In this part we discuss the process of actually doing research.

In Chapter 5 we discuss the design or methodology of the research. This includes the concepts of variables, controlling variables, taking samples, and experimental and control groups. In Chapter 6 we discuss the process of collecting data. We start with the design of reliable and valid instruments and then focus on data collection in the laboratory, through the use of models or simulation. In Chapter 7 we discuss the problems of collecting data from humans.
5 Research Design

Empirical research includes experimental, *ex post facto* and descriptive (case study) research. In this chapter we look at variables and how to control them, samples and how to take them, and other important design issues in empirical research.

Empirical research uses **induction**. Induction is the formulation of general theories from specific observations, as opposed to **deduction** which is the derivation of a new logical truth from existing facts. As an example of induction, if you observed 500 tomatoes and found in each case that the tomato was red, you might *induce* that all tomatoes are red. As an example of deduction, say you know as facts that (1) all stars contain hydrogen, and (2) the sun is a star, then you can *deduce* the new fact that (3) the sun contains hydrogen. Note that the results of deduction are always true (if the existing facts used are true) while the results of induction are not necessarily true (some tomatoes are green!).

Almost all scientific theories are based on induction. This leads to the possibility of such theories being wrong. As a result, a theory must be continually **tested**: one inconsistent observation is enough to disprove a theory, but no amount of supporting observations can prove it absolutely; they only add to one’s confidence in the theory.

Experimental research uses the experimental method. The experimental method can be thought of as **systematic trial and observation**: trial because the answer is not known before-hand, observation because the result must be carefully recorded, and systematic because all good research is planned and purposeful.

5.1 VARIABLES

In the experimental method one chooses a variable (also known as a **factor**) and manipulates it. This is known as the **independent variable**. The effect of this manipulation on other variables (the **dependent variables**) is measured. Say one wished to study the effect of temperature (independent variable) on cows’ heart rates (dependent variable). A basic approach would be to change the temperature surrounding the cows, and then measure their heart rates.

Of course, other factors can also affect heart rates—the humidity, perhaps, the cows’ diet or their general health. Such factors are called **nuisance variables**. Researchers must be sure that they are measuring the effect of the independent variable on the dependent variable, rather than being misled by the effects of nuisance variables.
Consider the following:

‘In Saudi Arabia murderers are publicly executed. In Britain the death penalty has been abolished. Saudi Arabia has proportionately fewer murders than Britain. Therefore the death penalty is an effective deterrent to murder.’

Scientific researchers will realise that the above is not sound reasoning. While we have an independent variable (treatment of murderers) and a dependent variable (number of murders committed per capita), it is unclear whether the change in the independent variable is the cause of the change in the dependent one. Saudi Arabia bans alcohol for example, and many of the murders in Britain are committed by people under the influence of alcohol—the nuisance variable (alcohol availability) thus affects the dependent variable (murders committed). The dominant religions of the two countries are different and this introduces another nuisance variable.

5.1.1 Controlling nuisance variables

A nuisance variable which is manipulated to have no effect on the dependent variable, or has an effect which can be determined and so separated from the effect of the independent variable, is called a controlled variable.

Some methods which can be employed to ‘damp out’ the effect of nuisance variables are:

1. where possible, keep the value of the nuisance variable constant (when measuring cows’ heart rates one might keep humidity at 42% throughout);
2. incorporate nuisance variables into the design of the experiment;
3. after the experiment use a statistical technique known as analysis of covariance to detect the effects of nuisance variables so that these can be separated out.

Techniques for method 2 will be discussed in later sections. A discussion of covariance techniques (method 3) is beyond the scope of this text; the interested reader should consult Net82, Hog93 or Fre90 for details.
5.1.2 Control and experimental groups

Experimental research often involves doing something new—the experiment—and comparing it with something standard—the control. The experiment and control are identical except for the two values of the independent variable (for example present/absent or high/low). This is effectively a special case of method 1 described above. The following example shows how this works:

In testing the effectiveness of a new drug for headaches, the control group receives a placebo (a ‘sugar pill’ which has no medical effect) and the experimental group receives the actual drug, but both groups are otherwise treated the same. Furthermore usually they are not told which group they are in. (This is known as a blind test; in a double-blind test not even the doctors dispensing the drugs know whether it is a placebo or not.)

A number of nuisance effects can be cancelled out in this way. The effect of people psychologically ‘feeling better’ after taking medicine and the effect of changes in outside conditions during the course of the experiment (from weather conditions to the state of the economy) should be felt equally by both groups. Any difference in the groups with regard to the dependent variable (presence and severity of headaches) can thus be attributed to the independent variable (use or non-use of the new drug).

