, where possible, the reality of the study should be explained to the participants afterwards in a debrief. This is often much harder to do in a field experiment than a laboratory experiment and it may even be impossible. The use of deception should, however, be avoided if possible.



Figure 1.7: Researchers must achieve a balance between good ethics and good science

In all experiments, privacy and confidentiality are important. **Privacy** can be respected in laboratory experiments because the tests or questions used are pre-planned. In the natural settings of a field experiment, however, there is a risk of invading a participant's personal space, or perhaps discovering something that they wouldn't want the researcher to know so researchers must be careful of this. **Confidentiality** can be respected in all experiments by keeping the participants' data secure and anonymous, although if the participants are unaware that data has been collected, as in a field experiment, it is important to ensure that they cannot be individually identified, for example by their place of work.

KEY WORDS

informed consent: knowing enough about a study to decide whether you want to agree to participate

right to withdraw: a participant should know they can remove themselves, and their data, from a study at any time.

protection from harm: participants should not be exposed to any greater physical or psychological risk than they would expect in their day-to-day life

deception: participants should not be deliberately misinformed (lied to) about the aim or procedure of the study. If this is unavoidable, the study should be planned to minimise the risk of distress, and participants should be thoroughly debriefed. It may be done to reduce the effects of demand characteristics but should be avoided

privacy: participants' emotions and physical space should not be invaded, for example they should not be observed in situations or places where they would not expect to be seen

confidentiality: participants' results and personal information should be kept safely and not released to anyone outside the study

RESEARCH METHODS

A psychology department ethical committee is looking at a research proposal for a study about the effect of cognitions on a therapy designed to help people to relax. The researchers only plan to ask for *consent* about part of procedure they will use—listening to an imagery-based relaxation recording—and not their full aim. Although they intend to tell all the participants they will have pulse monitors and be in darkened rooms, they also intend to use deception. They will *deceive* the participants about which level of the independent variable they will be allocated to. Rather than telling them what will really happen—their pulse rate should fall—the researchers will give them false information by telling them that some people see disturbing flashing lights. When the participants are given the limited information at the start of the study they will also be told that they can leave at any time, thus giving them the *right to withdraw*. The instructions on the recording tell the participants to imagine relaxing, intimate thoughts. The participants will also be reassured that they will not be asked about these thoughts, which ensures their *privacy* is protected. When the participants join the study, each will be given a number, which will be used to identify their data so that their names do not have to be used, ensuring their *confidentiality*.

ACTIVITY 1.5

Look at the Research Methods box. Which of the following can you identify?

- The type of *experiment* being planned
- The independent variable
- The dependent variable
- The experimental design

Can you suggest one way in which participants would not be *protected from harm* in this study?

Suggest why participants may want to withdraw from the study.

Why might it be necessary for the researchers to *deceive* the participants?

What is wrong with the hypothesis 'an imagery-based recording will help people to relax'?

Applying your knowledge of experiments to novel research situations

You should be able to recognise experiments (including the IV and DV, and be able to operationalise them, i.e. define them in detail) to decide whether an experiment is a laboratory or field experiment and to evaluate them, e.g. in terms of controls, standardisation, ethics and reliability and validity.

In addition, you should be able to plan an experiment; choosing the type of experiment (laboratory or field), deciding on and operationalising an IV and a DV, choosing an experimental design and implementing suitable controls.

Questions

Barry and Anouk are deciding how to test whether gender affects artistic ability. Barry suggests doing a study in the psychology department where they ask students to come in for a study about memory in which they must redraw a complicated image. Barry and Anouk can then see how well they do it. Anouk thinks it would be better to persuade the art teacher to use an art class and set a lesson where students have to copy the same complicated image.

- 1 Explain the type of experiment that is being suggested:
 - a by Barry
 - b by Anouk.
- 2 The independent variable is the same in Barry and Anouk's studies, as is the dependent variable.
 - a Describe the independent variable (IV).
 - b Describe the dependent variable (DV).
- 3 Explain **one** ethical issue that is clear from the procedure they have suggested.
- 4 Suggest **one other** ethical issue and how they could avoid problems with this issue.

1.2 Self-reports

In a **self-report**, the participant gives the researcher information about themselves directly. This is different from experimental tests or observations where the researcher finds the data out from the participant. There are two main ways to conduct a self-report: using a questionnaire or an interview. Both allow the researcher to ask the participant questions.

Questionnaires

In a **questionnaire**, the questions are presented to the participant in written form. The two techniques for doing this are a 'paper and pencil' exercise or online. There are several different types of question. The two most important question formats are **closed questions**, which have a fixed set of possible responses, and **open questions**, which ask for descriptive answers in the participant's own words. Closed questions can take the form of simple choices, such as those asking for yes/no answers or items from a list. Other forms of closed questions include rating scales (where a number is chosen, e.g. between 0 and 5) and Likert scales, which ask the respondent to say how much they agree with a statement such as 'Obesity is not important' or 'Exercise is a necessity' using the choices 'strongly agree / agree / neither agree nor disagree / disagree / strongly disagree'. Some examples of closed questions are as follows:

- What is your gender: boy or girl?
- How do you travel to school? walk / bicycle / bus / train / car
- Indicate which animal(s) scare you: dog, spider, cat, rat, fish, rabbit, bird. [You may tick as many as you like]
- How much do you like psychology on a scale of 0–4? (0 = not at all, 4 = very much)

KEY WORDS

self-report: a research method, such as a questionnaire or interview, which obtains data by asking participants to provide information about themselves

questionnaire: a self-report research method that uses written questions through a 'paper and pencil' or online technique

CONTINUED

closed questions: a question format in questionnaires, interviews or test items that produces quantitative data. They have only a few, stated alternative responses and no opportunity to expand on answers

open questions prompt the respondent to give detailed answers, which may be quite long. They contain more depth than the answers to closed questions and are more likely to be able to explore the reasons behind behaviours, emotions or reasoning. They typically ask 'Why...?' or simply 'Describe...'. Some examples of open questions are as follows:

- Why do you believe it is important to help people who suffer from phobias?
- Describe your views on the use of social media sites with regard to encouraging helping behaviour.
- Explain how you would respond if you were told to hurt another person.

KEY WORDS

inter-rater reliability: the extent to which two researchers interpreting qualitative responses in a questionnaire (or interview) will produce the same records from the same raw data

generalisability: how widely findings apply, e.g. to other settings and populations

Evaluating questionnaires

Questionnaires, which often use mainly closed questions, are fairly easy to analyse as they produce a total of each category of answers making it simple to summarise the findings. It is also possible to work out averages, which can help to describe the patterns in the results. Where qualitative data is gathered from questionnaires, it produces more detailed, in-depth information. This is an advantage, although it also leads to a problem. Answers to open questions have to be interpreted, and this can lead to a lack of reliability as the researcher may not be consistent in their interpretation. If more than one researcher is involved, there may also be differences between them. This would be a lack of **inter-rater reliability**.

Another problem with questionnaires is that it is easy for participants to ignore them, which means the return rate may be very low. Importantly, the people who do reply to a questionnaire may all be quite similar, for example have time to spend because they are unemployed or retired. This would mean all the people who filled out the questionnaire would have shared characteristics making applying the findings to other people less appropriate, i.e. could result in poor **generalisability**.

Another problem with questionnaires is that participants may lie. They may do this because they want to look more acceptable; this is called a **social desirability bias** (Figure 1.8). Participants may also lie if they believe they have worked out the aim of the study. To avoid this, researchers sometimes include **filler questions** among the real questions. The answers to filler questions are not analysed in the research since they serve only to hide the real purpose of the study.



Figure 1.8: People may lie in questionnaires, lowering validity, for example giving socially desirable responses to questionnaires about eating habits

KEY WORDS

social desirability bias: trying to present oneself in the best light by determining what a task requires

filler questions: items put into a questionnaire, interview or test to disguise the aim of the study by hiding the important questions among irrelevant ones so that participants are less likely to work out the aims and then alter their behaviour

RESEARCH METHODS

Dr Huang is a psychology teacher. She wanted to know how her students were progressing on the course. She decided to use the *self-report* method and used an online *questionnaire* that the students did in their free time to collect data. This included several *closed questions* (see questions 1–4 next), which collected quantitative data, and some *open questions* (question 5 and 6), which collected qualitative data. She asked her colleague to help her to interpret the responses to the open questions and to help to ensure that they had good *inter-rater reliability*, she devised a list to help them to interpret questions 5 and 6. For question 5, it included looking for comments about:

- reading up notes
- copying out notes
- reading the textbook
- looking things up online
- asking friends
- checking with the teacher.

For question 6, it included looking for comments about:

- copying out notes
- making summary notes
- making mind maps
- using past paper questions
- making test cards.

CONTINUED

Some of the questions on the questionnaire were:

- 1 How often do you do the homework set?
always sometimes never
- 2 Have you written yourself a research methods glossary?
yes no
- 3 'Psychology is a difficult subject'. Do you:
Strongly agree agree neither agree nor disagree disagree strongly disagree
- 4 Rate from 0 to 6 how well you understand the topic we have just completed:
0 = don't understand at all
6 = completely understand
- 5 Explain what you do after each lesson to help you to remember what you have learned.
- 6 Describe how you will plan your revision for the next test.

ACTIVITY 1.6

Look at the Research Methods box.

- Explain the difference between the *open* and *closed* questions.
- Suggest one more open question.
- Suggest one more closed question.
- Suggest why Dr Huang may have chosen to use an *online* questionnaire rather than one the students did on paper in the classroom.
- Explain why it was important that Dr Huang took steps to raise *inter-rater reliability*.

Interviews

An **interview** is a research method in which the researcher can use two alternative formats. They are typically face-to-face with the participant but the interview can also happen via the telephone. Interviews can, however, be conducted in other ways that allow real-time interaction, such as through a chat facility. The same formats of

questions can be asked in interviews as in questionnaires, although often interviews use more open questions than closed questions.

The schedule of questions, that is the range of questions that are asked and the order of them, differs between different interview formats. In a **structured interview**, the questions asked are the same for every participant and the order is fixed. There may even be instructions for the interviewer about their tone of voice, how to sit or how to dress so that the procedure is standardised each time data is collected. In an **unstructured interview**, in contrast, the questions asked depend on what the participant says, so the questions may be different for each participant. This is a very flexible format but it may be hard to compare data collected from different participants or by different researchers. A compromise is a **semi-structured interview**. Here, there are some fixed questions, which make sure that there is some similar information from every participant. This means that comparisons can be made between them, and averages can be calculated if this is appropriate. In addition, it is possible to ask some questions that are specific to individual participants. This allows the researcher to develop ideas and explore issues that are particular to that person.

KEY WORDS

interview: a research method using verbal questions asked directly, using techniques such as face-to-face or telephone

structured interview: an interview format using questions in a fixed order that may be scripted. Consistency might also be required for the interviewer's posture, voice, etc. so they are standardised

unstructured interview: an interview format in which most questions (after the first one) depend on the respondent's answers. A list of topics may be given to the interviewer

semi-structured interview: an interview format using a fixed list of open and closed questions. The interviewer can add more questions if necessary

Evaluating interviews

As with questionnaires, interviewees may lie either because they want to seem more acceptable (a social desirability bias) or because they think they know the aim of the study and are either trying to help the researcher by giving the answers they need, or to disrupt the research by doing the opposite. Interviewing is often time consuming and this can be a problem if it restricts the types of participant who volunteer for the research because it would give a narrow representation of feelings, beliefs or experiences.

When interpreting participants' responses to questions in an interview, researchers must be careful not to be **subjective**, that is, to produce findings which are based on their personal perspective. Instead, they should aim for **objectivity**, i.e. taking a view that is not led by their own feelings or beliefs. To achieve this, the interviewer may ask other researchers, who are experienced but unaware of the aims of their research, to interpret the findings.

KEY WORDS

subjectivity: the effect of an individual's personal viewpoint on, for example, how they interpret data. Interpretation can differ between individual researchers as a viewpoint may be biased by one's feelings, beliefs or experiences, so is not independent of the situation

objectivity: the impact of an unbiased external viewpoint on, for example, how data is interpreted. Interpretation is not affected by an individual's feelings, beliefs or experiences, so should be consistent between different researchers

Applying your knowledge of self-reports to novel research situations

You should be able to recognise self-report studies and decide whether they use questionnaires or interviews as a method. You should also be able to choose which of these to use in a new situation. In addition, you should be able to recognise questions using different formats (open and closed) and to identify and evaluate different interview formats (structured, semi-structured and unstructured) and techniques (telephone or

face-to-face). When doing this, it is important to consider how the method used affects the availability of different types of participant and their honesty, as this affects the validity of the findings.

You should also think about the kinds of data that are produced, and the way the data will be used. Although numerical data from closed questions can be analysed mathematically, data from open questions provides more in-depth information that may be more valid. For example, a closed question might not have a response close to a person's view, whereas an open question would allow that person to express views that they could not do in the choices available in the closed question.

The reliability of self-report data is also important. Questionnaires and structured interviews may be higher in reliability because they are likely to be administered in a consistent way and because they generate numerical results that do not need interpretation. Responses to open questions, in contrast, have to be interpreted by the researcher and since the researchers may differ in their opinions there is the possibility that they will be subjective.

Finally, you should be able to plan a study using a self-report and be able to write questions in different formats (open or closed) for either research method (a questionnaire or an interview). For a questionnaire, you should be able to describe how to plan a questionnaire with a pencil and paper or online format. For an interview, you need to be able to describe how to conduct an interview using each of the formats (structured, semi-structured and unstructured) and either of the techniques (telephone or face-to-face).

RESEARCH METHODS

Dr Singh is planning an *interview*-based study because he wants to confirm that a new shopping centre layout is making people more helpful to each other. He wants to collect objective data about the number of times people are helpful so has devised a *structured interview* with a list of specific questions such as 'How many times has someone held a door open for you?', 'Have you helped anyone carry their shopping?' and 'Have you seen anyone assisting a parent with a buggy?'. However, he is worried that this may produce very limited data so has an alternative plan to use an *unstructured interview*. This would begin with the question 'Please can you describe how friendly or helpful you have found people to be at the new shopping centre', after which he would base his



CONTINUED

questions on what they said. A colleague, Dr McArthur suggests that both methods have limitations. Interpreting the responses to the unstructured interview might lead to very *subjective* data, especially as Dr Singh already believes that the finding the new shopping centre layout encourages helpfulness. Although the data from the closed questions in the structured interview might produce more objective measures, this would limit opportunities for asking participants to expand on their answers. The colleague suggests that a **semi-structured interview** might be better.

ACTIVITY 1.7

Look at the previous Research Methods box.

- Why is the first of Dr Singh's suggestions a *structured interview*?
- Why would the data from these questions be more *objective*?
- Why is the second plan an *unstructured interview*?
- What is the problem with *subjective* interpretations of the participants' responses in the unstructured interview?
- Suggest why a *semi-structured interview* would be better in this case.

Questions

- 5 Shareen and Judith are investigating people's phobias. They have decided to use self-reports. Shareen is suggesting using a questionnaire and Judith wants to interview people instead.
- Suggest **one** closed question and **one** open question that Shareen could use.
 - Suggest **one** reason why Judith might want to conduct an unstructured interview.
 - Describe **one** ethical problem that might arise in **either** Shareen's **or** Judith's version of the study.

