

# An Introduction to the Research Process

# 1

## Overview

In this chapter we will introduce you to concepts that are important for understanding the research process, including:

- Research hypotheses
- Hypothesis testing
- Evidence-based practice
- Typical research designs

We do not assume any prior knowledge of statistics or of conducting research. All that you need to understand the concepts in this chapter is your brain.

Brains at the ready here we go .... Today there was a report on the radio suggesting that eating more blueberries will reduce the chances of getting cancer. This is not an uncommon type of report in the media these days. How do we know whether to believe all of the reports relating to health that the media present to us? Well, the best thing to do is to read the original research reports and make up our own minds about the adequacy of the research itself, as well as the validity of the conclusions that authors have drawn from their research.

This is how science works. Of course if you wish to work as a health professional there is an even greater need for you to be able to evaluate research evidence. This book will provide you with all of the tools necessary to be able to critically evaluate the research of other health professionals. Not only that, you will gain a working knowledge of how to conduct your own research and how to run some quite sophisticated statistical analyses on your data.

‘But I don’t want to be a researcher ... I have no need to know about statistical analyses’ we hear you say. This is a comment we often hear and in many ways it is a valid point. Many students who are training for careers in the health sciences will not be conducting their own research. However, they will need to have an understanding of how research is conducted and how to evaluate that work in order to make appropriate decisions about various forms of treatment. We would hope that the most appropriate form of treatment for any ailment would be chosen on the basis of research evidence. This in essence is what is meant by *evidence-based practice* and you will come across this term quite a bit in this book. An additional benefit of reading and working through this book is that you will then be able to evaluate the many claims about health that are thrown at us by the media, and you will be making people like Ben Goldacre<sup>1</sup> very happy indeed.

So why would you want to learn about statistics? Well, there are many reasons that we can think of:

- 1) It is a very interesting subject ☺
- 2) You will learn crucial skills underpinning evidence-based practice
- 3) You will be able to understand much of the jargon printed in published research
- 4) You will be able to evaluate the quality of published (and unpublished – where you find it) research
- 5) You will be able to design and conduct your own research
- 6) You will be able to draw valid conclusions from any research that you care to conduct
- 7) You will be able to impress your friends at parties

We think that statistics is a very interesting topic if only because it leads you to appreciate that many of the phenomena that we may observe in our lives are simply down to chance factors. This is the theme of an interesting book by Mlodinow (2008), and it is a book that is worth reading as it shows the pervasive influence of chance in all our lives. Because of this pervasive influence of chance we need to be able to somehow measure it, so that we can discount it as a reason for our research findings. For example, suppose that you attended Pilates classes and you noticed that not a single member of your group had a cold or flu during one particular winter. You might reasonably conclude that Pilates has some kind of protective effect against common viruses. How, though, do you know that all the people in your class that year were not simply lucky enough to avoid these common viruses for the whole winter? How do you know how common such an occurrence is in everyday life? For example, perhaps in the pub next door one of the regular drinkers had also noticed that none of her drinking companions had contracted a cold or flu that winter. We thus need to take account of all sorts of factors when drawing conclusions from our observations. This is

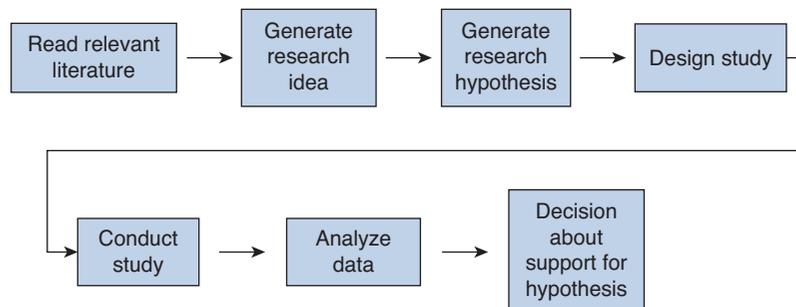
exactly the same for research as it is for anecdotal evidence such as that described in the above examples. One of the key issues that you will discover from working through this book is the importance of taking account of chance findings and assessing the probability that observed research results occurred due to chance.

## THE RESEARCH PROCESS

What is research? Well, let us start answering this question by posing another. Why do we want to conduct research? The reason we as researchers want to conduct research is because we wish to answer interesting (well interesting to us anyway) questions about the world. For example, *Is smoking related to cancer? Does eating blueberries protect us from developing cancer? Do simple cognitive strategies increase the likelihood of people eating blueberries? Do sugar pills (placebos) make people feel better?* These are all research questions and we as researchers would design and carry out research in order to find evidence to help us answer such questions.

### Deriving Research Hypotheses

In Figure 1.1 you can see one possible conceptualisation of the research process. Usually a researcher has some experience of the research in a particular field, whether it be research into the effectiveness of an intervention for the common cold or the possible causes of in-growing toenails. It is likely that researchers will have taken the time to read the previously published research in a particular area. To do this they will have searched for papers published in peer-reviewed journals using a variety of databases such as *Pubmed* in order to identify the most relevant and important research in the area. Knowledge of previous research has a number of benefits for researchers planning to conduct their own research. First of all they can see how others have tackled similar research questions. This saves them having to re-invent the wheel every time they have an interesting question to answer. Second, when researchers publish their work they often flag up future avenues of research and this can then



**Figure 1.1** Schematic representation of the research process

guide the would-be researcher in their choice of research questions. Third, knowledge of previous research lets the researchers know whether or not they might be heading up a blind alley, or whether other researchers have already answered their research question, which may save them a lot of time and effort in running a study that is unlikely to demonstrate anything useful or interesting. It cannot be stressed enough that before conducting your own research you should make sure that you know what others have done before you. To paraphrase a great scientist (Isaac Newton): make sure you stand on the shoulders of giants, that way you can see a whole lot more before you set off on your research journey.