Sampling and matching are important concepts in the design of control and experimental groups—these concepts are discussed in the following sections.

5.2 SAMPLING

Here the concepts of a ‘population’ and a ‘sample’ are introduced, and the why and how of taking scientifically useful samples are discussed.

5.2.1 Populations and samples

A population is any group that is the subject of research interest. Oxygen molecules in the universe, supercomputers in the world, guppies in South African rivers or the dogs in a particular city could all be populations—groups a researcher wants to study.

It is often not practical or possible to study an entire population—someone trying to determine the average length of adult guppies in South Africa would find it impossible to do this by measuring each and every guppy! In such cases it is necessary to make
general findings based on a study of only a subset of the population. Such subsets are called samples.

Samples must be representative of the population of concern, otherwise no general observations about the population can be made from studying the sample. A study of the incidence of mange in dogs in a city based on a sample of puppies is unlikely to produce meaningful results. Two key features of samples determine how representative of the population they are, these being size and bias.

5.2.2 Sample size

Say a gambler is playing a dice-game and believes that one die is ‘loaded’: it has been weighted so that the number 1 appears more often than others. She decides to roll the die a number of times (the sample of rolls of the die) to test whether it does in general favour the number 1 (in the infinite population of rolls of the die). If she rolls once, and gets a 1, we could certainly say that on 100% of tests the die came up 1. Most people would have problems accepting that just one roll is a valid test of the die. If she rolls 5000 times and she gets a 1 every time, then most people would be convinced that the die was in fact loaded.

As the above example shows, the sample must be ‘large enough’ to correctly represent a population. Chapter 9 gives the statistical details on how to determine what size sample is necessary to test a population correctly.

5.2.3 Sample bias

Many newspapers hold ‘phone-ins’ in which people’s views on particular topics are requested: some people phone in their views, and a selection of responses are published the next day. These responses are normally prefaced with editorial comments on the lines of ‘People in Soweto overwhelmingly rejected...’ or ‘Durbanites support...’. Many people never realise that this is complete nonsense. Think about it for a moment—who would participate in such a phone-in? People who don’t read that particular newspaper? Not likely. People who don’t have telephones? Not likely. People who don’t have strong views on the subject? Not likely. People who are naturally shy and retiring? Again, not likely. So, in fact, these newspaper surveys are not testing the population of people in their city, but are in fact testing the population of people in the city who read the specific newspaper on the day the phone-in topic was presented, who own telephones, who are confident enough to phone in their opinions, and who are interested enough in the topic to take the time to phone.
The above are just some of the more obvious problems in the sampling method employed. Similarly the question must be asked when ‘researchers’ do ‘man-in-the-street’ opinion testing – which man? And what street?

A sample is said to be biased if it represents only a specific subgroup of the population or if particular subgroups are over- or under-represented. The next section describes how to avoid bias.

5.3 SAMPLING METHODS

Random selection is the basic principle used to try to avoid bias in a sample. The random selection must ensure that each member of the population has as much chance as any other of being included in the sample. Thus taking a sample of Bloemfontein adults by randomly selecting names from the telephone book would be biased (people who don’t have telephones do not have an equal chance of being in the sample), whereas taking a sample of students at the University of Fort Hare by randomly selecting student records would not be biased.

We discuss three standard random sampling techniques.

5.3.1 Simple random sampling

Here one assigns numbers to each member of the population (enumerates the population)—for example, people’s I.D. numbers could serve as numbers for a study of South African adults, if everyone had an I.D. After performing this enumeration, one generates as many unique random numbers as the size of sample required, and the corresponding members of the population become the sample. Random numbers can be generated by a computer either by special computer programs, or by using the random functions available in programming languages such as Basic, Pascal and C.

Such enumeration is ideal; it is, of course, not always possible and other methods must be used. (The population of oxygen molecules would take quite some time to enumerate even if it stayed constant!)

5.3.2 Stratified random sampling

Sometimes one has prior information regarding certain characteristics of the population’s composition, and one wants the selection of sample items to reflect this. For
example, if we were studying housing construction in South Africa, and the population of interest was houses in the country, we might know certain rough proportions of housing types—30% informal, 50% brick, 15% cement and 5% wood, for example. A simple random sample would be unlikely to arrive at exactly these proportions (think about it). In stratified random sampling one uses simple random sampling within each group (stratum), ensuring that appropriate numbers are selected from each group so that the overall sample reflects each group in the known proportions.