1.3 Case studies

A **case study** is a detailed investigation of a single instance, usually just one person, although it could be a single family or institution. The data collected is detailed and in-depth, and may be obtained using a variety of different techniques. For example, in one case study, a participant may be interviewed, observed, given tests and asked to fill in questionnaires. Case studies are particularly useful for looking at rare cases where a detailed description is useful, and for following developmental changes where the progress of a child, or a person with a disorder, can be tracked through their improvement or decline. Case studies are therefore sometimes linked to therapy but it is important to remember that when the case study as a research method is being discussed, the therapeutic purpose is not the main aim.

In addition to using different methods or techniques within a case study, different sources of evidence may be used. These could include interviewing the participant themselves, their relatives or other people such as colleagues to obtain a wide range of information about the individual's history, as well as accessing pre-existing information medical or school records.

Evaluating case studies

In some ways, the findings from case studies are highly valid, as the individual is explored in great depth and within a genuine context such as their work or family. Validity may be improved further using **triangulation**, where the use of different techniques should produce similar findings, for example observations and interviews with the participant and questionnaires for their family should all lead to similar conclusions. The research includes details such as their past as well as their present situation, their social interactions, their thinking and their emotions as well as their behaviours. Such detail, however, carries risks. One potential problem is the development of a close relationship with the researcher. This may make the researcher subjective in their outlook, which would reduce the validity of the study. The level of detail can also be an ethical threat, as the questions asked may intrude into the participant's private life and they may feel unable to refuse to answer them. The detail about the individual may make it hard

to disguise their identity, even if they are not referred to by name, which would risk breaking the guideline of confidentiality.

Reliability is also an issue, as there is a single participant and perhaps one or only a few researchers. This, and their involvement with the case, means that they may find it hard to be objective, that is to take an external, unbiased view of the findings, for example when they interpret what the participant has said. So, a different researcher, with a truly external viewpoint, might interpret the findings differently. In addition, the findings are specific to the individual being studied so may only apply to very few other instances, or to none at all, i.e. the generalisability of findings from a case study is very limited.

KEY WORDS

case study: a research method in which a single instance, e.g. one person, family or institution, is studied in detail

triangulation: is when different techniques, e.g. observations, interviews and tests, are used to study the same phenomenon. If they produce similar results, this suggests the findings are valid

Applying your knowledge of case studies to novel research situations

You should be able to recognise case studies, and when it is appropriate to use one. You should also be able to suggest possible techniques that could be used in a case study. When making these decisions, it is important to consider the validity and reliability of the findings. One way that the validity can be improved is through triangulation, where different methods are used within the case study to obtain the same information: for

example, finding out about the participant's behaviour by observing them, interviewing them and asking their relatives to fill out a questionnaire. If the same results are obtained by all the methods, this suggests that the results are valid. Another consideration in planning case studies is an ethical one. Participants should be aware of their commitment so that they can give their informed consent, and particular attention should be paid to ensuring their privacy is not invaded and that confidentiality is maintained.

In addition to being able to evaluate the case study method, you should be able to plan a case study. This could involve giving details about the participant (or instance such as a family or institution) you would choose and describing a range of methods and techniques you would use in your study to enable triangulation. You would also need to be able to explain how you would plan the analysis and interpretation of the results.

RESEARCH METHODS

A psychologist in a sleep clinic has been conducting a case study on a patient, PMA, who has had very bad dreams for several years. The clinic has a sleep laboratory where sleeping patients can be videoed to record behaviours such as sleep-talking and sleep-walking. The psychologist uses an EEG (a machine that records brain waves) to follow PMA's sleep cycles and to detect when the patient is dreaming. The patient is then woken up and asked about their emotional state and what the dream is about. A test is also conducted when the patient is not dreaming, they are again woken up and asked about their emotional state. PMA has been asked to keep a dream diary to record when the bad dreams occur. Members of PMA's family have been interviewed to find out when the problems with nightmares started and how often they occur.

ACTIVITY 1.8

If you can, work in a pair. Individually, look at the Research Methods box and answer the following questions or share the questions between you with one person doing the even-numbered questions and one person doing the odd-numbered questions.

- 1 How many different methods and techniques can you identify?
- 2 Give two reasons why this is a case study.
- 3 The psychologist is concerned that if the patient PMA wanted to *withdraw* from the study, this might be quite difficult. Explain why this might be so.
- 4 Explain the *ethical* reason for the researcher referring to the patient as PMA.
- 5 Suggest one **other** ethical issue that might be a problem in this study.
- 6 Suggest **one** methodological issue that might be a problem in this study.

Peer assessment

In your pairs, look at the answers given by your partner. Consider the following points for each question:

- 1 Is each of the ideas given a research method or a technique? Are there any others?
- 2 There are only two features that make a study a case study. Do they have them both? Are other features, which are not strictly necessary for a case study, also included?
- 3 Is the explanation complete?
- 4 Check the reason given. It should be 'confidentiality' not 'privacy'. Why would 'privacy' be incorrect?
- 5 Is the ethical issue suggested relevant to this study? If so, why? If not, why not?
- 6 Is the methodological issue suggested relevant to this study? If so, why? If not, why not?

Questions

- 6 Damon and Inka are planning a case study of a person living with a disability, to investigate responses to situations involving helping.
 - a Suggest **three** techniques that Damon and Inka might use in their case study.
 - b
 - i Describe **two** ethical problems that might arise in the study that Damon and Inka are planning.
 - ii For **one** of these problems, suggest a possible solution.
 - c Explain whether the results from Damon and Inka's study would be typical of the way everyone would respond to emotional situations.

1.4 Observations

Observations involve watching human or animal participants. There are many different choices in the way that observations are conducted and we will consider each in turn. An awareness of these choices is important both for understanding and planning observational studies. The setting in which the observation takes place is one such choice. A **naturalistic observation** is conducted in the participants' normal environment, without interference from the researchers in either the social or physical environment. A **controlled observation** is conducted in a situation that has been manipulated by the researchers. This may be in terms of the social environment, such as varying group sizes or adding a model, or in terms of the physical environment, such as providing objects for play, different foods or new locations. Controlled observations can be done in either the participants' normal environment or in an artificial situation such as a laboratory.

KEY WORDS

naturalistic observation: a study conducted by watching the participants' behaviour in their normal environment without interference from the researchers in either the social or physical environment

controlled observation: a study conducted by watching the participants' behaviour in a situation in which the social or physical environment has been manipulated by the researchers. It can be conducted in either the participants' normal environment or in an artificial situation

CONTINUED

unstructured observation: a study in which the observer records the whole range of possible behaviours, which is usually confined to a pilot stage at the beginning of a study to refine the behavioural categories to be observed.

structured observation: a study in which the observer records only a limited range of behaviours

behavioural categories: the activities recorded in an observation. They should be operationalised (clearly defined) and should break a continuous stream of activity into discrete recordable events. They must be observable actions rather than inferred states

At the beginning of a study, observations may be non-focused, that is, the whole range of possible behaviours are considered. If this continues throughout the study, it is called an **unstructured observation**. More often, the range being studied is narrowed to a fixed set of behaviours, a technique is called a **structured observation**. The specific activities to be recorded are clearly defined in **behavioural categories**. This helps the observers to be consistent, i.e. it improves **inter-observer reliability**.

Another decision to be made is the role of the observer in the social setting. This may be participant or non-participant. A **participant observer** is part of the social setting so may, for example, engage in conversation with adults or play with children. A **non-participant observer** does not become involved in the situation being studied. This can be achieved by watching through one-way glass or by keeping apart from the social group of the participants, for example by sitting in a separate place such as a bench in a park.

KEY WORDS

inter-observer reliability: the consistency between two researchers watching the same event, i.e. whether they will produce the same records

participant observer: a researcher who watches from the perspective of being part of the social setting

non-participant observer: a researcher who does not become involved in the situation being studied

The role played by the observer may be either **overt** (it is obvious that they are an observer) or **covert** (they are hidden or disguised so the participants do not know the individual is an observer). A participant observer would be overt, if they were holding a clipboard. When a participant observer is disguised as a member of the social group (Figure 1.9), or when a non-participant observer is physically hidden (e.g. by using CCTV), they are covert.

Participants cannot be aware that they are being watched if the observer is covert. This increases validity as it is unlikely that participants would be affected by being observed, so demand characteristics and the effects of social desirability are reduced. However, covert observations raise practical issues, as the observer must be either hidden, far away or disguised in their role. This may make data collection more difficult, potentially reducing validity and reliability. Furthermore, covert observation raises ethical issues as the participants cannot give informed consent, and if they work out a participant observer's role this can cause distress.

KEY WORDS

overt observer: the role of the observer is obvious to the participants

covert observer: the role of the observer is not obvious, e.g. because they are hidden or disguised



Figure 1.9: A covert participant observer is disguised by being part of the social group: which one is the observer?

Evaluating observations

Naturalistic observations have the advantage that the behaviours seen are true to life. They are more likely to reflect the way the individuals really behave than if there is interference in the situation from researchers, as is the case in controlled observations. However, there is no guarantee that the behaviours being studied will actually occur in a naturalistic situation, so it may be necessary to use a controlled observation.

Using an unstructured observation ensures that any important behaviours that happen are recognised and recorded. However, it may be very difficult to collect data on all the activities accurately and many behaviours may be irrelevant to the aims of the study. It is therefore likely, especially when only specific activities are of interest, that a structured observation will produce more reliable data as the observer will only need to focus on a small number of categories.

The role played by the observer, and the participants' awareness of this, affect validity. If participants are unaware of the observer, or can ignore them, their activities are more likely to reflect their normal behaviour. This means that covert observers would produce more valid results than overt ones. However, in the case of covert participant observers there is an ethical issue of deception. The participants may interact with the observer in ways that they might have chosen not to if they had been aware that the individual was a researcher. This could invade their privacy and cause harm by distressing them. However, an overt observer may alter the behaviour of the participants as they are aware that they are being watched. This would reduce the validity of the findings as the activities being recorded are less likely to reflect real-world behaviour.

Applying your knowledge of observations to novel research situations

Observations can be used either as a research method or as a means to collect data in other research methods such as case studies, experiments or correlations. You should be able to distinguish between these two situations. Observation is being used as a research method in itself when the study consists solely of a means to collect data by watching participants and recording their behaviour directly to provide data. Observations are used as a technique to collect data

about variables in other research methods when they are used to measure the dependent variable in an experiment or one or both variables in a correlation. In a case study, observations can be used alongside other techniques to explore a single instance in detail.

You will need to be able to decide when it is appropriate to use observations as a method, or as a technique within other methods. You should be able to recognise and justify choices about naturalistic versus controlled, structured versus unstructured, participant versus non-participant and covert versus overt observations. You will also need to be able to suggest ways of achieving these, such as how to make an observer participant or covert.

In addition to being able to select and evaluate the observational method and the different techniques, you should be able to plan studies using any of the techniques in a way that is appropriate to a given aim. All of the 'technique' decisions (e.g. naturalistic, structured, participant, overt) need to be made for every observation. So, you may need to choose and describe the location for a naturalistic observation or how the environment will be controlled; the behavioural categories to be recorded in a structured observation; how observers would aim to be included in the social setting (participant observation) or avoid the social setting (non-participant) and how an observer could be made covert.

RESEARCH METHODS

Dr Huang is interested in whether her students detect each other's emotions and plans a *controlled observation*. She asks three students to act as confederates. They are told to take it in turns to appear quite sad in the common room at lunchtime. Dr Huang is on lunch duty with a colleague so they can act as *non-participant, overt* observers as they walk through the common room. The students will take no notice of them as they are used to them being there. Dr Huang suggests that she and her colleague use a list of specific behaviours to record, so they are doing a *structured observation*. This will also help to raise *inter-observer reliability* as they will be working from the same definitions, such as recording 'shows concern without action' if people look at the confederate without moving towards them, 'verbalises concern' if someone goes up to the confederate and asks them if they are OK and 'takes action' if they engage in a behaviour such as putting their arm round the confederate or buying them a drink.

ACTIVITY 1.9

Look at the Research Methods box.

- Suggest **one** other *behavioural category* that might have been included in the structured observation and define it.
- If Dr Huang had conducted an *unstructured observation*, how would the method have differed?
- An alternative plan would have been to have conducted a *naturalistic observation* and watched to see if any students appeared to be sad and how others responded. Outline **one** way it would have been ethically more acceptable and **one** way in which it would have been ethically less acceptable.
- Suggest how Dr Huang could have used a *participant*, *covert observer* rather than being an overt observer.

not be practical or ethical to conduct an experiment which controlled children's long-term exposure to television and it would not be ethical to increase real-life exposure to violent television programmes. Both of these could, however, be investigated using correlations. Real-life exposure of a group of children to violent television could be measured and correlated with another measured variable, such as how aggressive their behaviour was in school.

KEY WORDS

co-variables: the two measured variables in a correlation

correlation: a research method that looks for a relationship between two measured variables. A change in one variable is related to a change in the other (although these changes cannot be assumed to be causal)

causal relationship: a link between two variables such that a change in one variable is responsible for (i.e. causes) the change in the other variable, such as in an experiment

Questions

- 7 Debra and Jin want to use observations to find out about the behaviour of animals. Debra wants to go to the park and hide in a tree to observe the animals that live there. Jin thinks it would be better to set up an artificial situation and watch laboratory rats interacting with objects they would put in a special box.
- Who is suggesting a naturalistic observation and who is suggesting a controlled observation?
 - Explain whether the observers in Debra's study would be overt or covert.
 - For either Debra's or Jin's suggestion, decide whether it should be conducted as an unstructured or a structured observation and justify your choice.

To look for a correlation between two variables, each variable must exist over a range, sometimes called 'continuous data', and it must be possible to measure them numerically. Suitable variables would include durations of time, totals of tallies, numbers from a rating scale or test scores. Several techniques can be used to collect data for correlations, such as self-reports, observations and different kinds of tests and tasks.

It is important to recognise that any link found between two variables in a correlation cannot be assumed to be a **causal relationship**, that is, we cannot know whether the change in one variable is *responsible* for the change in the other variable (Figure 1.10). This is possible in the case of an experiment, as we can know that there is a causal relationship between the IV and the DV because changes in the IV are shown to be the cause of changes in the DV.

However, we cannot say from one correlation that an increase in one variable has caused an increase (or decrease) in the other because it is possible that the changes in both variables could be the result of another factor. Imagine that two variables are being measured: attention in class and score on a test. If these co-variables correlate it would be tempting to say that paying attention in class is responsible for good test

1.5 Correlations

A correlational analysis is a technique used to investigate a link between two **co-variables** (measured variables). **Correlations** are useful when it is possible only to measure variables, rather than manipulate them, i.e. when an experiment cannot be conducted. This may be because changing the variables would not be practical or would be unethical. For example, it would



results but we cannot be sure of this. It could be that both of these factors depend on a third variable, such as how hard-working the individual student is. The sort of student who pays more attention in class might also study much harder for the test. All we can conclude is that the two factors we have measured vary together, not that there is a cause-and-effect or causal relationship between them. As a consequence, it is important that you refer to 'co-variables' or 'measured variables' in a correlation and not independent and dependent variables. To make judgements about causality, an experiment must be used, so that we can be more certain that it is the manipulation of one variable that is responsible for the change in the other. If, on the other hand, we conduct a correlational study and find that there is *no* link between two variables, then we can conclude that there is no causal relationship.



Figure 1.10: A bizarre positive correlation has been reported between ice cream consumption and murder rates. This relationship is a correlation, however, so we cannot conclude that eating ice cream causes people to commit murder.