One of the biggest benefits of knowing the previous research in the area is that it allows you to ask the most important and relevant questions yourself. For example, suppose we want to try to encourage people to give up smoking. Knowledge of the factors that best predict quitting would be essential for us to design an effective health intervention. We might look at the effectiveness of patches, of psychological interventions such as hypnosis, or health promotion activities such as advertising on TV. We would want to look at the published evidence of the effectiveness of all these previously used interventions before designing our own intervention. In addition to this we would want to draw on any research expertise that we had in order for us to accurately measure the effectiveness of our own intervention. Only through a well-designed study could we tell whether our intervention has led to an increase in people quitting smoking.

### Activity 1.1

Take a good look at Figure 1.1 and see if you can think of any problems with the way it suggests that research is conducted (the answer can be found at the end of the book).

## The Research Question

Once you are familiar with a topic that interests you, and you have a thorough grounding in what has been researched before and what theories have been proposed to account for the findings of that published research, you are ready to start asking research questions.

Research questions can be framed in a number of ways. A useful way of categorising these is that some questions you ask will focus on differences between groups of individuals and other questions will focus on how concepts might be related to each other. For example, you might ask the question: Is treatment X useful for the treatment of tinnitus? In this sort of research question we would be looking to see if participants given treatment X have fewer symptoms of tinnitus. Alternatively, we could ask whether the degree of stress that a person experiences is related to the severity of symptoms of psoriasis. In this sort of question we are interested in whether the symptoms of psoriasis seem to be related to the amount of stress that a person is experiencing. We will come on to these different ways of framing research questions later in this chapter when we discuss typical research designs in more detail.

You need to understand that the way we frame our research question has a dramatic effect on our research design, and indeed the type of statistical analysis that we can conduct on the research data that we end up collecting. For example, if we were interested in the relationship between stress and symptoms of psoriasis we need to measure participants' stress (say in the form of a questionnaire) and also get an independent assessment of the severity of their symptoms of psoriasis. We can then run some statistical analyses to see how closely these two factors might be related to one another. Similarly, if we were interested in whether a new treatment for tinnitus was effective we would want to give the treatment to one group of participants and compare this group with another group who don't get the treatment. We would then run statistical analyses which can tell us if there are any differences in the severity of tinnitus across these two groups of participants. You will find in this book that we have different statistical techniques for these sorts of research situations.

Students often ask us what statistical test is best for a particular research topic, and our initial answer will invariably be to ask the student to think about their research question. But more specifically we would ask the students to think about something called a research hypothesis. We discuss this a little further in the following section.

## The Research Hypothesis

Once you have suitable research questions you can then start to formulate testable hypotheses. There is a subtle difference here between research question and research hypothesis. A research question may be a little vague in nature, for example: Is there a link between personality and ability to quit smoking? A research hypothesis should really be much more precise. Thus we would need to identify which aspect of personality we think might be related to ability to quit smoking, e.g. whether there is a relationship between neuroticism and ability to quit smoking or whether participants who are high in extroversion and low in neuroticism will find it easier to quit smoking than those high in neuroticism and low in extroversion. It is extremely important to be as precise as possible with research hypotheses in quantitative research, as the sort of research hypothesis that you have will determine the research designs you use and the statistical techniques that will be appropriate to analyse your data. You should always remember that in order to analyse your data you have to test a hypothesis. We will cover hypothesis testing in more detail in Chapter 4, but when you precisely set out your hypotheses it helps you make decisions about how to design your research and to choose appropriate statistical techniques to test those hypotheses.

## CONCEPTS AND VARIABLES

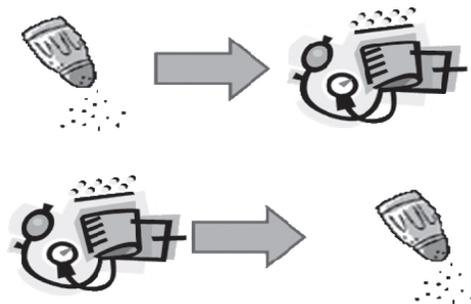
When we try to understand the world around us we often conceptualise the phenomena of interest. For example, we might have a *concept* of 'health' or perhaps 'illness' or 'treatment'. These are all concepts that a health scientist/practitioner might be interested in. Concepts can be thought of as the focus of our research. We might want to know how one concept

relates to another concept, e.g. how a particular treatment relates to a particular illness. Concepts can be abstract, e.g. the concept of ‘health’, or they can be more concrete, e.g. ‘heart-rate’. When we conduct research we have to *operationalise* these concepts into something that we can observe and measure. These measured concepts are called *variables*. Thus variables can be thought of as concepts that have been measured in some way. They are called variables simply because they vary, i.e. take on different values, from one person or situation or time to another. We might operationalise the concept of ‘health’ by asking people to give a rating for how healthy they feel on a scale of 1 to 7, or we might operationalise the concept of ‘heart-rate’ by using a heart-rate monitor. In the remainder of this book we will be focussing on variables, but it is important to appreciate this relationship between concepts and variables.