To continue our example, if we wanted a sample of 1000 houses, we would select 300 informal houses randomly, then 500 brick houses, 150 cement houses and finally 50 wooden ones. Stratified random sampling is preferable to simple random sampling if members within each stratum are fairly similar (homogeneous) but there are marked differences between members of one stratum and those of another.

### 5.3.3 Cluster sampling

Simple random sampling can be impossible in large populations. It can also be unviable economically. Say we wished to determine the average level of toxins in South African streams, and have the time and resilience to actually enumerate all streams in the country. A simple random sample of 200 streams would probably have us flying and driving over great distances to individual streams to measure toxin levels, and this travel would be extremely expensive as well as time-consuming.

In cluster sampling, one subdivides the population into subgroups called clusters. One then randomly select a sample of clusters, and then randomly select members of the cluster sample to serve as the population sample.

For example, we might break our South African stream population into geographically separate clusters, with each of the nine provinces forming one cluster. We could then randomly select three (say) provinces and perform simple random sampling within each province. Cluster sampling can be done to more than one level—in each of the selected provinces (clusters), we could enumerate all magisterial districts (subclusters), and randomly select a sample of districts from which the actual sample streams is taken. And so on. Eventually, we would be able to take our sample of 200 streams by travelling to, say, only three provinces, and only four districts within each province, rather than having to travel to 200 completely separate locations.

Note that clusters are different from the strata of stratified random sampling. A population is divided into strata so that each stratum is defined by a known characteristic, whereas the division into clusters is based on spatial separation alone.
Cluster sampling is not quite as reliable as simple random sampling or stratified random sampling, but is often the only possible approach. A more detailed treatment of these issues is available in Coc77.

5.4 MATCHING

The placement of participants into the control or experiment group can be performed at random, and if the sample size is large enough this should be effective. Another common approach is called matching. Here, say we want experimental and control groups of 40 people (or dogs, or crystals) each, a total of 80 sample items. We select the 80 sample items by random sampling, but then, instead of randomly assigning each sample item to a group to make up the two groups, we instead find sample pairs which are in some way(s) similar, and randomly assign one member of each pair to the experimental group and the other to the control group. Matching can be extremely useful if the sample size must be small and there is thus a risk that nuisance variables’ effects will be pronounced in one group and not in the other.

A matching example: A pharmaceutical firm employs 200 trainee scientists in its research section. A new training course has been developed to teach certain skills, and the company wishes to determine if this course will be more effective than the existing one (effectiveness to be measured by an end-of-course test). They decide to do an experiment with a random sample of 20 scientists, with ten taking the old course (control group) and the other ten taking the new course (experimental group). However, age is known to affect how easily people absorb new knowledge, and the firm wants to eliminate this nuisance variable. They therefore divide the 20 into two groups by creating pairs according to age—the two oldest scientists as one pair, the third and fourth oldest as the next, and so on down to the two youngest. One member of each pair is now put in the control group and the other in the experimental group (at random), so that the effects of age are felt equally between the groups, and do not contribute to any differences observed between the two groups’ success.

5.5 EXPERIMENTAL DESIGN

One of the simplest experimental designs is the experimental group and control group design. If you have more than one independent variable, and you want to determine the effect of each independent variable on the dependent variable, then you will need a more complex design.

One such type of design is the factorial experiment. (Recall that ‘factor’ is an-
other name for a variable.) In a factorial design, one takes each independent variable (or controlled variable) and selects several levels for each. Then all possible combinations are tried. For example, the researcher interested on the effect of surrounding temperature on cows’ heart rates might decide that the humidity and breed of cow are two nuisance variables which she needs to control. So a simple factorial experiment might try two different breed of cows, three levels of humidity, and five levels of temperature, the result being 30 readings.

To analyse the data collected in a factorial experiment one needs advanced statistical techniques that are not covered in this book. So if you intend to use such a design you should consult a book on experimental design such as Mea88 or Hic93. Also, in many situations one cannot afford to try each possible level of each variable. So there are more complicated designs which try to extract the same information but using fewer measurements. Again a book on experimental design is your best guide.

Empirical research can also deal with changes over time. Such research can sometimes be categorised as one of cross-sectional, longitudinal or time-lag design. Say we are studying children’s attitudes towards homework. In a cross-sectional design we could compare the attitudes of 10-, 12- and 14-year-olds. In a longitudinal design we could take 10-year-olds and see how their attitudes change as they age over four years. In a time-lag design we could, over a four-year period, determine the attitudes of each year’s 10-year-olds.