The nature of the relationship between the two variables in a correlation can be described in terms of its *direction*. In a **positive correlation**, the two variables increase together. The change is in the same direction, so higher scores on one variable correspond with higher scores on the other. For example, in a positive correlation between exposure to aggressive models and violent behaviour, greater exposure to models would be linked to higher levels of violence. When two variables are **negatively correlated**, higher scores on one variable correspond with lower scores on the other. For example, a negative correlation might exist between number of years in education and level of obedience: people with fewer years of education are more obedient

(see also Section 1.9 on how to draw a scatter graph and a discussion of the *strength* of a correlation).

KEY WORDS

positive correlation: a relationship between two variables in which an increase in one accompanies an increase in the other, i.e. the two variables increase together

negative correlation: a relationship between two variables in which an increase in one accompanies a decrease in the other, i.e. higher scores on one variable correspond with lower scores on the other

Hypotheses in correlational studies

As with hypotheses in experimental studies, a hypothesis in a correlational study needs to operationalise the co-variables as it predicts that there will be a relationship between them. The hypothesis 'There will be a correlation between amount of sleep and emotional reactivity' does not operationalise either variable. This could be improved by saying 'There will be a correlation between the number of hours a person sleeps for and their emotional reactivity indicated by how loudly they cry during a sad film'.

At the beginning of this chapter, we considered Dr Huang's idea that there could be a link between computer-game playing and A Level performance. A non-directional hypothesis in a correlational study testing this idea could predict that *There will be a correlation between the number of computer games a student plays and their final A Level grade*. Both of the co-variables are operationalised here.

A directional hypothesis for this study could say: *There will be a negative correlation between the number of computer games a student plays and their final A Level grade*. We might make this prediction if we believed that the time spent playing games might stop students working. However, a different directional hypothesis could be: *As the number of computer games a student plays increases, their A Level grade increases*. We might make this prediction if we believed that students who engaged more with technology, even through games, were also more likely to benefit from technology-based learning aids. Remember that hypotheses in

correlational studies should not say that one factor *causes* the change in the other.

Evaluating correlations

A correlational study can only be valid if the measures of both variables test real phenomena in effective ways. To achieve this, the variables must be clearly defined and relate directly to the relationship being investigated, making them valid. The reliability of a correlation depends on the measures of both variables being consistent. So, for some correlations, such as those using scientific scales (such as volume in cm³ or time in seconds), the measures will be highly reliable. In other cases, such as studies correlating variables measured using self-reports or observations, there is a risk that reliability will be lower. This is because results from these measures may be less objective than from scientific measurements. The main issue with correlations, however, is to remember that the conclusions do *not* necessarily reflect a causal relationship.

Applying your knowledge of correlations to novel research situations

Correlations provide a good starting point for research. They can indicate whether a relationship exists that might be worth pursuing with other research methods, such as experiments. Correlations are also useful because they enable researchers to explore problems when it is not practically or ethically possible to conduct experiments. You should be able to distinguish between correlations (with *two* co-variables which are both measured) and experiments (where there is one variable—the IV—that is manipulated by the researcher and only *one* that is measured: the DV). You should also be able to recognise the difference between positive and negative correlations.

You will need to be able to decide when it is appropriate to use a correlation rather than any other method, for example when it is impossible to manipulate variables for practical or ethical reasons. You should also be able to justify choices about ways to measure the variables in a correlation and to suggest whether you would expect a positive or a negative correlation in a study. Finally, you will need to understand how to display the results of a

correlational study on a scatter graph. This is discussed in Section 1.9.

In addition to being able to select and evaluate the correlational method, you should be able to plan studies using a correlation in a way that is appropriate to a given aim. Remember that you must create two co-variables that can each be measured on a scale. So, you may need to choose a technique for measuring these, which could be different for each variable, for example by using a test or a questionnaire, interview or observation that produces quantitative data.

RESEARCH METHODS

Professor Chao is studying phobias. She thinks that there may be a *correlation* between how long a phobia has lasted and how severe it is. She is asking her sample of participants with phobias to record how many years they have suffered with their fear and rate how much the phobia interferes with their life on a scale of 1 (hardly at all) to 10 (almost constantly and prevents me from functioning normally). When looking for a link between the co-variables, there are several possible outcomes. One is that there may be a *positive correlation*: phobias that have lasted longer may be more severe. Another is that there is no relationship between the two variables. If so, when she plotted the results on a *scatter graph*, the points would appear randomly placed, rather than lying on a line.

ACTIVITY 1.10

Look at the Research Methods box.

- An alternative outcome might have been that phobias that have lasted longer are less severe. Explain why this is a negative correlation.
- If Professor Chao had written a hypothesis for a negative correlation saying *Phobias that last longer will be less severe* what would it need to make it complete?
- If Professor Chao found a positive correlation, it would be tempting, but incorrect, to say that the passage of time makes phobias worse. Why would this conclusion be incorrect?



Questions

- 8 Ekua and Takis are going to find out if there is a correlation between the amount of coffee people drink and the number of dreams they recall.
- Explain why this is a correlational study and not an experiment.
 - Suggest whether the results will show a positive correlation or a negative correlation.
 - Suggest **one** way to measure the amount of coffee that people drink.
 - Explain either **one** strength or **one** weakness of the way you have suggested measuring this variable.

1.6 Longitudinal studies

A **longitudinal study** is one that follows a single group of participants over time, studying one or more variables at intervals. By measuring variables over weeks, months, years or even decades, a longitudinal study can detect changes in individuals. People change over time, not just as they age but as they accumulate life experiences. In addition, longitudinal studies can explore the effect of specific 'experiences' such as interventions or events on development.

KEY WORDS

longitudinal study: a research method that follows a group of participants over time, weeks to decades, looking at changes in variables to explore development or changes due to experiences, such interventions, drugs or therapies

cross-sectional study: compares people at different ages or stages by comparing different groups of participants at one point in time

cohort: a group of participants selected at the same age or stage

The use of longitudinal studies is an alternative to cross-sectional studies as a way to look at effects over time. In a **cross-sectional study**, differences between people of differing ages, or at different stages (such as

different periods after an intervention) are explored by comparing different groups at one point in time. For example, to explore changes in mindfulness with age, a cross-sectional study could look at 10, 20, 30, 40 and 50-year-olds. Although a cross-sectional study allows researchers to investigate long-term influences, it is difficult to separate changes over time from changes due to differences due to the individuals growing up at different times. For example, if a cross-sectional study looking at helpfulness in people of different ages, these could be due to differences in upbringing at the time when they were children or the changing expectations of society rather than their age. Such problems are avoided in a longitudinal study as a single **cohort**, a group of participants all selected at the same age or stage, are chosen and tested at different time points. A cohort might be individuals born in a certain year, people about to embark on a course of treatment such as a new drug regime or people experiencing a life change, such as at the beginning of a pregnancy.

Longitudinal studies may have different intentions or designs. Some studies track patterns of change in variables to investigate them, such as changing emotions or beliefs as we get older. Other studies record longitudinal changes in two or more variables to investigate possible correlations between them, such as exposure to a multicultural society and prejudice. Yet others use an experimental design where one or more variables, measured over an extended period of time, are expected to change as a consequence of a deliberate manipulation. In the last case, the variables being measured are dependent variables and the testing at different time points in relation to the deliberate manipulation is an independent variable. To provide a baseline for tracking long-term change, the variable is usually measured at the outset of the study, and again after one or more periods of time. This procedure can be thought of as a very long-term repeated measures design, so is sometimes referred to as a 'quasi-experimental' design. This procedure can be used, for example, to measure the consequences of interventions such as educational, health, mental health or occupational ones that aim to improve outcomes. One level of the independent variable is the baseline test or 'pre-intervention' time point and one or more repeat testing sessions occur, providing a 'post-intervention' level of the independent variable and possibly more time points during, or long after the intervention as a follow-up.

KEY WORDS

longitudinal design: an experimental design where the same participants are tested on two or more occasions over a long time, e.g. before and after a six month intervention

situational variable: a confounding variable caused by an aspect of the environment, for example the amount of light or noise.

Intelligence quotient (IQ): a measure of general reasoning and problem solving ability

Evaluating longitudinal studies

A longitudinal study has the clear advantage over a cross-sectional study that researchers can be confident that any changes they detect are due to the passage of time and not the effect of differences between the cohorts of participants representing each age or stage. This has two benefits. **Longitudinal designs**, such as repeated measures designs, retest the same individuals. This means that the researchers can be confident that differences are not due to participant variables. A second consequence is that it also overcomes potential confounding variables. For example, in the Terman study of highly intelligent children (see Figure 1.11), differences in educational methods, career opportunities and other societal factors that could have been problematic in a cross-sectional study would not influence this longitudinal study. Confounding variables such as this, which are aspects of the environment, are called **situational variables**. This method is therefore the most valid test of developmental changes.



Figure 1.11: Lewis Terman, who began a long-running longitudinal study in 1922, and psychologist Lee Cronbach, one of his participants

A very long-running longitudinal study is the Terman Life Cycle Study of Children with High Ability, which began in 1922. Although the study eventually debunked the myth that a high **Intelligence quotient (IQ)** inevitably leads to success, some of its participants did become famous, such as the psychologist Lee Cronbach, who—perhaps unsurprisingly specialised in psychological testing and measurement—became president of the American Psychological Association.

However, precisely because longitudinal studies last for a long time, the sample size is likely to fall; this is called **sample attrition**. Terman's initial sample, in 1922, was 1528 children. In a comparison between data collected in 1996 and 1999, although 162 questionnaires were returned in 1999 just 119 participants provided data on both occasions (Holahan & Velasquez, 2011). Where did they go? Participants may leave studies for many reasons, but these tend to be more apparent in a longitudinal study. They include: growing older and deciding to withdraw consent (e.g. when consent was originally given by a parent), boredom (from being repeatedly tested), moving home and becoming uncontactable, life events that prevent participation (poor physical or mental health, lack of time, becoming homeless or going to prison) and, ultimately, dying.

KEY WORD

sample attrition: the loss of participants from a sample over time. This may be due to many reasons, such as losing contact, the desire to discontinue, for example though boredom, being unavailable or death

A smaller sample is likely to be less representative, but especially so in this case. Sample attrition means that generalisability is progressively threatened as the sample becomes concentrated on willing, motivated, people who live stable, healthy, law-abiding lives. These people may not be the most psychologically interesting! The people who stay in a longitudinal study may even do so because they feel 'special' by doing so, creating an issue for validity. For example, Terman's participants knew they had high IQs and the sense of being important was reinforced by having a name: they were referred to as Terman's 'Termites'. Participation could, for example, have given them a motive to compete or succeed, or a sense of community i.e. the very act of participation will have made these individuals different from non-participants, even ones of similarly high intelligence.

Another problem caused by the duration of longitudinal studies relates to reliability. To ensure that data collection at each time point is consistent, the measures used need to be standardised. However, over time, ways to measure variables may be improved, potentially leading to the need to introduce different, better measures. In addition, researchers working on the study may move on so, for example, the individuals administering interviews may not remain constant. Conversely, if researchers continue working on a study for a long period of time, they are likely to get to know the participants if they have direct contact with them. The changing development of these relationships can also be confounding variables.

Finally, longitudinal studies can present particular ethical issues. For studies starting with children, obtaining consent in a way that the children understand is desirable, as well as from their parents or guardians. Furthermore, it is harder to decide whether a child wants to withdraw from a study than it is with adults, who are more likely to speak up for themselves. Participants must also be repeatedly asked for consent (and reminded of their right to withdraw) but must also be recontacted at each time point. This requires the maintenance of an extensive bank of contact details, which presents an issue for confidentiality.

Applying your knowledge of longitudinal studies to novel research situations

Longitudinal studies provide information about long-term changes. They are therefore suitable for investigating aims that relate to development and ageing, but also to any other life changes. Such changes include both those that are the result of a life event, such as recovery from an illness or the effects of a marriage or divorce, and those that are the consequence of a manipulation, such as attending a course or planned exposure to social groups. You should be able to identify the variables being measured and whether these are simply being tracked over regular time points to monitor

possible changes, or whether there is a manipulation or influence which is actively intended to cause change. In the latter case, time points could be 'before, during, after and follow-up' in relation to the manipulation. This situation would be a longitudinal study with an experimental design.

You will need to be able to decide when to use a longitudinal study rather than any other method, and which type of longitudinal study is appropriate in a given situation. You should also be able to suggest variables to study and ways to measure them, e.g. using a test, questionnaire, interview or observation, and how the results could be analysed and used.

In addition to being able to select and evaluate the longitudinal method, you should be able to plan a longitudinal study in a way that is appropriate to a given aim, such as whether or not there is a manipulated independent variable. Remember that you need to identify one or more variables to investigate, consider how they will be measured, and decide on the intervals between the testing time points and how you will standardise procedures to ensure consistency over time, as well as planning how to re-contact participants to retest them.

Questions

- 9 Hanif is planning a longitudinal study to measure changes in professional ability as adults progress through their careers. He has a sample of participants from different types of employment and has various measures of career progression such as how much they earn, the seniority of their position in their organisation and how many staff in the organisation are above and below them.
 - a Suggest how often Hanif should contact his participants to track their progress.
 - b Hanif wants to measure how happy they are in their career. Suggest how Hanif could measure this.
 - c The participants will need to be recontacted each time Hanif wants to test them. Suggest how Hanif could do this.

1.7 The definition, manipulation, measurement and control of variables

Variables are factors that change or can be changed. In experiments, these are the independent and dependent variables as well as variables that are or are not controlled. In correlations, there are two measured variables called co-variables.

Experiments look for changes or differences in the dependent variable (DV) between two or more levels of the independent variable (IV), which are set up by the experimenter. It is important that the IV is operationalised, so that the manipulation of the conditions represents the intended differences. At the beginning of this chapter, we discussed the idea of the **aim** of a study, that is, what the research is trying to find out. The aim refers to the variables but not necessarily very clearly. Operational definitions of variables help to add clarity to the intention expressed in the aim.

Consider a study testing the effect of age on susceptibility to false memories. The IV would be age, with, for example, 'young', 'middle-aged' and 'old' groups. It is important to know *how* old the people in the groups are; this is operationalisation. You might operationalise 'young' as under 20 years old, 'middle-aged' as 40–50 years old and 'old' as over 70. The DV must also be operationalised, so it can be measured effectively. We could operationalise the DV by counting the number of details 'remembered' about the false memory or how convinced the participants were that it was true.

KEY WORD

Aim: the intention of the study, the idea being tested or the purpose of the research, such as to investigate a question or solve a problem

RESEARCH METHODS

Dr Huang is thinking about buying new chairs for her classroom. Her **aim** is to explore whether hard or soft chairs help her students to work better (Figure 1.12). She wonders whether to predict a *non-directional (two-tailed) hypothesis*: *There is a difference in work rate of students sitting on comfortable and uncomfortable chairs*. Another psychology teacher says that students respond well to other comforts like access to a drinks machine or snack bar, and the soft chairs might make them happier, so they work harder. Dr Huang rewrites her prediction as a *directional (one-tailed) hypothesis*, saying: *Students on comfortable chairs will have a higher work rate than ones sitting on uncomfortable chairs*. A third teacher is not convinced and suggests that if the students are too comfortable, they will become sleepy and lazy, so work less. The hypothesis would then be: *Students on comfortable chairs will have a lower work rate than ones sitting on uncomfortable chairs*. This is also a directional (one-tailed) hypothesis. Her *null hypothesis* (in all cases) would be: *Any difference in work rate of students sitting on comfortable and uncomfortable chairs is due to chance*.

Now imagine a study that aims to look for a link between sleep and emotions. A non-directional hypothesis might be: *There will be a correlation between amount of sleep and emotional reactivity*. The possible directional hypothesis could say: *There will be a positive correlation between amount of sleep and how emotional someone is*. Note that the hypothesis does *not* say that one factor *causes* the other to change. The null hypothesis here would be: *Any correlation between amount of sleep and emotional reactivity is due to chance*.