When we conduct research we are interested in variables. We are interested in variables because we want to try to find out how they might vary and why they might vary. Thus, we are not really interested in blood pressure for the sake of it, but rather we want to understand what causes blood pressure to be too high or too low and perhaps how we can prevent this from happening. In order to try to identify why variables vary as they do, we often need to look at other variables to see how they might vary in relation to our target variable. For example, we might look at salt intake to see how that is related to blood pressure. If we find that high salt intake tends to be associated with high blood pressure then we might suggest that lowering salt intake might lower blood pressure.

You should be able to see from the foregoing discussion that in science we are interested in variables and more specifically we are interested in the relationships between variables. More often than not we are trying to identify causal relationships between variables. We have to be very careful when thinking about which variable causes changes in another variable. For example, if we simply measured salt intake and found that it seems to be related to blood pressure we could not conclude that high salt intake causes high blood pressure. It might be that high blood pressure causes high salt intake. It might be, for example, that when we have high blood pressure we crave salty foods more and this leads to an increase in salt intake. We could look to see the direction of the causal link between these two variables by setting up an *experiment* (we take a closer look at experiments later in the chapter). We could deliberately alter the amount of salt that people have in their diets and see if this leads to a change in blood pressure. Alternatively, we could try to manipulate a person’s blood pressure to see if it leads to an increase in salt intake. In this way we can establish the causal relationship between these two variables.

Sometimes it is not possible to look at variables by conducting experiments because we are not able to manipulate the variables that we are interested in. For example, we cannot manipulate a person’s age, they are the age they are and we cannot change that. Also, it is often unethical to manipulate variables, for example we would not want to burn someone to see what effect this has on their



heart-rate. Often our ethical guidelines mean that we just would not want to manipulate the variable that we are interested in. For example, we would not want to manipulate a person's blood pressure because of the potential harm it might cause them. In such situations we would simply measure naturally occurring levels of such variables and see how they might be related to other variables of interest. In such studies we are simply observing and measuring variables and then establishing how they might be related to each other. These are often called *correlational* studies.

When you focus on variables, you will begin to see that not all variables have the same characteristics. For example, the sex of a person is a variable (that is, it varies from one person to the next). This is classed as a *categorical variable* as the values that it can take are simply categories, in this case the categories are *male* and *female*. Other examples of categorical variables are illness diagnosis, thus a person could have muscular dystrophy or not have muscular dystrophy. Or you might classify participants in your study as having generalised anxiety disorder, social phobia or panic disorder. In these cases the diagnosis is the variable. The category that people are placed in varies as function of the symptoms they may or may not have. Another type of variable might be the actual number of symptoms a person has. For example, if we look at the criteria for a diagnosis of chronic fatigue syndrome (CFS) there are quite a large number of symptoms<sup>2</sup> including:

- Severe fatigue
- Muscular pain, joint pain and severe headaches
- Poor short-term memory and concentration
- Difficulty organising your thoughts and finding the right words
- Painful lymph nodes (small glands of the immune system)
- Stomach pain and other problems similar to irritable bowel syndrome, such as bloating, constipation, diarrhoea and nausea
- Sore throat
- Sleeping problems, such as insomnia and disturbed sleep
- Sensitivity or intolerance to light, loud noise, alcohol and certain foods
- Psychological difficulties, such as depression, irritability and panic attacks
- Less common symptoms, such as dizziness, excess sweating, balance problems and difficulty controlling body temperature

People who have CFS will vary in the number of these symptoms that they have at any particular time, and thus the number of symptoms of CFS could be a variable of interest in our research. It is clear that this variable is different from the categorical variables described earlier as we are counting the number of symptoms rather than classifying things or people. We might suggest that such a variable is called *discrete* as we are counting whole numbers of symptoms here and thus the values that the variable can take are only in terms of discrete whole numbers of symptoms.

A final form of variable is a variable that we might call *continuous*. Such a variable can take on any value on the scale that we are measuring. A good example might be reaction times.

Let us suppose that we are wishing to test the effects of a new hay fever treatment. We are concerned that it might have the effect of slowing a person's responses. We would want to obtain a measure of patients' responses when they have taken the medication and also when they are medication free. In such a study we might ask the participants to press a response button as fast as they can when they see a certain target picture appear on the computer screen. We would then record how long it took the participants to respond in this task before and after taking the new medication. We would get the computer to measure the time between the target being presented and the instant that the participant responds. Usually in such tasks we measure reaction times in milliseconds. But response times might be measured even more precisely than this if we had suitable equipment. Response times in this study would be classed as a continuous variable.

It is important to note here that there is a difference between the underlying concept and the way we measure it. It might be that the underlying concept can be considered to vary on a continuous scale (e.g. time) but we choose to measure it on an interval scale (e.g. in days or seconds) or on a categorical scale (BC and AD). Just because we have measured a variable in a particular way does not mean that the concept varies on the same scale of measurement.

## LEVELS OF MEASUREMENT

The sorts of statistical tests that we conduct on our research data depend very much on the sort of variables we are measuring. Usually, in order to determine which tests might be most appropriate we look at the *level of measurement* that we have.<sup>3</sup> Level of measurement relates to how we have measured the variables that we are interested in. For example, if we are interested in response times we might classify participants as being 'like lightning' if their response to a question was faster than one second or as being 'slow coaches' if their response was slower than one second. Alternatively, we could ask a judge in a study to give a rating on a five-point scale of how fast they thought participants had responded (where 1 indicates extremely slow and 5 indicates super quick). Alternatively, we could simply use a stopwatch to measure response times. The point that we are making here is that when we conduct research we will need to make decisions about how we measure the concepts that we are interested in (remember this is called operationalisation). These decisions that we make can have a big impact on the types of statistical tools that we are able to use to analyse our data and this is largely because we have different tools for different levels of measurement.