These designs can be used together. Someone studying whether younger children are keener on homework by using a longitudinal design could find that this seems to be true. However changes to society could have affected all children over the period of the study—for example, the introduction of television, video games or personal computers. By using a time-lag design combined with the longitudinal design, the researcher would be aware of the effects of the nuisance variables.

Discussion questions and exercises

1. What is bias? How should sampling be conducted to avoid bias?

2. Suppose you used a greenhouse, which keeps temperatures constant, to investigate the amount of water needed to grow tomatoes using a particular fertilizer. What is:
   
   (a) the independent variable(s);
   
   (b) the dependent variable;
   
   (c) some possible nuisance variables;
(d) the controlled variable(s)?

3. What type of sampling would you use for the following situations and why?
   
   (a) the tusk size of African elephants;
   
   (b) which is more popular in South Africa: rugby or soccer?
   
   (c) the average age of your family (think carefully about this one!);
   
   (d) the average weight of Isaac Williams’ 1000 cows.

4. Can one estimate the quality of diamonds in the world by sampling only from Kimberley? What about the composition of diamonds?

5. QuikDry Manufacturers has built prototypes of two different robotic car-painters. They wish to test which is faster on average. They have a sample of 10 cars, with sizes 120, 110, 140, 230, 190, 130, 120, 150, 200, 100. Use matching to design a control group and an experimental group. How appropriate is the design?
6 Data Collection

In this chapter we first discuss the concepts of reliability and validity of data measurement devices. We then describe methods of obtaining data from laboratory experiments and models. In the following chapter we explore the problem of obtaining data from people.

6.1 INSTRUMENTS

One has to measure data somehow. Any device used for this measurement is called an instrument (e.g. thermometer to measure temperature, a voltmeter to measure potential difference, an I.Q. test to measure intelligence).

There are two fundamental criteria for instruments:

- Reliability
- Validity

The term reliability means that measurements made are consistent: if the same experiment is performed under the same conditions, the same measurements will be obtained. The term validity means that the measurements are correct: the instrument measures what it is intended to measure, and that it measures this correctly.

*Example 1:* Every month a househusband has 10 kg of fish delivered to his home. The fish is packed in ice-filled cartons to ensure freshness. He diligently weighs each delivery to ensure that he is getting the right amount of fish, and every month the scale measures 10 kg exactly. While the scale might be reliable, the measurements are not valid, as he is weighing the fish together with the packaging and the ice.

*Example 2:* A particularly controversial issue is the use of I.Q. tests to measure intelligence. One problem with an I.Q. test is that people do not consistently obtain the same scores: the I.Q. measured varies according to people’s mood, the conditions under which the test is written, people’s health and many other factors. Questions therefore arise as to the reliability of I.Q. tests.
Even if IQ tests were reliable (a person always got the same score), serious questions have been raised concerning the validity of these tests: Do I.Q. tests in fact measure intelligence? Some say that I.Q. tests simply measure how skilled people are at solving the specific types of questions asked in I.Q. tests! (See Cur90 and Jor89.)

### 6.2 LABORATORY WORK

As an undergraduate you probably did some practical work. There were (we hope) three goals:

- Learning about practical work.
- Becoming familiar with a certain instrument, apparatus or technique.
- Demonstrating theory by specific experiments.

In postgraduate study you are likely to have to deal with apparatus you have never seen before while performing experiments where nobody knows what is supposed to happen. You may be called upon to design and build apparatus yourself. None the less, all the principles of undergraduate experiments still apply:

- **Plan**: Decide what you are going to do, how you are going to do it, and how you will measure what has occurred, before you do anything.

- **Keep full records**: All data must be retained in a notebook. Detail the date, the conditions, the experimental steps and the results—don’t be vague. Diagrams can be useful.

- **Avoid errors you can avoid**: Most errors can be avoided by good technique: being careful and methodical. Look for places where errors can occur. One problem is conditions changing while a series of comparative experiments are performed. For example, in measuring the elasticity of a spring by using a series of weights, successive weights might stretch the spring and so actually alter its elasticity. Use theoretical calculations to check that various outcomes during the experiment make sense.

- **Estimate remaining inaccuracy**: Bad data must be treated circumspectly. Often data has ‘noise’ (inaccurate measurements amongst the accurate ones). If noise cannot be completely eliminated, then you must try to estimate the error in your results.
➢ Try alternatives: Try alternative measuring devices, alternative orderings. Apart from providing a check of the reliability of your results, differences amongst such alternatives could suggest new research problems.

➢ Stay safe: Always take appropriate precautions and use your common sense.