Figure 1.12: Would you work harder in lessons if you had more comfortable classroom chairs?

Controlling variables and standardising procedures

Controlling of variables

Psychologists need to control variables in their studies to be more certain about their findings, i.e. to improve validity. In particular, in experiments, it is important to control confounding variables, that is, ones that have a consistent effect so they confound, i.e. confuse, the results. Confounding variables act on the DV selectively in one level of the IV so can interfere with the results in one of two ways. One possibility is that they can work against the effect of the IV, counteracting its effect on the DV, so preventing a real effect from being identified. Alternatively, they can increase the apparent effect of the IV and so suggest that there is an effect that doesn't really exist. These variables are the most important to control.

Other uncontrolled variables, which have a random effect on the DV across all levels of the IV, are not so problematic. The difficulty is to identify which variables will be important to control before the experiment starts. This is one function of a pilot study, a preliminary test of the procedures of a study. The intention is to identify potentially problematic uncontrolled variables and control them. This prevents or limits their effect to avoid making the results difficult to interpret if it is hard to separate their effects from those of the IV.

Consider Dr Huang's study of students and chairs. Perhaps Dr Huang compares one class in a room with the new (soft) chairs and another class in a different room with the old (hard) chairs. If the room containing the new chairs happens to have better lighting, Dr Huang may find that the students in the 'soft chairs' condition perform better. However, this may be due to the confounding variable of brighter lighting rather than the comfort level of the chairs. This is an example of a situational variable, because lighting is an aspect of the environment. Another possible confounding variable is how hard-working the individual students are. We might expect normally hard-working students to be randomly distributed among the different classes, in which case this variable is not a problem. However, suppose that all the students in the 'soft chairs' class do arts and humanities subjects and all the students in the 'hard chairs' class do maths and sciences. If Dr Huang

happens to use a test of data analysis as her measure of the DV, she might find that the students in the 'hard chairs' level of the IV perform better. This would suggest that the soft chairs make students perform worse but could in fact be due to the uncontrolled variable of subject groups. This is an example of a participant variable, because the difference has been caused by a feature of the individuals, i.e. their ability in maths.

Standardisation

Controls make sure that the levels of the IV represent what they are supposed to, i.e. that the differences between them are going to create the intended situations to test the hypothesis. This helps to ensure validity. It is also important that every participant is treated in the same way. This is the process of standardisation. One way that this is achieved is by having **standardised instructions** that give the same advice to every person in the study. Imagine a questionnaire testing attitudes to helping behaviour. All participants would need to have the same advice about how to fill it in, so that any effects of social desirability—the influence of needing to give answers that were acceptable to society—were equally likely.

KEY WORD

standardised instructions: the written or verbal information given to participants at the beginning and sometimes during a study that ensure the experience of all participants, regardless of level of the IV, is as similar as possible

The procedure itself also needs to be standardised. This involves having equipment or tests that are consistent, i.e. that measure the same variable every time and always do so in the same way. Consider the questionnaire about attitudes to helping again. All the questions should focus on the same aspect of behaviour, i.e. helping, rather than some looking at a different but possibly related factor, such as being friendly or happy. In laboratory experiments, standardisation is easier than in other studies, as equipment is likely to be consistent, for example stopwatches or brain scans. However, some of these measures, such as brain scans, may need to be interpreted and this must also be done in a standardised way (Figure 1.13).

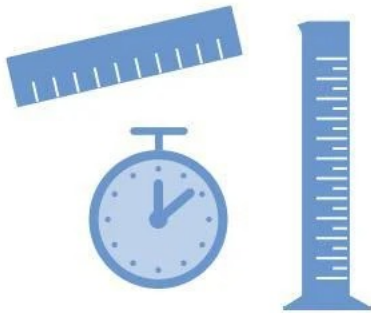


Figure 1.13: Scientific instruments are likely to produce objective, reliable data

Applying your knowledge of variables and controls to novel research situations

In experiments it is important to be able to decide how to operationalise the IV to produce the different conditions (to achieve validity) and to measure the DV in a consistent (reliable) way and to be able to justify these choices. You will also need to be able to decide what controls it is appropriate to use and to suggest how these can be implemented.

RESEARCH METHODS

In Dr Huang's experiment about classroom chairs, the IV of hard and soft chairs must be operationalised. The text also referred to them as 'comfortable' and 'uncomfortable' chairs, but this still does not make clear what is meant by 'hard' and 'soft'. This could be done by saying 'chairs with wooden/plastic seats' and 'chairs with padded seats'. Operationalisation of the DV is also needed. The text referred to working 'better' and 'harder' but this is also incomplete. We need to expand on the idea of *work rate*, which was also used. This might be measured by counting the number of pieces of homework handed in late, or the time spent doing extra work. Either of these would indicate the amount of work being done. There are many *confounding variables* that could be important in this study, for example, some of

CONTINUED

the students might work harder anyway or the rate of work might vary with the weather. If students worked harder on sunny days, this would be a *situational variable*. The important variables to control are those that could confound the results. For example, if there was a choice of chairs, the students who chose to sit on comfy ones might be the laziest. If left as an *uncontrolled variable*, this could alter the results by making it look as if soft chairs made students work less.

In the description of the correlation on sleep and emotions, the co-variables were the 'amount of sleep' and 'how emotional someone is' or their 'emotional reactivity'. It is important to operationalise variables in correlations too. An operational definition of 'amount of sleep' could be the hours and minutes spent unreactive, or with an EEG indicating REM or nREM sleep. A questionnaire could be used to measure their emotions, with an operational definition of 'emotional reactivity' as how likely a person is to respond negatively to innocent comments, or how often they argue.

ACTIVITY 1.11

Look at the Research Methods box.

For the experiment:

- Suggest **one other** way the IV could be operationalised.
- Suggest **one other** way the DV could be operationalised.
- Would a possible difference between how lazy students were be a *situational variable* or a *participant variable*?
- Suggest **one other** possible *uncontrolled variable*.

For the correlation:

- Suggest **one other** way the variable of 'emotions' could be operationalised.



Questions

Umar and Saira are conducting an experiment on revision strategies. They are comparing students who revise alone and students who revise with other people. They are testing the hypothesis 'Revising alone is more effective than revising with other people'. Their teacher has told them that they are unlikely to find out anything useful as there will be too many uncontrolled variables to produce useful results.

- 10 a Identify the IV and the DV in Umar and Saira's experiment.
- b Explain whether Umar and Saira's hypothesis is directional (one-tailed) or non-directional (two-tailed).
- 11 a i Umar suggests that one confounding variable could be that students who work with other people find the company motivating, so are able to work for longer.
Explain how this confounding variable could affect the results of their experiment.
- ii Suggest how this variable could be controlled.
- b Saira suggests that another confounding variable could be that students who work with other people find them distracting, so work less effectively.
- i Explain how this confounding variable could affect the results of their experiment.
- ii Suggest how this variable could be controlled.
- 12 A friend says that student ability could be an uncontrolled variable. Umar and Saira hope that this will have a random effect on their results.
Explain whether this would be more or less important than the uncontrolled variables that Umar and Saira identified.

1.8 Sampling of participants

A **population** is a group of people (or animals) with one or more characteristics in common. For example, the population of a country is all the people who live there, the population of internet users is everyone who can access the internet. A population could also be people who share a particular interest, such as 'all football supporters' or who have a particular feature, for example

'all left-handed people'. The **sample** is the group of people who participate in a study. They are taken from a population and should ideally be representative of that group so that the findings will be generalisable. Details about the sample, such as age, ethnicity and gender are often important in investigations, so are commonly reported in the core studies. This is because these features may have an influence on psychological differences. Other characteristics of the sample, such as socio-economic status, education, employment, geographical location or occupation, may also be relevant.

The size of the sample also matters. Small samples are less reliable and are likely to be less representative. This is because a smaller group of people is unlikely to contain all the differences or variations that exist with the populations and some of those will be important to any study.

The different **sampling techniques** described next produce samples that differ in terms of how well they represent the population. The extent to which they are representative of the population determines how effectively generalisations can be made.

KEY WORDS

population: the group, sharing one or more characteristics, from which a sample is drawn

sample: the group of people selected to represent the population in a study

sampling technique: the method used to obtain the participants for a study from the population

Opportunity sampling

Studies are often conducted with the people who are around at the time. Selecting participants in this way is called **opportunity sampling**. An opportunity sample is unlikely to represent the population fairly because readily available people will tend to be alike so will probably not include the variety that exists. For example, many studies are conducted using university students as they are convenient for the researchers. However, this means that the sample will be predominantly young, with a better than average education. The results may not, therefore, reflect the scores that people of different ages or educational opportunities might produce. Despite this potential

problem, opportunity sampling is a very commonly used method, even for professional psychologists, because for many investigations the results are unlikely to be affected by age or education.

Volunteer (self-selected) sampling

Rather than the researcher choosing individuals to ask, they may invite people to take part in their study. They might put up an advertisement, make an announcement or post a request on the internet. In this way, the people who respond and become participants choose to do so, i.e. are volunteers, so are described as a **volunteer sample** (Figure 1.14). As the individuals are self-selected, that is they choose whether to join in, this sampling technique is unlikely to be representative of the population. Volunteers may have more free time than average and, apart from being willing, often have other characteristics in common such as being better educated. Nevertheless, it is a useful technique when looking for participants who are unusual in some way, for example in Baron-Cohen et al.'s study (see Chapter 3), where people with autism spectrum disorder were needed.

KEY WORDS

opportunity sample: participants are chosen because they are available, for example university students are selected because they are present at the university where the research is taking place

volunteer (self-selected) sample: participants are invited to participate, for example through advertisements via email or notices. Those who reply become the sample

random sample: all members of the population (i.e. possible participants) are allocated numbers and a fixed amount of these are selected in an unbiased way, for example by taking numbers from a hat

Public Announcement

WE WILL PAY YOU \$4.00 FOR ONE HOUR OF YOUR TIME

Persons Needed for a Study of Memory

*We will pay five hundred New Haven men to help us complete a scientific study of memory and learning. The study is being done at Yale University.
*Each person who participates will be paid \$4.00 (plus 50c carfare) for approximately 1 hour's time. We need you for only one hour: there are no further obligations. You may choose the time you would like to come (evenings, weekdays, or weekends).

*No special training, education, or experience is needed. We want:

Factory workers	Businessmen	Construction workers
City employees	Clerks	Salespeople
Laborers	Professional people	White-collar workers
Barbers	Telephone workers	Others

All persons must be between the ages of 20 and 50. High school and college students cannot be used.
*If you meet these qualifications, fill out the coupon below and mail it now to Professor Stanley Milgram, Department of Psychology, Yale University, New Haven. You will be notified later of the specific time and place of the study. We reserve the right to decline any application.
*You will be paid \$4.00 (plus 50c carfare) as soon as you arrive at the laboratory.

TO: PROF. STANLEY MILGRAM, DEPARTMENT OF PSYCHOLOGY, YALE UNIVERSITY, NEW HAVEN, CONN. I want to take part in this study of memory and learning. I am between the ages of 20 and 50. I will be paid \$4.00 (plus 50c carfare) if I participate.

NAME (Please Print).....

ADDRESS

TELEPHONE NO. Best time to call you

AGE OCCUPATION SEX

CAN YOU COME:

WEEKDAYS EVENINGS WEEKENDS

Figure 1.14: Would you respond to this advert?

Random sampling

Opportunity and volunteer samples may be biased – they will probably contain very similar people so are unlikely to include the spread of characteristics in the population. In **random sampling** each person in the population has an equal chance of being chosen so the sample is much more likely to be representative. Imagine a researcher is looking for a sample of students at a school and put an advert for volunteers on the library notice board. Students who never go to the library cannot be included so their sample might be biased towards those who work the hardest. Similarly, if they took an opportunity sample from the common room, it would only include students who are relaxing. Now their sample might be biased towards the *least* hard-working. To obtain a representative sample the researcher could

instead use a numbered list of all students and use a random number generator to choose the participants. This would be a random sample as any individual is equally likely to be chosen. If the population is small, such as all the members of one class, the research can simply give each person a number, put pieces of paper with each number on in a hat, and draw out numbers until there are enough for the sample.

Applying your knowledge of sampling techniques to novel research situations

An early step in any research is to obtain an appropriate sample. The extent to which generalisations can be made from research depends in part on how representative the sample is. It is therefore important to get the best possible sample. However, practical constraints prevent researchers from using random samples most of the time and, for many psychological phenomena, it is reasonable to believe that processes happen in a fairly universal way so some sample bias is unproblematic. However, to assume there are no differences in emotional responses, cognitive processing or behaviour between populations would be misleading. Indeed, cross-cultural research, the

psychology of individual differences and developmental psychology are three areas specifically devoted to the study of such differences. It is therefore important that you can recognise limitations in the sampling technique used. This means that you should be able to identify possible differences between individuals or groups that might matter for the particular phenomena being explored in a study. Imagine two researchers at different universities are both studying obedience and both want samples from people nearby but not their own students. One university is near a police college and the other is next to a hospital and the researchers both obtain opportunity samples with the same age and gender spreads from these workplaces. Even though the samples are similar in age and gender, the difference in occupations may mean that the results of their studies may be different, for example if police officers were more obedient than nurses.

You will need to be able to explain how you would use each sampling technique and to explain reasons for choosing each technique. For example, why it might be difficult to use a technique in practice or why generalisations could or could not be made from the sample obtained. Table 1.3 will help to remind you of the key advantages and disadvantages of one sampling technique compared to another.

Sampling technique			
	Opportunity sampling	Volunteer (self-selected) sampling	Random sampling
Strengths	Quicker and easier than other methods as the participants are available, therefore a larger sample can be readily obtained.	Relatively easy because the participants come to the researcher. They are also likely to be committed, e.g. willing to return for repeat testing.	Likely to be representative as all types of people in the population are equally likely to be chosen.
Weaknesses	Likely to be non-representative as the variety of people available is likely to be limited, so they will tend to be similar and the sample could therefore be biased.	Likely to be non-representative as people who respond to requests may be similar, e.g. all have free time.	In reality everyone may not be equally likely to be chosen, e.g. if they cannot be accessed (if the original list is incomplete) or if mainly one type of participant, e.g. girls, happen to be selected. This is particularly important if the sample is small.

Table 1.3: Strengths and weaknesses of sampling techniques

RESEARCH METHODS

Your teacher has asked you to do a cognitive psychology investigation. You want it to work well so you want a *sample* of 30 people from which you can generalise. You think it would be easiest to take an *opportunity sample* from your classes at your college but you realise this might produce a biased sample as all your subjects are sciences. This might matter as the investigation is about using logic to solve problems and you think that science students might be especially good at the task. If so, their results might not be representative of the college *population* as a whole. A friend suggests making an announcement in the canteen asking for students studying all different subjects. This *volunteer sampling* technique might be better but there would be no guarantee of getting a range of people. You decide that the best sampling technique would be to choose individuals at random from a list of all the students in the college. You hope that this would mean you would be equally likely to get students taking each subject. You decide to use this *random sampling* technique based on the students' examination candidate numbers. You enter the candidate numbers into a random number generator and use the first 30 numbers that are generated.

CONTINUED

- Dr Singh is investigating the effects of ageing and is asking all the residents at two local care homes for their help.
- Dr Huang pulls student numbers on the college register from a hat to select a sample for a new study on homework and part-time jobs.