The lowest level of measurement is called a *nominal* scale. Such measurements are typically frequency counts of participants in a category. For example, if we were interested in sex differences in the diagnosis of autism we would count up the number of males and females with the diagnoses and compare them, perhaps using a test like chi-square (see Chapter 9). The crucial characteristic of nominal level data is not only are they classed as categories but there is also no order to the categories, you couldn't say that one category is higher or lower than another (such variables are thus categorical). Thus, we wouldn't be able

to say that being female is better or worse, higher or lower, than being male. We are simply counting how many cases there are in each category. Another good example of a nominal level variable is religion. Here we can't say that being a Christian or Muslim is higher or lower on the scale than say being a Jew or a Hindu. They are simply different categories.

The second level of measurement is the *ordinal scale*. Here we have some sort of order to the different categories on our scale. A good example is the rating scales that are often used to get participants' opinions about things. So we might be interested in how good patients rate their accident and emergency (A & E) department to be on a five-point scale where 1 equals *an absolute shambles* and 5 equals *absolutely fabulous*. Take a look at the rating scale below. How good do you think your A & E department is?

An absolute shambles	Not very good	Neither good or bad	Quite good	Absolutely fabulous
1	2	3	4	5

Using this scale we can see that someone rating the A & E department as a 5 considers this better than someone giving a rating of 3 or 4. Also someone giving a rating of 1 thinks the department is worse than someone giving a rating of 2 or 3. Thus there is some order of magnitude to the data from the lowest rating to highest. The important point to note about such a scale though is that we do not have equal intervals between adjacent points on the scale. Thus we couldn't necessarily say with confidence that the difference between a 1 and 2 on the scale is the same as the difference between 3 and 4 on the scale. That is, the difference between *An absolute shambles* and *Not very good* is not necessarily the same as that between *Neither good or bad* and *Quite good*. Thus, although there is an ordering of the categories on the scale we do not have equal intervals between the adjacent scores. This means that many of the statistical tests that are discussed in this book are not appropriate for such data. We would usually, when dealing with data from ordinal scales, use what we call *non-parametric* tests (see Chapter 4, for example), although as we suggested earlier there is still some debate concerning this.

The next level of measurements is those that involve *interval* scales. In these sorts of measurement scales the difference between adjacent points on the scale are equal. That is, there are equal intervals along the scale. Perhaps the best example of an interval scale is one of the scales that we use to measure temperature, e.g. the Fahrenheit scale. Measuring temperature on such a scale we know that the difference between 0° and 1° Fahrenheit is exactly the same as the difference between 11° and 12°, which is the same as that between 99° and 100°. Once we start using scales that are interval level we have much greater choice in terms of the statistical tools available to us for data analysis. Provided we meet certain assumptions we are able to use both parametric and non-parametric statistical tests (see Chapter 4). One of the problems with interval level scales like temperature is that they do not have a fixed zero. Realistically speaking we don't have an absolute point where we have zero temperature (e.g. zero Fahrenheit and Centigrade do not equate to zero temperature), the zero points on

the scales that we use to measure temperature today in many ways are arbitrary zero points (Centigrade reflecting the freezing point for water and Fahrenheit being even more arbitrary). Why is not having a fixed zero important? The answer to this is that if we do not have a fixed zero then we cannot calculate ratios using the measurement scale. Thus we would not be able to say that  $10^\circ$  is twice as hot as  $5^\circ$ , or that  $50^\circ$  is half as warm as  $100^\circ$ . When you have a fixed zero on the scale you can calculate such ratios. An example of a scale that has a fixed zero is number of symptoms of an illness. When someone has a score of zero on this scale they have absolutely no symptoms. With such a scale we can say that someone who has eight symptoms has twice as many symptoms as someone who has four symptoms and four times as many symptoms as someone who has only two symptoms. Thus, when we have such scales they allow us to calculate such ratios. Not surprisingly such scales are called *ratio* level measurements.

We can view the different levels of measurement in order of level as follows:

- Nominal
  - Ordinal
  - Interval
  - Ratio
- 
 Increasing levels of measurement

### Activity 1.2

Have a go at categorising the following variables in terms of their level of measurement:

- Types of jobs undertaken by staff in an intensive care ward
- Ratings for job satisfaction of A & E staff
- Number of visits to a family doctor after a hospital stay for heart transplant patients
- Length of time to regain consciousness after a general anaesthetic
- Number of fillings given to primary school children
- Temperatures of children after being given 5 ml of ibuprofen
- Ethnicity of patients

## HYPOTHESIS TESTING

Once we have set out our research hypothesis we can then proceed to design research which tests this hypothesis. When we have a clear hypothesis this will have a big influence on how we design our study and which statistical tests we should use to analyse our data. Let us look at a general research question, say: Is high salt intake linked to high blood pressure? We might frame our hypothesis in two ways. We might say that people who have a high salt intake will have higher blood pressure than those who have a low salt intake. Here we are

interested in differences between groups of people – those who have high salt intake compared to those who have low salt intake. If we set the study up in this way we would use a statistical technique which tests for differences between groups of people (e.g. the *t*-test or the Mann–Whitney test, see Chapter 7). We might frame the research hypothesis in a slightly different way. We might simply state that we think there will be a relationship between salt intake and blood pressure such that increasing salt intake is associated with increases in blood pressure. In this sort of study we would use statistical techniques which measure relationships among variables (e.g. Pearson’s Product-Moment Correlation coefficient, see Chapter 10). In these two examples we are designing studies and running statistical analyses to test hypotheses. The statistical analyses help us decide whether or not we have support for our hypotheses.

## EVIDENCE-BASED PRACTICE

What is evidence-based practice? Well, we guess we have to go right back to ask the question: What is the purpose of scientific research? One answer to this question is that we use scientific research to understand our world better. If we understand our world then we can behave more appropriately in response to our new understanding. For example, if we found out that improving sanitation leads to much lower levels of infections then we would want to ensure that we had the highest levels of sanitation possible. If we find that using the MMR vaccine leads to increased incidence of autism we would want look for other ways of inoculating against measles, mumps and rubella. The changes that we make in response to research evidence constitute evidence-based practice (EBP). Thus, given that there appears to be no link between the MMR vaccine and autism we should look to using this route to inoculation rather than separate inoculations for these three diseases, as the latter is associated with much higher risks of infection and long-term harm to children. These are both examples of EBP. Essentially, the ethos of EBP is that we look at the available research evidence and we base our plans, behaviours and practice on such evidence.

There is one important requirement for engaging with EBP, and that is that you need to understand what constitutes evidence. Generally, evidence comes out of conducting scientific research and testing research hypotheses. Thus improving your knowledge of research will enhance your ability to engage with EBP.

## RESEARCH DESIGNS

In this section we wish to introduce you to some of the main ways in which researchers design and conduct their research. We will cover experimental and correlational designs as well as single-case designs. In our research we might be interested in differences between conditions, e.g. the difference in blood pressure between a no-salt and low-salt group of high

blood pressure patients. Alternatively, we might want to focus on relationships between variables, e.g. the relationship between anxiety/stress and waiting times in an A & E centre. Let us first look at differences between groups.

## Looking for Differences

Quite often in health research we are interested in differences between the means of different groups. For example, we might be interested in the difference between a new treatment group and a standard treatment group in recovery from septicaemia. We might compare the length of time it takes participants to recover from the illness in these two groups. Another example of looking at differences would be to compare the same group of patients under two separate conditions. For example, we might want to see if we can improve brain injured patients' ability to navigate a hospital by training them using a virtual reality (VR) tool. In this sort of study we might assess navigation ability before and after training with the new VR tool. If we had undertaken the type of research in the first example above we would have used what we call a *between-groups* or *between-participants* design. If we had conducted that research in the second example then we would have used a *within-groups* or *within-participants* design.

## Between-Groups Designs

The key feature of a between-groups design is that you have different participants in each condition that you are comparing. By 'condition' we mean the conditions under which people participate in the research. In a between-groups design these conditions will be different for each group of participants in the study. The beauty of this sort of study from the perspective of statistical analyses is that each group is usually independent, that is a person in one group cannot influence the results of a person in another group. The observations on the variables that we are interested in are completely independent from each other. Most of the statistical tests we use assume that the scores from participants are independent of each other. When we have two separate groups of participants like this it is sometimes called an *independent groups* design to emphasise the fact that the data from each group are independent of each other.

## Randomised Control Trials

A classic example of a between-groups design is the *randomised control trial*. Let us look at the first example given above, the difference in septicaemia recovery times between a new treatment and a placebo treatment group. In such a study we would randomly allocate participants to each of these two groups. We would then give patients treatment (either the new treatment or the placebo treatment) and compare the groups in terms of recovery times. There are some important features of this sort of design that make it the gold standard for research in the health sciences (we discuss these in much more detail in Chapter 14). First of all we need to randomly allocate patients to the various conditions in such a design.

Using such a process, when we came across a patient who was willing to take part in the study we would use random number tables or toss a coin to decide whether they got the new or the standard treatment. Such research designs are often also called *experimental designs*. In an experimental design the researcher manipulates one variable called the *independent variable* to see if there is an effect on another variable called the *dependent variable*. In the example given earlier in this paragraph the treatment group is the variable being manipulated: we have decided to give one group of patients a new treatment and another group of patients a placebo treatment. We have manipulated what sort of treatment each group of participants receives and this is therefore the independent variable. Also in this example we want to see if there is a difference between the new treatment and placebo treatments in the recovery times from septicaemia. Recovery time is thus the dependent variable in this study. Students learning about experimental designs often have difficulty working out which variables in a study are the independent and dependent variables, and so it is worth making the effort now to understand these.

In an experimental design it is important that the participants are randomly allocated to the various conditions of the independent variable. The reason for such random allocation of participants is that it reduces the risk of there being systematic differences between your two treatment groups, which may end up compromising the conclusions that you can draw from your study. For example, suppose we allocated the first patients to volunteer for the study to the new treatment condition and then all other participants to the standard treatment condition. It could be that the first volunteers were the more-urgent cases and thus we would expect longer recovery times than for the less-urgent cases. If we used such a process of non-random method of participant allocation and we found a difference between our conditions in terms of recovery times we would not know if it was a result of the treatment or the severity of the illness. We would have introduced a *confounding variable* into the study.

Confounding variables are variables that are not central to your study but which may be responsible for the effect that you are interested in. Whenever you let confounding variables into your study designs then you have less ability to draw firm conclusions about the differences between your treatment conditions. Random allocation helps us guard against potential confounding variables, and if we do not randomly allocate participants to conditions we have to be acutely aware of the increased problem of potential confounds.