Questions

- 13 Ichiro is planning a study about the benefits students gain from doing non-academic work, such as clubs, music, art, sports or part-time jobs. He is asking his friends for advice on how to obtain his sample.
- a Kiaria says he should use an opportunity sample.
Explain a disadvantage of using an opportunity sample in Ichiro's study.
 - b Haniya says he should use a volunteer sample.
Explain a disadvantage of using a volunteer sample in Ichiro's study.
 - c Habib says he should use a random sample.
Explain a disadvantage of using a random sample in Ichiro's study.

ACTIVITY 1.12

Look at the Research Methods box. What problems would the following situations lead to?

- You revert to your idea of an *opportunity sample* but use your neighbours, who are mainly retired people.
- You followed your friend's advice about *volunteer sampling* but lots of the younger students were in detention at lunchtime that day.

Read the examples next. Which *sampling technique* is being used in each situation?

- Professor Chao is doing some internet research and is recruiting participants by posting on social media asking for people to help with her study.

1.9 Data and data analysis

Psychologists, like all scientists, often produce numerical results from their investigations. These results are called the 'raw data'. As it is difficult to gain an understanding from large amounts of figures, the results are often simplified mathematically so they can be represented visually on graphs. This makes the meaning of the results easier to interpret. We will discuss a range of analytical techniques in this section. It is important that you are confident that you can count up scores, find the mode or the range of a data set, make simple comparisons and interpret data from tables or graphs.

Types of data

As you may know from the core studies or from earlier parts of this chapter, psychologists use a variety of different research methods. These methods can produce a range of different types of data. The main types are quantitative and qualitative data. When psychologists collect data, they can collect either numerical results, called **quantitative data**, or they can collect **qualitative data**, which is detailed and descriptive.

KEY WORDS

quantitative data: numerical results about the amount or *quantity* of a psychological measure, such as pulse rate or a score on an intelligence test

qualitative data: descriptive, in-depth results indicating the *quality* of a psychological characteristic, such as responses to open questions in self-reports or case studies and detailed observations

Quantitative data

Quantitative data indicates the *quantity* of a psychological measure, such as a total or frequency. Many psychological variables are measured as the

strength of a response and tends to be measured on scales, such as time or ratings. Alternatively, quantitative data can be a numerical score on a test such as for IQ or personality. Quantitative data is associated with experiments and correlations which use numerical scales but it is also possible to obtain quantitative data from observations, questionnaires or interviews. For example, a record of the number of times a behaviour is seen or the total number of different responses to a closed question in an interview would be quantitative data. The sources of quantitative data are typically highly objective, as the scales or questions used need little if any interpretation making them high in validity. In addition, the measures used are generally highly reliable, as the measures are fixed quantities.

Qualitative data

Qualitative data indicates the *quality* of a psychological characteristic. Such data is more in-depth than quantitative data and includes detailed observer accounts and responses to open questions in questionnaires, interviews or case studies. Although there is a risk of subjectivity in the interpretation of such data by the researcher, qualitative data may be more representative as the participant can express themselves fully, so in some senses qualitative data can also be valid. Table 1.4 will help to remind you of the key advantages and disadvantages of quantitative compared to qualitative data.

Data type		
	Quantitative	Qualitative
Strengths	Typically uses objective measures. Scales or questions often very reliable. Data can be analysed using measures of central tendency and spread making it easy to compare.	Data is often valid as participants can express themselves exactly rather than being limited by fixed choices. Important but unusual responses are less likely to be ignored because of averaging.
Weaknesses	Data collection method often limits responses so the data are less valid, e.g. if the participant wants to give a response that is not available.	The data is often relatively subjective so findings may be invalid as data recording or interpretation may be biased by the researcher's opinions or feelings. If detailed data can only be gathered from one or a few individuals it may not generalise to the majority.

Table 1.4: Quantitative versus qualitative data

RESEARCH METHODS

Dr Singh and Professor Chao disagree over the best way to collect data about people's emotions when their personal space is invaded. Dr Singh wants to collect *quantitative data* by measuring their pulse rate and give them scales of *very / quite / not at all* to rate how stressed they feel. Professor Chao thinks it would be better to collect *qualitative data* by interviewing each participant and getting them to describe their feelings. She plans to ask questions such as 'How do you feel if you think someone is standing too close to you?' and 'Describe how you felt the last time you had a professional very close to you such as a doctor, dentist or optician'.

ACTIVITY 1.13

Look at the Research Methods box. What are the advantages and disadvantages of each suggested data collection method?

- 1 Are the questions suggested by Professor Chao open or closed questions?
- 2 Using ethical ideas only, suggest two more questions that could be used in the *interview*, **one open question** and **one closed question**. State which question format would be used to collect *qualitative data* and which would collect *quantitative data*.

Descriptive statistics

This section explores different ways that mathematics can be used to simplify and understand the data produced by studies in psychology. Studies produce 'raw data', the original information from the participants, which may be quantitative or qualitative. This section discusses how quantitative data is analysed. Studies often produce a great deal of raw data, and it is important to summarise it, simplify it and interpret it. There are several ways to do this, some of which—such as inferential statistics—are beyond the scope of this

book. Here, we will focus on 'descriptive statistics', that is, ways to simplify and describe the data so that it is easier to summarise, compare and understand what the data show.

The first step in this process is often to produce a summary table. This might be the totals from a tally chart in an observation, percentages from a questionnaire or summary information about averages and the spread of data, such as described next. Whenever a table is produced it should have a title, and the rows (going across) and the columns (going down) should always have headings that indicate what the figures are about. For an experiment, these are likely to be the measure(s) of the DV and the levels of the IV. If figures, such as scores on the DV, are measured in units (seconds, Volts etc) the units should appear just once, in the heading. They should not be repeated in each 'cell' ('square' in the table).

This advice will help you to draw tables correctly but will also help you to interpret the data they contain. You can, for example, compare data between columns, or between rows. Imagine a study that compared A level students and degree level students on several measures, such as the time they spent studying or socialising and how stressed they felt. The information in Table 1.5 would enable you to make different comparisons. You could compare across columns, e.g. to see which group of students spent the most time working, or were most stressed or compare the rows, e.g. to see that A level students spend more time working than socialising.

Student group		
	A Level	Degree Level
Average time spent working per week (hours)	6	14
Average time spent socialising per week (hours)	4	20
Average stress level	10	11

Table 1.5: A summary data table for a study about student time use and stress



Measures of central tendency

A set of quantitative results can be summarised to one number that represents a 'middle' or typical score, called a **measure of central tendency** or 'average'. There are three different measures of central tendency: the **mode**, **median** and **mean**.

KEY WORDS

measure of central tendency: a mathematical way to find the typical or average score from a data set, using the mode, median or mean

mode: the measure of central tendency that identifies the most frequent score(s) in a data set

median: the measure of central tendency that identifies the middle score of a data set, which is in rank order (smallest to largest). If there are two numbers in the middle, they are added together and divided by two

mean: the measure of central tendency calculated by adding up the values of all the scores and dividing by the number of scores in the data set

The mode

The mode is the most frequent score in a data set. It can be used with numerical data, such as scores on a

test, and also with data sets that are choices that can be counted. Data that can be counted to produce a frequency also includes results in separate (discrete), named categories, such as in response to a question 'What is your favourite subject: maths, English or psychology?'). Importantly, the mode is the *only* measure of central tendency that can be used on data in discrete categories. However, although the mode considers all the scores in a data set, it doesn't consider their values, so it is less informative than the median or mode.

If two (or more) responses are equally common there will be two (or more) modes. For example, the faces in Figure 1.15 could be used in a test to compare people with a diagnosis of autism spectrum disorder with a control group. The participants could be asked 'Which face looks the happiest?' The face that was chosen most frequently as the happiest in each group would be the mode for that group.



Figure 1.15: A facial expression test

A self-report in a school produced the data set in Table 1.6. The mode for subject choice is 'Psychology', because more people said this subject was their favourite: ten compared to four and six for the other subjects.

	Subject		
	Maths	English	Psychology
Number of people	4	6	10

Table 1.6: Number of people choosing each subject as their favourite

	Night of the week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Number of boys	0	1	2	3	3	0	11
Number of girls	1	1	2	4	6	0	6

Table 1.7: Main homework night

Another question asked ‘On which day of the week do you do most homework?’ (see Table 1.7). The responses from girls and boys were compared.

In Table 1.7 more boys have ‘Sunday’ as their main homework night, so this is the mode for boys. Girls have said that ‘Friday’ and ‘Sunday’ are their main homework nights, but these two categories are the same, so there are two modes, each containing six girls. We could also combine the totals to work out an overall mode. Adding together the totals for each day shows that for all students, Sunday is the most popular homework night, with 17 students in total.

The median

Unlike the mode, the median cannot be used with data in separate (discrete) categories, it is only used with numerical data on a linear scale (i.e. points in a sequence). To find the median, all the scores in the data set are put in a list from smallest to largest (ranked). The middle one in the list is the median. If there is an even number of participants, so there are two numbers in the middle, these are added together and divided by two to find the median. The median is more informative than the mode because it considers the rank order of the values of scores in a data set rather than just how many there are. It also is uninfluenced by a small number of outlying scores. This can be a benefit, in that it is less swayed by a small number of extreme scores, but on the other hand, it does not take these values into account so is less representative than the mean.

Another question in the school survey asked participants to rate how hard they thought they worked, from 1 to 10. Students in their AS and A Level years were asked.

AS student data:

8, 6, 9, 1, 5, 6, 2, 7, 3, 6, 9, 8, 5, 6, 3, 8, 5, 10, 2, 3

A Level student data:

7, 9, 6, 7, 9, 7, 10, 10, 7, 10, 9, 4, 9, 6, 10, 10, 7, 9, 7, 7

Putting these data into order for the two groups separately:

AS students:

1, 2, 2, 3, 3, 3, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 8, 9, 9, 10

There are two numbers in the middle: $6 + 6 = 12$, $12/2 = 6$ so the median = 6

A Level students:

4, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10

There are two numbers in the middle: $7 + 9 = 16$, $16/2 = 8$ so the median = 8

The median for the A Level group, of 8, is higher than the median for AS students, which is 6. This suggests that A Level students believe they are working harder than AS students do.

The mean

The mean is the measure of central tendency that we usually call the ‘average’. It can only be used with numerical data from linear scales. The mean is worked out by adding up the values of all the scores in the data set and dividing by the total number of scores (including any that were zeros). The mean is the most informative measure of central tendency because it considers the values of each of the scores in a data set (rather than just the rank order of those values, as the median does), so can be influenced by outlying scores. This can be a problem, as the mean can be swayed by a small number of extreme scores but, on the other hand, by taking these values into account the mean is more representative than the median or mode.

Looking back at the data used in the section on the median, mean could be calculated instead. There were 20 students in each group. For the AS students, the calculation is therefore all the scores added together, then divided by 20, i.e. $112/20 = 5.6$, so the mean is 5.6. For the A Level students, the calculation is again all the scores added together, and divided by 20, i.e. $160/20 = 8$, so the mean is 8. Like the median, this too shows that the A Level students believe they are working harder than the AS students do.

Measures of spread

A **measure of spread** is an indicator of how varied the results are within a data set: are they clustered together or widely dispersed? If two data sets are the same size, with the same average, they could still vary in terms of how close most data points were to that average. Differences such as this are described by measures of spread: the **range** and the **standard deviation**.

The range

The range is the simplest measure of spread and is calculated in the following way:

- 1 Find the largest and smallest value in the data set.
- 2 Subtract the smallest value from the largest value, then add 1.



You may have learned to calculate the range without adding 1. In psychology we do this because the scales we use measure the gaps between points, not the points themselves. Consider a scale of student happiness from 1 (sad) to 8 (very happy). This can be represented on a line (see Figure 1.16):



Figure 1.16: Happiness scale

If we say someone's happiness is at level of 3, they could be anywhere between 2.5 and 3.5, and someone scoring 6 has a happiness level somewhere between 5.5 and 6.5. So, if 3 and 6 were the lowest and highest scores, the real spread extends to those limits, i.e. from 2.5 to 6.5, so this spread works out as $6.5 - 2.5 = 4$. This figure is one bigger than the largest score (6) minus the smallest score (3) = 3.

The range for the two sets of data given in the section on the median would be calculated in the following way:

AS students:

1, 2, 2, 3, 3, 3, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 8, 9, 9, 10

$10 - 1 = 9$, $9 + 1 = 10$ so the range = 10

A Level students:

4, 6, 6, 6, 7, 7, 7, 7, 7, 7, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10

$10 - 4 = 6$, $6 + 1 = 7$ so the range = 7

So, not only are the medians and means for these two data sets slightly different, but the ranges are also different. This tells us that the diversity of opinion about how hard they are working is greater for AS students than A Level students. We could also say that although most A Level students think they are working very hard, AS student opinion varies from 'not working hard' to 'working very hard'.

One problem with the range is that it does not accurately reflect outliers. That is, it would not be clear from the range whether the most extreme scores, e.g. very large scores, were single odd scores or typical of the data set. Imagine that the least hard-working student in the A Level set had rated themselves as working at level 1 rather than 4. This would make very little difference to the mean (it would be 7.85 instead of 8), but would change the range from 7 to 10 (the same as the range for the AS group).

The standard deviation

In the same way that the mean can tell us more than the mode, a measure of spread called the standard deviation can tell us more than the range. Rather than looking only at the extremes of the data set, the standard deviation (given the name s , SD or σ) considers the difference between each data point and the mean. This is called the *deviation*.

KEY WORDS

measure of spread: a mathematical way to describe the variation or dispersion within a data set

range: the difference between the biggest and smallest values in the data set plus one (a measure of spread)

standard deviation: a calculation of the average difference between each score in the data set and the mean. Bigger values indicate greater variation (a measure of spread)

As the standard deviation tells us the spread of values of scores in a data set, sets with scores that are more spread out have larger standard deviations, while sets with closely clustered scores have smaller standard deviations. When the standard deviations of two groups of participants' scores are similar, this means that they have a similar variation around the mean.

Returning to the data about how hard students believe they are working (from the section on the median) the standard deviations for the two sets of data would be:

AS students: SD = 2.62

A Level students: SD = 1.72

These figures show, as the range did, that there is greater variation in the opinions of the AS students as the figure of 2.62 for the standard deviation for this group is larger than 1.72, the figure for the A Level group. Note that if the standard deviation is calculated for the A Level group with the opinion score of 4 for one student being replaced with a 1, the standard deviation increases, as this makes the group more varied. However, it still does not become as large as the figure for the AS group. This is different from the effect of this same change in one participant's score on the range. The range of the two year groups become the same (they would both be 10), whereas in fact *within* the groups, the AS students are more varied. This is reflected in the standard deviation

however, which changes from 1.72 when the student has a score of 4 to 2.16 if they have a score of 1. So, the main advantage of the standard deviation over the range is that it takes every score into account and therefore provides a representation of variation *within* the data set. As it is not just considering the extremes is not distorted by outliers like the range would be.

RESEARCH METHODS

A new psychology teacher, Jamal, has tested a group of ten of his students on their recall of the aims, methods, results and conclusions of one of the core studies. The test was marked out of 15. Two students scored 0, and one student scored each of the marks 3–10. Jamal works out the *mode* and is horrified to find that the modal score is zero! Jamal's colleague Nurul recommends he works out the *median* instead, so he finds the two middle scores in the group (5 and 6), adds them together and divides by 2. This is 5.5, so Jamal feels much better. He wonders if calculating the *mean* will make him feel better still, but it doesn't. When Jamal adds up all the scores and divides by 10 (because he had ten students), the mean is only 5.2. This is because the median does not consider the absolute value of all the scores, only their relative values. So the two zero scores do not affect the median because they are simply 'lower than the central values in the rank order'. However, for the mean, the absolute values of the zeros, being the smallest values possible, do affect the mean. These three figures are different *measures of central tendency*. Jamal wants to compare this group to his colleague Nurul's students, who scored 0, 1, 1, 2, 2, 3, 4, 4, 5. He calculates two different *measures of spread*. The biggest and smallest values in her set are 0 and 5, so he subtracts 0 from 5 and adds 1, giving a *range* of 6. The *standard deviation* for her group is 1.58.

ACTIVITY 1.14

- 1 Look back at Figure 1.15. For the data set in the table shown, which is the modal response?

	Face A	Face B	Face C	Face D
Number of participants selecting each face as the happiest	0	3	7	1

- 2 Look back at the data for AS and A Level students in the section on the median. What is the mode for each group?
- 3 Using the information about Jamal and Nurul's students in the Research Methods box, answer the following questions:
 - a What is the mode for Nurul's group?
 - b Which group does this suggest performed better on the test?
 - c What is the median for Nurul's group?
 - d Which group does this suggest performed better on the test?
 - e The mean for Nurul's group was 2.4. Which group does this suggest performed better on the test?
 - f What is the range for Jamal's own group?
 - g When compared to the range for Nurul's group, what does this tell you?
 - h The standard deviation for Jamal's group was 3.49. Does this tell you that his group was more varied or less varied than Nurul's?

Graphs

A graph is a visual illustration of data. There are many different types of graph and in this section we will consider only bar charts, histograms and scatter graphs, each of which you should understand well enough to be able to name, recognise, draw and use to interpret data.



Bar charts

A **bar chart** is used when the data is in separate categories rather than on a continuous scale. Bar charts are therefore used for the totals of data collected in named categories and for all measures of central tendency (modes, medians or means). To construct a bar chart, the bars must be separate, with each bar representing a total or average of a category of data. This is because the *x*-axis represents the distinct groups and not a linear scale. For a bar chart of the results of an experiment, the levels of the IV go along the bottom (on the *x*-axis) and the DV goes on the *y*-axis. To help you to remember which is the *x*-axis and which is the *y*-axis, think 'X is a-cross' (Figure 1.17).

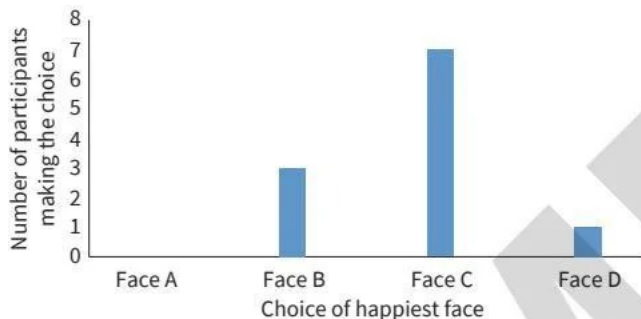


Figure 1.17: A bar chart showing the number of participants selecting each face as the happiest

Histograms

Histograms can be used to show the pattern in a whole data set, where this is continuous data, i.e. data measured on a scale rather than in separate categories. A histogram may be used to illustrate the distribution of a set of scores. In this case, the DV is plotted on the *x*-axis (across) and the frequency of each score plotted on the *y*-axis (up the side). The scores along the *x*-axis may be grouped into categories (e.g. if the DV is age, the data may be grouped into 0–5 years, 6–10 years, 11–15 years, etc.). To construct a histogram, each bar, e.g. representing a value on the scale of the DV, is plotted at the height representing the frequency of that value. As the scale represented on the *x*-axis is continuous the bars are drawn next to each other, unlike in a bar chart. This means that if there are no scores in a category, a gap must be left to show that the category is empty (see Figure 1.18, which shows 'empty' categories on the *x*-axis).

KEY WORDS

bar chart: a graph used for data in discrete categories and total or average scores. There are gaps between each bar that is plotted on the graph because the columns are not related in a linear way

histogram: a graph used to illustrate continuous data, e.g. to show the distribution of a set of scores. It has a bar for each score value, or group of scores, along the *x*-axis. The *y*-axis has frequency of each category

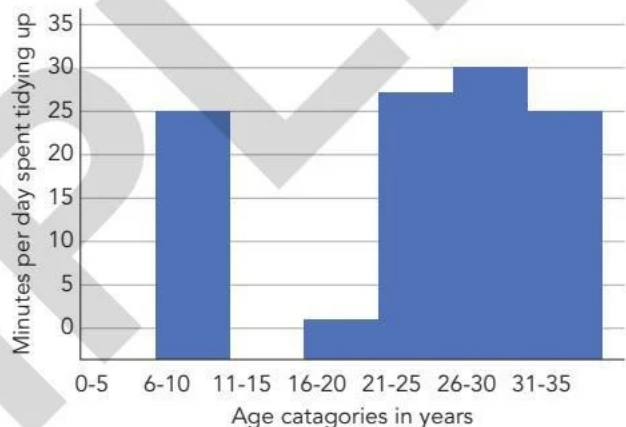


Figure 1.18: A histogram of the time spent tidying by people of different ages

Scatter graphs

The results from a correlational study are displayed on a scatter graph. To construct a **scatter graph**, a dot is marked at the point where an individual's scores on each variable cross. Sometimes you will see a 'line of best fit' drawn on a scatter graph. The position of this line is calculated and its line is drawn so that it comes close to as many points as possible (see Figure 1.19a–d). In a strong correlation all the data points lie close to the line, but in a weak correlation they are more spread out. Note that you will often see the strength of a correlation described as a number (and '*r*' value) from +1 to –1. Values close to +1 are strong positive correlations and values close to –1 are strong negative correlations. Lower or 'smaller' values (closer to 0) are weaker correlations. Where there is no significant correlation, the points do not form a clear line (and this has a '*r*' value of 0 or very close to zero).

KEY WORD

scatter graph: a way to display data from a correlational study. Each point on the graph represents the point where one participant's score on each scale for the co-variables cross

It is important to remember that you *cannot* draw a causal conclusion from a correlational study. Therefore, scatter graphs such as the ones in Figure 1.19 only tell you that there is a relationship between the co-variables but not which (if either) of them, is the cause of this link. An experiment could help to find this out.

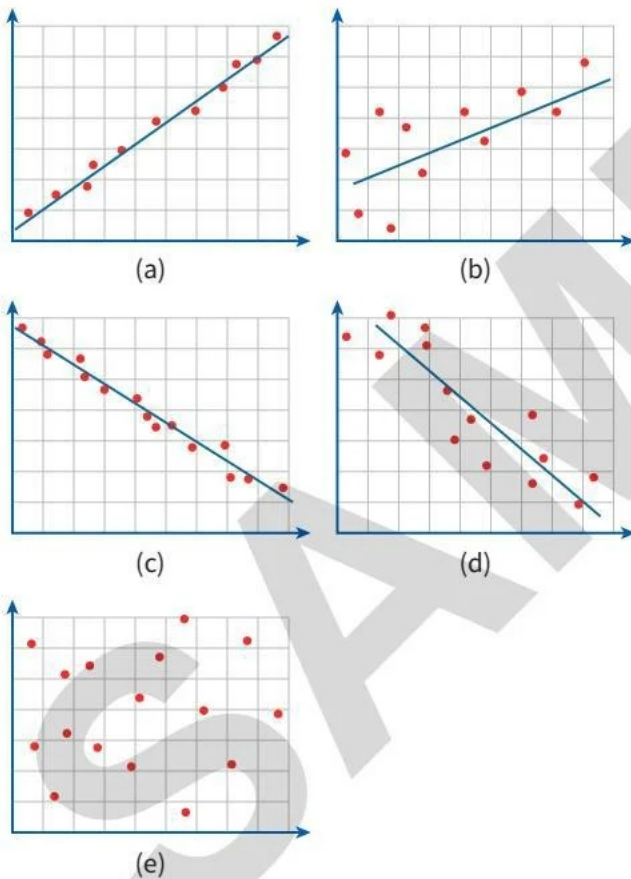


Figure 1.19: Scatter graphs showing (a) strong positive correlation, (b) weak positive correlation, (c) strong negative correlation, (d) weak negative correlation and (e) no correlation

RESEARCH METHODS

Groups of students in Dr Huang's class are doing different psychological investigations into altruism. Group A is counting the number of times boys and girls in the class say 'thank you'. They find that on average, the boys said thank you six times and the girls said thank you seven times. They are going to plot a *bar chart* of these averages. Group B is asking everyone the question, 'Which of these situations have you ever helped in: a parent with a pram / a elderly person with their shopping / a child crossing the road / an injured person?' The totals for each group are 14, 7, 16 and 3, respectively. They are also going to plot a *bar chart*. Group C is using an online computer game to test how many seconds it takes each student to respond to another player's request for help. They work out the percentage of students who took 0–5, 6–10, 11–15, 16–20 and 21–25 seconds. The results are 0%, 15%, 45%, 20% and 5%. They plot a *histogram* of their results. Group D asks every person to answer two questions 'How many times have you lent your notes to someone else?' and 'How likely on a scale of 1–10 are you to give away your lunch to somebody who is hungry?' Their results were:

Lend notes	5	3	9	4	2	6	9	7	3	4
Give lunch	7	2	7	10	1	8	8	7	4	5

They are correlating the two sets of scores and will draw a **scatter graph**.

ACTIVITY 1.15

Look at the Research Methods box. If you can work in a pair, each plot one bar chart (from either Group A or Group B's results). In addition, plot one other graph from the results of Groups C and D.

Peer assessment

In your pairs, look at the two graphs plotted by your partner. Consider the following points:

- Has the correct type of graph been chosen?
- Does the graph have a title?
- Is the x-axis labelled?
- Are there units or categories marked on the x-axis?
- Is the y-axis labelled?
- Are there units or categories marked on the y-axis?
- Are all the bars or points correctly plotted?

If you think any of these points are missing or could be better, suggest to your partner how their graph could be improved.

- 17 Ethelred has collected data by measuring two variables, hours of sleep and happiness. He has drawn a scatter graph of his data and found a negative correlation. Sketch a graph to illustrate the pattern of Ethelred's data.

1.10 Ethical considerations

As you will see from the discussion of the core studies, psychologists need to consider **ethical issues** when they conduct research. It is important that they take steps to make sure that their research follows ethical guidelines. In this section, we will look at some of these ethical issues and consider the ethical guidelines that help psychologists to deal with these issues effectively.

Ethical issues

As you will have seen from examples of psychological research, investigations using humans or animals have the potential to cause concerns about the welfare of the participants. Such concerns are called ethical issues. Problems may arise through the nature of the study, such as the potential for psychological discomfort caused by a study about stress, or from aspects of the procedure, such as the need to hide the real aim of the study. Ethical issues may also arise from the implications of the research, for example the possibility for results having a negative impact on part of society.

To help psychologists to cope with potential ethical issues that could arise in their research, many countries have an organisation that produces a code of conduct. In addition, research that is being conducted at a university is likely to require approval from the institution's ethical committee. An ethical code provides advice, for example as a set of **ethical guidelines**, that helps psychologists to work in a way that satisfies the primary concern of the welfare of individuals involved in the research as well as the perception of psychology in society. Participants who are deceived or distressed may not want to participate again, may view psychology badly and pass this message on to others, and are less likely to trust the findings of psychological research. These are all outcomes that should be avoided.

Questions

- 14 Read the data from the histogram in Figure 1.18. Draw a suitable table into which you can put this data.
- 15 Look back at the beginning of the section on descriptive statistics. Plot a bar chart for the data about 'average time spent working per week' for A Level and degree level students.
- 16 Dorothy has conducted a study looking at the development of language in a sample of ten infants. She has worked out the average number of different words the infants could say each month from the age of 1 year to 2 years old. The number of words varies from 0 to 60. At 18 months the ten infants could say 3, 45, 32, 14, 21, 18, 38, 29, 14 and 39 words.
- Work out the modal number of words the babies could say at 18 months.
 - Work out the range of the number of words the babies could say at 18 months.
 - Describe how the median number of words the babies could say at 18 months would be calculated.
 - Describe how the mean number of words the babies could say at 18 months would be calculated.
 - Draw and label the axes for a histogram for Dorothy's data.

KEY WORDS

ethical issues: problems in research that raise concerns about the welfare of participants (or have the potential for a wider negative impact on society)

ethical guidelines: pieces of advice that guide psychologists to consider the welfare of participants and wider society

Ethical guidelines relating to human participants

The discussion that follows is based on the British Psychological Society Code of Ethics and Conduct (2018), although there are many other similar ethical codes in use throughout the world. All of these guidelines help to contribute to the objective of minimising harm and maximising benefit, which is one of the key principles of the Code of Ethics and Conduct.

Protection from harm

Sometimes research carries risks for participants. Although this should be minimised in the planning of the study, some level of risk may be inevitable. A study may have the potential to cause participants psychological harm (e.g. embarrassment, self-doubt or stress) or physical harm (e.g. engaging in risky behaviours or receiving injections). Participants in such studies have the right to be protected and should not be exposed to any greater risk than they would be in their normal life. Care should be taken to eliminate such risks (e.g. by screening participants), experienced researchers should be used and studies should be stopped if unexpected risks arise.

Valid consent

Sometimes it is important in experiments to hide the aims from participants to reduce demand characteristics. However, participants have the right to know what will happen in a study so they can give their informed consent. The researcher's need to hide the aim makes it hard to get genuine consent. Ideally, valid consent

should be obtained from participants before the study starts. To do this, the researcher must get agreement or permission from potential participants by giving them sufficient information about the procedure for them to decide whether they want to participate, so that their decision is *informed*. This is often possible without needing to misinform i.e. lie to or otherwise mislead participants. Participants can be told what will happen to them in the study and any possible risks so that the participants can consent to the procedure without the researchers telling them the full aim. This can satisfy to needs of ethical requirements without leading to demand characteristics causing a change in the participants' behaviour.

For consent to be valid it must not only be informed but must also be freely given by a competent individual. The researcher must be sure that the participant understands what is being asked of them and feels that they have a choice about participating. This may be difficult for some groups of potential participants, such as people with mental health problems, learning difficulties or a low level of literacy, people who are working in an additional language or have amnesia, and for children. Furthermore, researchers should ensure that participants do not feel pressured to participate, such as prisoners may under some circumstances, or if the process of recruitment of participants is inappropriately persuasive.

When working with child participants, it may be difficult to obtain valid consent in the same way as with adults. It is therefore important to ask their parents or guardians for consent, but also the children themselves where possible. In this case, consent should be requested in a 'child friendly' way that is suitable for the level of understanding the child has.

In some situations, it is not even possible to ask for consent. This is often the case in naturalistic observations and field experiments. In such situations, a researcher may ask a group of people similar to those who will become participants whether they would find the study acceptable if they were involved. This is called **presumptive consent** because it allows the researcher to presume that the actual participants would also have agreed to participate if asked. Especially when participants have not been fully informed, it is important to **debrief** them at the end of the study.

KEY WORDS

presumptive consent: this can be obtained when informed consent cannot be obtained from actual participants. A similar group of people to the anticipated sample are given full details of the proposed study and asked if they would find the study acceptable or not. If they would be happy to be involved, the study can continue

debriefing: giving participants a full explanation of the aims and potential consequences of the study at the end of a study so that they leave in at least as positive a condition as they arrived

Right to withdraw

Participants should be able to leave a study whenever they wish. This is their right to withdraw and it must be made clear to participants at the start of the study. Although participants can be offered incentives to join a study, these cannot be taken away if they leave. This prevents participants thinking that they have to continue. Researchers should not use their position of authority to encourage participants to remain in a study if they want to stop. So in practice, participants may need to be reminded of this right and researchers should follow this guideline even if data will be lost.