You might ask why wouldn't every researcher use random allocation of participants to conditions when it is such a good safeguard against such confounds. Well, it is quite often the case that we want to compare groups of people who cannot be randomly allocated to conditions. For example, we might want to find out whether there is a difference in the number of back injuries suffered by male and female nurses. Here we are interested in the difference between male and female nurses. We cannot randomly allocate participants to our target groups as they are already either male or female. We thus have to be aware that there are more potential confounds with this sort of study. When you investigate differences between intact groups such as males and females, or those diagnosed with a disease compared to those without a diagnoses you are said to be undertaking *quasi-experimental* research.

This is not quite an experimental design as you have not been able to randomly allocate your participants to the conditions that you are interested in.

**Activity 1.3**

Try to identify the independent and dependent variables in the following example studies:

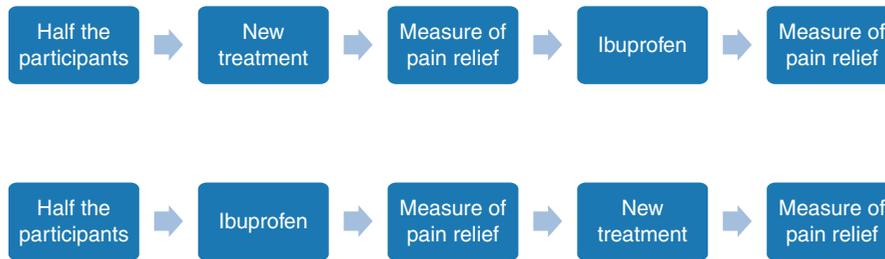
1. Examining the difference between paracetamol and aspirin in the relief of pain experienced by migraine sufferers
2. Examining the effects of consultants providing full information about a surgical procedure to patients (rather than minimal information) prior to surgery on time to be discharged post surgery
3. Examining the difference between wards with matrons and those without on in-patient satisfaction
4. Examining the uptake of chlamydia screening from family doctor surgeries with and without chlamydia health promotion leaflets

## Within-Groups Designs

Sometimes in research we might not be necessarily interested in comparing across different groups but rather in comparing one group of people across a number of different tasks or comparing the same group of people on a number of different occasions. For example, we might want to know whether patients with Alzheimer's disease have bigger decrements in memory when in new situations as compared with at home, or we might be interested in comparing the short-term memory capacity of Alzheimer's patients from one year to the next. Such designs are called *within-groups* or *within-participants* designs. One of the problems with between-groups designs is that you have different groups of people in each of your conditions. This means that by chance your groups might be different on some important variable that undermines your ability to draw causal inferences about how your variables are related to each other. Remember, we suggested that random allocation of participants to conditions is the best way of limited the impact of this sort of problem. Another way of limiting this is through the use of within-groups designs. In such designs you have the same group of people being measured on multiple occasions or under multiple conditions. This means that you don't get the differences across groups as a result of individual differences, as each participant is effectively compared against her- or himself, they act as their own control. Another positive feature of using within-groups designs is that because you only need one group of participants you need to recruit fewer of them to take part in your study. Imagine you had a study where you wanted to see if a new treatment for migraine was more effective than ibuprofen. You could recruit 40 migraine sufferers and randomly allocate them to either the new treatment or the ibuprofen group and then compare them to see which participants experienced the most pain relief. Alternatively, you could recruit one group of 20 participants, and the first time they had a migraine they would take one of the treatments and the second time they would take the other treatment, and on each occasion you would



**Figure 1.2** Order of events for the within-groups pain relief study



**Figure 1.3** An illustration of the order of conditions in a counterbalanced study

record how much pain relief they experienced. You can see here that you only need half the number of people for the within-groups design as you would for the equivalent between-groups design.

One of the problems with this arrangement of conditions is that all the participants have received the treatments in the same order. Because of this we do not know whether there might be some bias in the way they report their experiences of pain relief. Or perhaps some participants drop out of the study and so do not complete the second stage of the study. If this is the case then the lost participants would all be from the ibuprofen condition and this would lead to a less-sensitive study. One way around such problems in within-groups design is to use *counterbalancing*. In a counterbalanced study half the participants would receive the treatments in the order indicated in Figure 1.2, and the other half of participants would receive the ibuprofen followed by the new treatment as in Figure 1.3.