Lack of deception

Participants should not be deliberately misinformed, i.e. deception should be avoided. When it is essential to deceive participants, e.g. to avoid the effects of demand characteristics, they should be told the real aim as soon as possible and be allowed to remove their results if they want to. When participants have been deceived and they know they have been in a study, debriefing (see next) should follow immediately.

Confidentiality

All data should be stored separately from the participants' names and personal information held, and names should never be published unless the individuals have specifically agreed to this. Such information should be stored securely and should not be shared with anyone outside the study. These measures ensure confidentiality. The identity of participants should be protected by destroying personal information. However, where it is needed to re-contact participants or to pair up an individual's scores in each condition in a repeated

measures design, each participant can be allocated a number that can be used to identify them.

When conducting a case study or field experiment with institutions, confidentiality is still important and identities must be hidden. For example, the names of schools or hospitals should be concealed. One common way to anonymise individuals is to use their initials. However, where an individual is well known, this may not be sufficient to conceal their identity. For example, one participant in the study about sleep and dreams by William Dement and Nathaniel Kleitman had a participant with the initials WD.

The only exception to this is that personally identifiable information can be communicated or published when the participant gives their informed consent for this or in exceptional circumstances when the safety or interests of the individual or others may be at risk.

Privacy



Figure 1.20: Privacy should still be maintained even if consent cannot be given by participants

Observations, self-reports that ask personal questions, and any study that uses personal information all risk invading privacy. This means that the researcher may enter physical space or emotional territory that the individual would want to keep to themselves. In a questionnaire, interview or case study, participants should be aware of their right to ignore questions they do not want to answer. When completing a questionnaire in a laboratory situation, participants should be given an individual space. In observations, people should only be watched in situations where they would expect to be on public display, see Figure 1.20.

Debriefing

All participants who are aware that they have been in a study should be thanked and given the chance to ask questions. Debriefing participants provides them with an explanation at the end of the study that explains fully the aims of the study and ensures that they do not want to withdraw their data. If participants have been negatively affected by a study the researcher must return them to their previous condition. However, debriefing is not an alternative to designing an ethical study, so it is important to consider all the ways in which a study could cause distress and to minimise them.

Ethical guidelines relating to the use of animals

Animals are used in psychological research for a number of different reasons. Driscoll and Bateson (1988) suggested animals may be: convenient models (e.g. for processes such as learning), a way to carry out procedures that could not be done ethically on humans (e.g. isolation or brain surgery) or be good or interesting examples in their own right (e.g. communication in birds, bats or whales). As a consequence, much psychological research is conducted on animals and therefore their welfare needs protecting.

The discussion that follows is based on the British Psychological Society Guidelines for Psychologists Working with Animals (2020), although there are many other similar ethical codes in use throughout the world. Animals are also often protected by law. These guidelines encompass appropriate legal requirements and specifically consider the effects of psychological research in which animals may be confined, harmed, stressed or in pain, so suffering should be minimised. Veterinary advice should be sought in any case of doubt.

Researchers must aim to ensure that in any research, the means justify the ends, i.e. that the animal suffering caused by the planned experiment is outweighed by the benefits. One way to consider this question is to use Bateson's (1986) cube (see Figure 1.21). When the certainty of benefit (e.g. to humans) is high, the research is good and the suffering is low, the research can be argued to be worthwhile. The guidelines below help to contribute to the objective of minimising harm and maximising benefit.

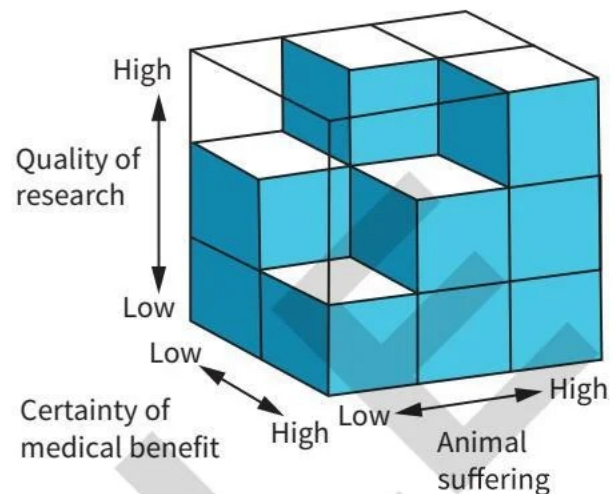


Figure 1.21: Bateson's cube (1986). According to Bateson, research should only be conducted if it falls within the clear part of the cube

Replacement

Researchers should consider replacing animal experiments with alternatives, such as videos from previous studies or computer simulations.

Species

The chosen species should be the one least likely to suffer pain or distress. Other relevant factors include whether the animals were bred in captivity, their previous experience of experimentation and the sentience of the species (its ability to think and feel).

Number of animals

Only the minimum number of animals needed to produce valid and reliable results should be used. To minimise the number, pilot studies, reliable measures of the dependent variable, good experimental design and appropriate data analysis should all be used.

Procedures

Research on animals is controlled by legal requirements and/or guidelines from relevant organisations because studies will, necessarily, affect the animals in some way. However, for research on animals to be effective the animals' experience should, as far as possible within the constraints of the research being conducted, be a normal and positive one. There are therefore guidelines to help psychologists to care for their research animals well.

Pain, suffering and distress

Research causing death, or suffering, such as disease, pain, injury, physiological or psychological distress and discomfort should be avoided. Where possible, designs which improve rather than worsen the animals' experience should be used (e.g. studying the effect of early enrichment on development compared to normal rather than early deprivation). Alternatively, naturally occurring instances may be used (e.g. where stress arises naturally in the animal's environment or lifetime). During research, attention should be paid to the animals' daily care and veterinary needs and any costs to the animals should be justified by the scientific benefit of the work (see Bateson's cube).

Housing

Isolation and crowding can cause animals distress. Caging conditions should depend on the social behaviour of the species (e.g. isolation will be more distressing for social animals than solitary ones). Overcrowding can cause distress and aggression (therefore also physical harm). The level of stress experienced by individuals should also be considered (e.g. the animal's age and gender). Between testing, animals should be housed with enough space to move freely and with sufficient food and water for their health and well-being, both in terms of their biological and ecological needs. However, the artificial environment only needs to recreate the aspects of the natural environment that are important to welfare and survival, e.g. warmth, space for exercise or somewhere to hide. Environments that are 'visually appealing' to people may not be those that are best for an individual species. Cage cleaning should balance cleanliness against avoiding stress, e.g. caused by unfamiliar smells or disturbance by humans.

Reward, deprivation and aversive stimuli

Deprivation is the removal of resources that are important to an animal. In planning studies using deprivation the normal feeding or drinking patterns of the animals should be considered so that their needs can be satisfied (e.g. carnivores eat less frequently than herbivores, young animals need greater access to food and water). The use of preferred food should be considered as an alternative to deprivation (e.g. for rewards in learning studies) and alternatives to aversive (unpleasant) stimuli and deprivation should be used where possible.

Ethical guidelines and their evaluation and application to research

All research with human participants should be ethical. Researchers should, for example, always follow the guideline of confidentiality. Although ethical guidelines should always be followed, sometimes it is necessary to accept some risks to participants in order for a study to achieve its objectives. When this is the case, the researchers must consider whether the risks can be justified and, if so, how they can be minimised. Ethical issues may arise because of the nature of the topic being studied, such as phobias, as this threatens participants' psychological well-being, or they may be at risk of physical harm. Alternatively, ethical issues may arise from the need to use controls which limit participants' knowledge or choices so threaten their privacy, informed consent or right to withdraw. This, in turn, has the potential to cause psychological harm. Although debriefing can help to reduce any harm that has been done, it is not a substitute for designing a more ethical study. Researchers must therefore consider all appropriate guidelines when planning a study, if necessary consulting with colleagues and following the advice of an ethical committee.

You will need to be able to make decisions based on ethical guidelines, and these can be applied to a study you already know about, a novel example that is presented to you or a study that you are asked to plan. So, in relation to informed consent, protection (physical and psychological), right to withdraw, deception, confidentiality, privacy and debriefing, you will need to be able to explain why issues have arisen and how and why each guideline has been broken or has been followed.

All research with animals should be ethical. However, for any good research to be conducted a compromise must be reached between animal suffering and the good that will come out of the research. In evaluating studies, it is important to remember that when the research is planned, decisions must be based on *expected* outcomes, i.e. the potential gains of the study. This is governed by the importance of the work, how certain it is that there will be a valuable benefit and this, in turn, is in part determined by the effectiveness of the research itself. So, in deciding whether a piece of research is ethical, we must think about:

- how much the animals suffer (e.g. in terms of pain, deprivation, distress)
- what the positive outcomes might be (e.g. the benefits for people and whether these are worthwhile)
- whether the research is sufficiently well planned to achieve these possible benefits.

You can consider these three elements in terms of Bateson's cube, illustrating each one with examples relating to the particular study. The same ideas can be used whether you are considering a study you already know about, a novel example that is presented to you or a study that you are asked to plan. So, in relation to the three points, you will need to be able to explain or justify:

- choices about the animals and their care, such as the species and how they are housed, and the procedures chosen, e.g. the number of animals used and the design of the study in terms of the aspects of feeding, access to companions or pain and distress that arise as a consequence of the study
- why the study is being done, for example in terms of the need for new research or the ways in which humans (or animals) could benefit from the findings
- the strengths of the design of the study, e.g. in terms of controls and ways that good objectivity, validity and reliability have been achieved.

RESEARCH METHODS

Professor Chao is planning an *experiment* on emotions in animals. She wants to see if, like us, non-human animals tend to approach things they like and avoid things they don't like. She considered *replacing* the use of animals with videos of animals in the wild responding to different stimuli but decided that she would not know for certain what they were reacting to so the findings might not be valid. She chose to use rats as the *species* is bred for laboratory use and rats can be *housed* alone without distressing them. She decided to use a *repeated measures design* as it would limit the *number of animals* used compared to an *independent measures design*. To reduce *pain and distress*, she used food they liked a lot (peanuts) in the 'approach' condition and food they did not like a lot (lettuce) in the 'avoidance' condition. Although the rats were tested before they were given fresh food in their cages, they were not *deprived* of food.

ACTIVITY 1.16

Look at the Research Methods box.

- Professor Chao considered using *aversive stimuli* such as loud noises or electric shocks in the 'avoid' condition but decided this was unethical. Why?
- Suggest one problem with Professor Chao's decision to use a *repeated measures design*, and explain how this problem might be solved.

Questions

- 18 Rashid is conducting a study to investigate people's helpfulness when they see someone who needs help to cross the road. Rashid is going to ask his friend Kinza to act as a stooge (confederate) and pretend to need help so that he can observe people's actions.
 - a Why could Rashid's study raise ethical issues?
 - b What could Rashid do to follow ethical guidelines and minimise these issues?
- 19 Dr Hadid is planning a study using laboratory mice to test their ability to learn to jump over a small wall.
 - a What could Dr Hadid do to ensure he follows animal guidelines relevant to his laboratory mice?
 - b What should Dr Hadid *not* do to ensure he follows animal guidelines relevant to his laboratory mice?

1.11 Evaluating research: methodological issues

As well as evaluating research in terms of ethics, it can also be considered in terms of whether it is 'good science', i.e. by looking at methodological issues. There are several key methodological issues that you have encountered elsewhere in the chapter, such as reliability, validity and generalisability. We will now explore these again in a little more depth and see why they, and **replicability**, are important in the evaluation of research.



Reliability

Whenever research is conducted data is obtained, researchers must attempt to ensure that the way in which these results are collected is the same each time, otherwise differences could occur (between participants, between conditions in an experiment or between the data obtained by different researchers). Such inconsistencies would be problems of reliability.

The reliability of the measures used to collect data depends on the 'tool' used. A researcher collecting reaction times or pulse rates as data will probably have reliability as the machines used are likely to produce very consistent measures of time or rates. One way to check reliability is to use the **test-retest** procedure. This involves using a measure once, and then using it again in the same situation. If the reliability is high, the same results will be obtained on both occasions, i.e. there will be a high correlation between the two sets of scores. Imagine an experiment on euphoria and anger in which a researcher is not sure whether their questionnaire is a reliable measure of 'happiness'. They use a group of participants and give them the questionnaire on two separate occasions. All the participants would need to be tested at the same time of day and the same day of the week to ensure that their happiness levels were indeed the same. If the 'happiness scale' was reliable, this test-retest procedure would produce a high correlation between the scores on the first and second tests. If the reliability was low, the test would need to be redesigned.

KEY WORDS

replicability: the extent to which the procedure of a study can be kept the same whenever the research is repeated. This is especially important when a study is repeated, either by the same or different researchers to verify results. It also enables researchers conducting other studies to follow exactly the same procedure to test different aspects of a problem, different participant groups etc.

test-retest: a way to measure the consistency of a test or task. The test is used twice and if the participants' two sets of scores are similar, i.e. correlate well, it has good reliability

Another reliability problem relates to subjective interpretations of data. For example, a researcher who is using a questionnaire or interview with open questions may find that the same answers could be interpreted in different ways, producing low reliability. If these differences arose between different researchers, this would be an *inter-rater reliability* problem. Similarly, if in an observation researchers gave different interpretations of the same actions, this would be low *inter-observer reliability*. If the reliability was low, the researchers in either case would need to discuss why the differences arose and find ways to make their interpretations or observations more similar. This can be done by agreeing on operational definitions of the variables being measured and by looking at examples together. These steps would help to make the researchers more objective.

To minimise differences in the way research is conducted that could reduce reliability, standardisation can be used, that is, the procedure is kept the same. This could include instructions, materials and apparatus, although remember that there would be no reason to change many of these. Important aspects of standardisation are those factors which might differ, such as an experimenter's manner towards participants in different levels of the IV, an interviewer's posture or tone in asking questions or an observer's success at concealing their presence.

Replicability

When a researcher conducts a useful study, they often publish their results and conclusions, along with clear details of how the study was conducted. The core studies are examples of such publications. One reason for the level of detail in these published studies is to allow other psychologists to repeat the research, i.e. to replicate it. This helps to ensure that researchers are honest about their work as it must be possible for others to reproduce their findings otherwise the research loses credibility. Furthermore, if this were not possible—and achieved—the impact on psychology in general would be to lower confidence in research as a whole, damaging the potential for a positive impact on society. It is therefore essential that details about how the study will be conducted are considered at the planning stage so that they can be rigorously followed by the researchers themselves, and by others. Such details include many of the key points in this chapter, such as the variables that were measured and manipulated and how they were operationalised; features of the sample and how this was

obtained; the procedure that was followed including, for example, controls, counterbalancing and allocation to groups; and the materials used.

Validity

Many factors affect validity, including reliability: a test or task cannot measure what it intends to measure unless it is consistent. Objectivity also affects validity: if a researcher is subjective in their interpretation of data, their findings will not properly reflect the intended measure.

To have face validity, a test or task must seem to test what it is supposed to. Imagine a test of helping behaviour that involved offering to assist people who were stuck in a bath full of spiders or worms. It might not be a valid test of helping because people who were frightened of spiders or worms would not help, even though they might otherwise be very altruistic. This would be a lack of face validity.