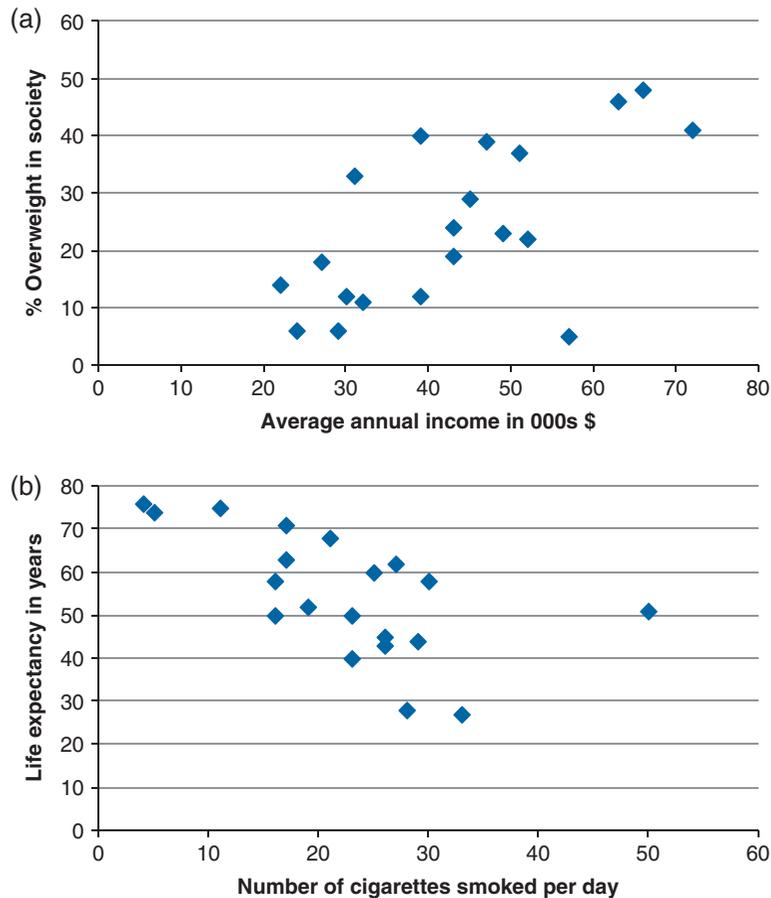
In the counterbalanced study, if there is bias in the study it is spread equally across both the new treatment and the ibuprofen conditions. If participants drop out from the study it is quite likely that they will do so from both of the pain relief conditions.

## Correlational Designs

Quite often in research we are not interested in looking at differences among groups but rather in how one variable might change as another variable changes. For example, it seems to be the case that as we increase the wealth in a society the weight of the citizens becomes greater. Thus, in most western societies we are currently experiencing a dramatic increase in rates of people who are overweight. Another example could be that as the number of cigarettes smoked by a person increases their life expectancy decreases. What we are dealing with here are *relationships* among variables. We want to know how one variable varies in relation

to another variable. In such research designs we simply take measures of the variables that we are interested in and then look to see how they vary in relation to each other. Such designs are called *correlational designs*. We can use statistical techniques such as Pearson's Product-Moment Correlation coefficient or Spearman's rho to give us a measure of how strongly any two variables are related to (or correlated with) each other. We cover these sorts of designs and analyses in Chapter 10. A useful way of representing the relationship between two variables is to plot a scattergram (see Chapter 10 for more on these) we have presented two examples of these in Figure 1.4 (these graphs were generated using hypothetical data).

In Figure 1.4(a) you should be able to see that as annual income increases (as you move to the right along the  $x$ -axis) there is a trend for the percentage of people classified as overweight to increase too. The dots on the graphs cluster around an imaginary line running



**Figure 1.4** Scattergrams showing: (a) the relationship between annual income and percentage of people overweight, and (b) between number of cigarettes smoked per day and life expectancy

from the bottom left-hand corner to the top right-hand corner of the graph. We call such a pattern of findings a positive relationship. In a positive relationship between two variables as one variable increases so does the other. In Figure 1.4(b) however, you should be able to see that as the number of cigarettes smoked per day increases, life expectancy seems to decrease. The dots on the graph seem to cluster around an imaginary line that runs from the top left-hand corner to the bottom right-hand corner of the graph. We call such a pattern of results a negative relationship between the variables. In such relationships as one variable increases the other decreases.

### Activity 1.4

See if you can describe a study which uses an experimental design looking at the link between childhood exercise levels and symptoms of Attention Deficit Hyperactivity Disorder (ADHD) (if you are unsure of this please have another look at the section on 'Randomised Control Trials' above). When you have done that see if you can design a study which uses a quasi-experimental design (please bear in mind what the difference is between experimental and quasi-experimental designs). Finally, do the same again but use a correlational design (remember when we are doing correlational research we simply measure the variables of interest and then see how they are related to each other).

## Causation

Quite often in research we want to know what causes a variable of interest to change. For example, we might want to find out what has caused an increase in asthma over the past decade or whether an increase in the dosage of a drug causes a decrease in the symptoms of a particular disease. If we are interested in such causal relationships we often run experimental studies. In an experimental study we manipulate one variable called the *independent variable* (IV) and see what effect this manipulation has on another variable called the *dependent variable* (DV). In such a study we can see what causal effect a change in the IV has on the DV. Thus we might manipulate the dosage of a drug to see what effect it has on the symptoms of a disease. When we deviate from such experimental designs we are less able to draw causal conclusions. For example, suppose we were interested in the difference between people who fracture their arms and those who fracture their legs in the uptake of physiotherapy. If we find that the people who break their legs are less likely to turn up for physiotherapy we couldn't say that the type of fracture has caused the difference in the uptake of physiotherapy. It could be for example, that certain types of people (e.g. males) are more likely to have leg fractures and it is their sex that is responsible for them not taking up physiotherapy opportunities. When we have quasi-experimental designs like this, more possibilities for confounding variables creep in.

If we look at correlational designs it is also very difficult to determine the causal direction of the relationship between variables. Suppose we find that there is a positive correlation between alcohol consumption and blood pressure, which of these variables has caused a change in the other? It could be that consuming higher levels of alcohol causes an increase

in blood pressure, or it is perhaps equally plausible that those with high blood pressure self-medicate by drinking more. The causal direction of the relationship between these two variables is unclear. We return to this issue again in Chapters 5 and 10.

## Summary

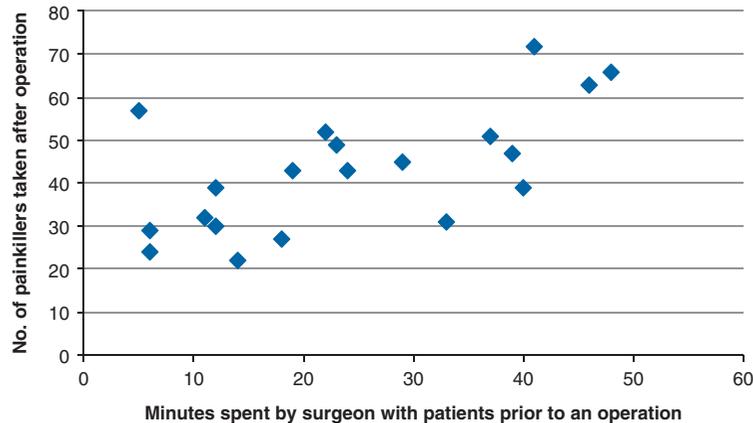
We have now introduced you to many of the basic concepts of research design. Armed with knowledge of these concepts, the statistical techniques that we cover in the rest of the book should make a little more sense. In addition, the research that you read about in journals may make more sense. Furthermore, you will be able to scrutinise the claims people make about the causal relationships among the variables that they have investigated.

# MULTIPLE CHOICE QUESTIONS

1. What is the difference between a research question and a research hypothesis?
  - a) Usually research questions are more precise than hypotheses
  - b) Usually research questions are more vague than research hypotheses
  - c) Usually research questions are exactly the same as research hypotheses
  - d) None of the above
2. What are the benefits of knowing about the previous research in a particular field of interest?
  - a) We can see how others have tackled similar research problems
  - b) We can see what other researchers have suggested needs following up in future research
  - c) It saves us undertaking research that may be pointless
  - d) All of the above
3. According to the description of the research process in this chapter how do we decide whether there is support for a particular research hypothesis?
  - a) We rely on previous research to test a new research hypothesis
  - b) We need to interview other researchers to see if they agree with our hypothesis
  - c) We design a study and then collect and analyse our data in such a way as to test our hypothesis
  - d) We see if the research hypothesis makes logical sense
4. Which of the following constitute the main focus of interest for us when conducting quantitative research?
  - a) Participants' demographic details
  - b) The questionnaires that we use
  - c) Publishing our research
  - d) Variables

5. Which of the following could be considered as a ratio level variable?
  - a) Occupation of participants
  - b) Time taken to complete a programme of physiotherapy
  - c) Ratings on a five-point scale for satisfaction with out-patient services
  - d) None of the above
6. Why is it that temperature scales cannot be classed as ratio level variables?
  - a) They are too complicated
  - b) They contain arbitrary intervals between adjacent values on the scales
  - c) There are too many scales for consistent measurements
  - d) They do not have a fixed/absolute zero
7. In the scheme outlined in this chapter which of the following represents the correct ordering of the levels of measurements?
  - a) Nominal, ordinal, interval, ratio
  - b) Ordinal, ratio, interval, nominal
  - c) Ratio, ordinal, interval, nominal
  - d) Interval, nominal, ordinal, ratio
8. What is the defining characteristic of interval level data?
  - a) You can put the categories you have in order of magnitude
  - b) You have a fixed zero
  - c) You have categories which cannot be ordered in a meaningful way
  - d) You have equal intervals between adjacent points on the scale
9. Correlational designs tell us about:
  - a) Differences between conditions
  - b) The causal relationships between variables
  - c) The relationships between variables
  - d) None of the above
10. What is a quasi-experiment?
  - a) A study where you simply measure the relationship between two variables
  - b) A study where you are interested in the difference between intact groups
  - c) A study where you randomly allocate participants to your experimental conditions
  - d) A study looking at the reflexes of the hunchback of Notre Dame
11. In an RCT study how should you allocate your participants to experimental conditions?
  - a) Randomly
  - b) By matching participants in each condition on the bases of demographic variables such as age
  - c) Put all those to volunteer first in one condition and then the remainder in the other condition
  - d) All of the above are appropriate ways of allocating participants to conditions
12. Why are confounding variables such a problem in research?
  - a) They are difficult for participants to give responses to
  - b) They make questionnaires too long for participants to complete
  - c) They lead to high attrition rates for studies
  - d) They make it difficult to draw conclusions about the relationships between the main variables in the study

13. In which of the following designs are you least likely to have a problem with confounding variables?
- Experimental designs
  - Quasi-experimental designs
  - Correlational designs
  - Both a) and c) above
14. Which of the following designs are best for establishing causal links between variables?
- Experimental designs
  - Quasi-experimental designs
  - Correlational designs
  - Both a) and c) above
15. Take a look at the following scattergram. What can you conclude about the relationship between the two variables?



- That there is a negative relationship between minutes spent with patients and painkillers taken
- That there is no relationship between the two variables
- That there is a positive relationship between minutes spent with patients and painkillers taken
- No conclusions can be drawn from this scattergram

## Notes

- Ben Goldacre writes a column for *The Guardian* called 'Bad Science' in which he critically appraises many of the claims made about health in the media ... it really is a good read, as is his book of the same name.
- Taken from the NHS Chronic Fatigue web pages (<http://www.nhs.uk/Conditions/Chronic-fatigue-syndrome/Pages/Symptoms.aspx>).
- There is though some debate about the importance of levels of measurement in the choice of statistical tests. However, we feel that it is instructive to discuss these here as they provide a useful framework for understanding the different types of variables that we deal with in health science research.