If participants think that they understand the aim of a study, their behaviour or responses are likely to be affected. This would also lower validity. In the design of a study, the researcher should aim to minimise demand characteristics, that is, those features which could indicate to the participants what is expected. For example, in a study about false memories, the researcher needs to hide the aim from the participants otherwise they will try to spot which memory is false. They might then try to remember this information particularly well, or might not report it at all if that is what they think the researcher expects.

Another problem for validity is whether the results obtained in one situation will apply to other situations. If not, then the test or task is too specific to be measuring the general phenomenon it was intended for. This is the problem of ecological validity, which applies more widely than considering just whether findings from the laboratory apply to the 'real world'. For example, a test of stress conducted in a laboratory may indeed not reflect the stress experienced in day-to-day life. But equally, a test of stress conducted at home might not reflect the stressful experiences people endure at work or during healthcare procedures. If so, the results may not generalise beyond the situation tested.

The task itself matters too. If the task participants are asked to do is similar to tasks in day-to-day life,

it has mundane realism. This is important as the findings are more likely to have high ecological validity if the task is realistic. For example, in an experiment on emotions, responses to dangerous polar bears or dangerous insects could be used. As fewer people would encounter polar bears, responses to insects are likely to have higher mundane realism and therefore higher ecological validity.

Generalisability

Ecological validity contributes to the generalisability of the results. Another factor that affects the ability to generalise is the sample. If the sample is very small, or does not contain a wide range of the different types of people in the population, it is unlikely to be representative. Restricted samples like this are more likely to occur with opportunity or volunteer samples than with a random sample.

Evaluating and applying concepts relating to methodological issues

You will need to be able to make evaluative decisions about methodology in relation to studies you already know about or a novel example of research or a study that you are asked to design yourself. In all of these cases, you need to ask yourself the following questions:

- Are the measures *reliable*? The study will collect data. Does the tool used to collect that data work consistently? Are the researchers using that tool in a consistent way? Is it *objective* or could they be *subjective* in their interpretation of the data?
- Is the study *valid*: Does it test what it is supposed to? This may depend on the reality of the task and the relevance of setting (*ecological validity*). Might the participants have been affected by *demand characteristics*?
- Are the findings *generalisable*: if not, why not? If so, how widely? This may depend on *ecological validity* or on the types and size of the sample.
- Could the research be *replicated*? What details would you need to know? Why might it be important to be able to do so?

You will also need to be able to suggest ways to improve the methodology used. You can consider improving the:

- *method* (e.g. a field versus a laboratory experiment or a questionnaire versus an interview)
- *design* (independent measures will have fewer problems with order effects but repeated measures could overcome issues with individual differences)
- *sample* (using opportunity sampling might allow a larger sample to be collected, volunteer sampling could help to find particular types of participant and random sampling would give better generalisability)
- *tool* (measuring the inter-rater reliability or test–retest reliability and changing procedures to make improvements)
- *procedure* (to raise validity by reducing demand characteristics, making the task more realistic, etc.).

RESEARCH METHODS

Dr Singh is planning an experiment on obedience. He wants to test whether drivers are more obedient to traffic wardens wearing white clothing or black clothing. He wants the test to be *valid*, so he uses the same male traffic warden wearing different clothing in each condition. He has four observers, one watching cars approaching a junction in his town from each direction. It is important that they are *reliable*, so he gives them operational definitions for the behavioural categories they are to observe:

- *Obedient behaviours*
 - slowing down: visibly reducing speed
 - stopping: coming to a halt before the line on the road.
- *Disobedient behaviours*
 - stopping late: coming to a halt past the line on the road
 - driving on: failing to stop when instructed to do so by the traffic warden.

CONTINUED

He aimed for these categories to be very *objective*. He believes that it is unlikely that the participants (the drivers) will respond to *demand characteristics* as they would not know that they were in an experiment. The drivers who stop past the line are given a note by a confederate while they are stationary. This *debriefs* them and asks if they would be happy to answer questions by telephone. Dr Singh's office number is given for them to call.

ACTIVITY 1.17

Look at the Research Methods box.

- Dr Singh is concerned about the *generalisability* of his findings. He has two ideas for changes to the procedure: conducting the same test in a village as well as his town and repeating the procedure using a female traffic warden in both conditions. Explain how each idea would improve generalisability.
- The proposed study has high *ecological validity*. Explain why.
- Dr Singh thinks that of the observations in one of the behavioural categories, visibly reducing speed could be subjective. Explain why this is likely.
- Dr Singh wants to measure the *inter-observer reliability* of his four observers. Explain why this is important.
- In the final part of the study, some participants find out that they have been in a study. Suggest one ethical problem that could arise from this.
- By giving the drivers a number to call, rather than taking their number and calling them, Dr Singh is giving the participants their *right to withdraw*. Why is this important?

CONTINUED

- Dr Singh asks the participants who do call him why they stopped and why they stopped over the line. He asks two of his colleagues to interpret the reasons they give but wants to ensure that they have high inter-rater reliability. He gives them a list of possible interpretations including a numerical scale to indicate how strongly the participant felt they may be punished. As both colleagues interpreted the responses from all the available drivers, Dr Singh can *correlate* the score given to each driver by the two colleagues to see if they are similar. What can he conclude if this produces a strong positive correlation?
- What other details would you need to know about this study to be able to *replicate* it in a town near you?

He is teaching the same course to two different groups. He gives the 'more frequent test' group a mini test every week and the 'less frequent test' group a test every two months.

- At the end of the course, Mr Nguyen assesses all the students with the same essay-based exam. His school expects pairs of teachers to 'double mark' all exams, so Mr Nguyen and his colleague both mark the exams from Mr Nguyen's classes using a detailed mark scheme. They mark one student's exam together to make sure that they are marking in the same way. Explain whether this procedure will improve reliability or validity.
- After his study is complete, Mr Nguyen discovers that the students in the 'more frequent test' group have been sharing their mini tests with the students in the 'less frequent test' group each week. Explain whether this will have been a threat to reliability or validity.
- Mr Nguyen hopes the study can be conducted again, the following year, asking students not to share the tests. However, he is also concerned about the generalisability of his findings. Suggest what changes could be made to improve generalisability.

Questions

- 20 A teacher, Mr Nguyen, is planning to conduct an experiment using his students as participants. He is aiming to investigate whether more frequent or less frequent testing of the students during his course produces better results at the end of the course.

SUMMARY

Psychologists can use several different research methods; experiments (laboratory and field), self-reports (questionnaires and interviews), case studies (detailed investigations of a single instance, e.g. one person), observations, correlations and longitudinal studies.

In experiments, there is an independent variable (IV), which is manipulated, changed or used to create different conditions and a measured dependent variable (DV). By imposing controls, the experimenter can be more certain that changes in the IV are the cause of changes in the DV. There are three experimental designs. In an independent measures design there are different participants in each level of the IV, in a repeated measures design the same

participants are used in all levels of the IV and in a matched pairs design the participants are paired up with one member of each pair in each level of the IV. In a repeated measures design counterbalancing helps to overcome order effects (fatigue and practice effects) and in an independent measures design random allocation helps to overcome the effects of individual differences. In experiments it is important to control variables to raise validity. The most important are confounding variables. If these are left as uncontrolled variables, they can alter the apparent effect of the IV on the DV. Uncontrolled variables can be described as participant variables (due to differences between individuals or between the same individual at different times) or situational variables (due to differences in physical setting or the social situation).



CONTINUED

In self-reports, different question formats can be used, including open questions (producing qualitative data) and closed questions (producing quantitative data). An interview can have different formats: structured, unstructured or semi-structured. Different techniques can also be used. 'Paper and pencil' or online for questionnaires and telephone or face-to-face for interviews. Observations have many different features, for example structured (observing known categories) or unstructured (recording any events) and naturalistic (observing whatever is happening) or controlled (constructing events to observe). The role of the observer may be obvious to the participants (overt) or hidden (covert) and the observer themselves may be part of the social situations (participant) or not (non-participant).

Correlations are used to look for relationships between two co-variables that are measured. The findings of correlational studies can produce a positive correlation (the two variables increase together) or a negative correlation (as one variable increases the other decreases) but conclusions cannot be drawn about causal relationships between the variables. All variables, e.g. those in correlations, the IV and DV in experiments and behavioural categories in observations should all be operationalised.

Any research begins with an aim, which is developed into a testable hypothesis. This can be directional (one-tailed) or non-directional (two-tailed). This is compared to a null hypothesis, which proposes that there is no difference or relationship (or that any pattern in the results has arisen due to chance). To test the hypothesis, a group of participants (the sample) is selected from the population. This can be done by opportunity sampling (choosing people who are available), random sampling (selecting participants so that each individual has an equal chance of being chosen) or volunteer (self-selecting) sampling (inviting participants, e.g. by advertising).

Studies can collect different types of data. Quantitative data is numerical and qualitative data is descriptive. Data analysis of quantitative data includes using various measures of central tendency

(the mean, median and mode) and measures of spread (the range and standard deviation). Data can be displayed graphically using bar charts, histograms or scatter graphs.

Research in psychology raises ethical issues. Some important issues relate to informed consent (knowing about the study and agreeing to do it), protection of participants (physically and psychologically), the right to withdraw (being able to leave a study), deception (being misled), confidentiality (keeping participants' data anonymous), privacy (not invading physical or mental space) and debriefing (explaining the study to participants afterwards and returning them to their previous state). There are also ethical guidelines relating to the use of animals, including issues relating to the species used, number of animals, the pain and distress they experience, the way they are housed and rewarded or deprived and their suffering.

Two very important methodological issues are validity and reliability. Ecological validity relates to how well the findings from one situation, e.g. a laboratory, represent what would happen in other situations. Subjectivity threatens validity because it causes researchers to interpret findings from their personal viewpoint, whereas objectivity allows researchers to measure variables in ways that are independent of their own perspective. Demand characteristics also threaten validity because they inform participants about the aim of the study which can alter their behaviour. Results of studies should be generalisable, that is they should apply to other people, situations and times. Reliability refers to the consistency of measures. In an experiment it is important to use standardisation of procedures to ensure that all participants are treated in the same way. This raises reliability. When researchers interpreting data are consistent, they have good inter-rater reliability (e.g. due to practice of operational definitions). Inter-observer reliability is the consistency in the records made by observers who are watching the same events. The reliability of a test, e.g. a questionnaire or a task in an experiment can be evaluated using a procedure to measure test-retest reliability, by conducting the test twice and correlating the two sets of data.

SELF-EVALUATION CHECKLIST

After studying this chapter, complete a table like this:

I can	Needs more work	Almost there	Ready to move on	Coursebook	Workbook
Describe and use the general concepts and terminology relating to research methods, including applying these to novel scenarios					
Describe, evaluate and apply experiments (laboratory and field)					
Describe, evaluate and apply self-reports (questionnaires and interviews)					
Describe, evaluate and apply case studies					
Describe, evaluate and apply observations					
Describe, evaluate and apply correlations					
Describe, evaluate and apply longitudinal studies					
Explain and use aims and hypotheses					
Explain and use variables, including operational definitions, uncontrolled variables, participant and situational variables and controls					
Explain and use types of data including qualitative and quantitative					
Explain and use samples, populations and sampling techniques					
Explain and use ethical issues and guidelines for humans and animals in psychological research					
Explain and use the concept of reliability					



CONTINUED

I can	Needs more work	Almost there	Ready to move on	Coursebook	Workbook
Explain and use the concept of validity					
Explain and use the concept of generalisability					
Explain and use descriptive statistics including measures of central tendency and spread, tables and graphs					

PROJECT

Plan a laboratory experiment and produce a written document containing:

- Your aim.
- Your IV and an operational definition for each level of the IV.
- Your DV with an operational definition and full details of how it will be measured.
- A statement of which experimental design you would use, and a justification of why you would choose this design.
- Details of controls you would impose, with a justification of why they are necessary.
- How you could, if you were to conduct the study, obtain a sample, including details of the sampling technique, the sample size and any special features of the sample.
- The materials that would be needed, such as any stimulus materials (such as word

lists or images), questions that would be asked in a questionnaire or interview and a description of the format if appropriate as well as a scoring system for questions producing quantitative data.

- A description of any ethical issues that could arise and the steps that would be taken to follow all relevant ethical guidelines.
- A table that could be used to collect the raw data.
- A table that could be used for descriptive statistics and an indication of what these would be, including the measure of central tendency you would use.
- The axes for a graph, with a title, axis labels and units that could be used to display the data.

Note that, unless you have explicit approval from your teacher, you must **NOT** conduct this experiment.

REFLECTION

Think back to how you planned your study. This is the same process that psychologists use in research. They work from an aim, to devising a study that tests their aim in an ethical way that will produce valid, reliable and generalisable results. Can you identify any steps you included in your plan that contributed to reliability? To validity? Or to generalisability? If not, think about ways that you could change your study to improve it.

EXAM-STYLE QUESTIONS

- 1 A hypothesis in a study says 'Greater mindfulness will be achieved after a period in mindfulness training than after a period of waiting'.
 - a Is this an alternative hypothesis or a null hypothesis? Explain your answer. [1]
 - b Is this a directional (one-tailed) hypothesis or a non-directional (two-tailed) hypothesis? Include a reason for your answer. [1]
- 2 Jun is conducting a self-report study about attitudes to people with phobias. He cannot decide whether to use a questionnaire or an interview.
 - a **Suggest** one advantage and one disadvantage of using a questionnaire for Jun's study. [4]
 - b Write one closed and one open question that Jun could ask. [2]
 - c Jun is concerned that his interpretation of the responses to questions might not be consistent. Is this mainly a reliability or a validity issue? **Explain** your answer. [2]
- 3 Priya is planning an experiment to find out whether boys or girls in her school doodle more.
 - a **Describe** how Priya could conduct her experiment. [10]
 - b Identify one possible weakness / limitation with the procedure you have described in your answer to part a and suggest how your study might be done differently to overcome the problem. [4]
- 4 A student is designing an experiment that aims to test whether parrots are more intelligent than cats. He has three parrots and two cats, which he plans to use as his sample. To find out which is most intelligent, he is going to hide their food bowl inside a box and time how long it takes the animal to get to the food.
 - a **Identify** and operationalise the independent variable in this experiment. [1]
 - b Identify and operationalise the dependent variable in this experiment. [1]
 - c Identify and outline the sampling technique used in this study. [2]
 - d Which measure of central tendency would be best to find out the average time taken to find the food? [1]
 - e A friend suggests that this is not a very valid test of intelligence because it might depend on how well the animal can smell the food. Explain this criticism. [2]

COMMAND WORDS

suggest: apply knowledge and understanding to situations where there are a range of valid responses in order to make proposals / put forward considerations

explain: set out purposes or reasons / make the relationships between things evident / provide why and/or how and support with relevant evidence

describe: state the points of a topic / give characteristics and main features

identify: name / select / recognise





CONTINUED

- 5 A study into sleep obtained participants by placing advertisements in shops near to the university. The participants who responded were a sample of nine females and one male and were mainly retired people. The study was testing a new way to help people to fall asleep, using a recording of bubbling stream. Half the participants were told it would help them to sleep, the others were told it would keep them awake.
- a Identify and outline the sampling technique used in this study. [2]
 - b Explain **one** possible problem with generalisability in this study. [2]
 - c Describe **one** ethical issue that would arise in this study. [2]
 - d How well the participants slept was measured in two ways, by self-report and by how many minutes they stayed asleep for. Which of these measures is more reliable and why? [2]
- [Total: 40]

Additional references

Holahan, C. K., & Velasquez, K. S. (2011). Perceived strategies and activities for successful later aging. *The international journal of aging and human development*, 72(4), 343-359. doi:10.2190/AG.72.4.d