Where possible, an experiment should be repeated several times under several conditions. This is particularly true of laboratory experiments (Squ85). For results to be accepted, they must be reproducible—other researchers in the field should be able to read your report and from that perform similar experiments which yield similar results.

6.3 MODELLING AND SIMULATION

6.3.1 Models

Many problems can be formulated or translated into models. A system is a subset of the world that is considered to be self-contained. A model is a simplified representation of a system. Armed with a suitable model, one may try either or both of:

- mathematical analysis to solve or optimize;
- computer simulation to approximate.

There are several good reasons to use a model.

1. It would be too expensive to build the real thing to ‘see if it works’ (e.g. a petrochemical plant).
2. The real system exists but cannot be experimented on (e.g. a nuclear reactor).
3. One can use the model for ‘what-ifs’ (e.g. ‘what will happen to sales if the price is increased?’)
4. One can use the model for forecasting (e.g. a change in the global climate).

Sometimes the model itself is a worthwhile goal— for example, to aid our understanding of superconductivity (and explain the recent high-temperature ceramics). More often, the model is just a first step in solving a problem.

Modelling is often used as a research tool because it can be inexpensive. Paradoxically, one limitation is cost (computer time is not free). Other difficulties include
getting a sufficiently accurate model and the validation of the results (even if you ‘know’
the answer is correct you have to justify this to others).

McHaney (McH91) states ‘the best approach to model development is to incorporate
the least amount of detail while still maintaining veracity’ (emphasis added). Clogging
up a model or simulation with fine details can add a lot of time to the development
process and such details can introduce their own problems. On the other hand, essential
detail cannot be omitted if the model is to validly represent the real-world situation.

6.3.2 Simulations

The process of model creation and usage is just the scientific method in miniature.
Steps in a simulation are:

1. Define the system and the objectives;
2. Determine the model’s scope and scale (what’s in it and how much detail);
3. Choose a programming language and code the model;
4. Run the model;
5. Gather data and analyse it.

An example is Volterra’s rabbit-and-fox model of the relationship between predator
and prey in an area (McH91). The model is based on the fact that the greater the
number of foxes the more rabbits are eaten:

Say one starts with a large number of foxes and few rabbits. Then the
rabbit population decreases. After a while the scarcity of rabbits means that
there is not enough food for the foxes, foxes die, and so the fox population
decreases. This decrease in the fox population allows an increase in the
rabbit population, which in turn leads to an increase in the fox population
and a new cycle begins.

The above model:

- uses numeric calculations to predict behaviour over time: the number of foxes at
  any given time is used in the calculation of the number of rabbits in the next
  period, and vice versa;
is a simplified representation of the system as the presence of other predators and prey, as well as other ways in which foxes and rabbits could die, are not included.

Another example is a queuing system. Consider a bank with several customers queuing for the next available teller. Simulation can be used to determine the behaviour of the system (e.g. length of wait) as more and more customers are added.

The maxim ‘Garbage in, garbage out’ applies especially to simulations. If the input data is not truly representative of the real world, the simulation is useless.

6.3.3 Writing simulations

A simulation can be written in a standard computer language (such as C or PASCAL). There are also many special simulation languages (such as SLAM and GPSS/H) which have built-in commands and structures for some of the common situations one may encounter.

Software simulators are also available for specific fields. If your problem involves a situation that a simulator correctly models, the software package can obviously be enormously helpful in your research. On the other hand, if your problem has specific features which the simulator is not equipped to deal with, then it is a mistake to try to ‘change the problem to fit the simulator’.

The result of a simulation can be just a number, or even a simple yes/no to the question of whether the spaceship reaches the moon. But it is much more useful to have intermediate data displayed, preferably through the use of animation (since the eye is the best input mechanism we have). The commercial simulation languages and simulators have graphical commands built in. Beware of proof by picture, however—a neat graphic showing the spaceship landing on the moon does not by itself guarantee that there are no bugs in the simulation and that the real version won’t explode on take-off.

Discussion questions and exercises

1. What criteria would you use for assessing the worth of a research instrument? Discuss.

2. Take one example of research interest, list some of the problems of collecting data, and suggest ways to avoid them.

3. Construct a model of a student’s success in a course. (Will she pass or fail?)
4. How would you go about experimenting on:

(a) the ability of mice to find their way through a maze?

(b) the colour of metals when burning?
7 Data from People

In this chapter we discuss techniques of collecting data from people. The computer analyst wishing to study user satisfaction after introducing a new system, the psychiatrist wanting to determine stress levels of executives, the bio-engineer wanting to determine the acceptability of a meat substitute—they all need data from people.

7.1 INSTRUMENTS

Just as one would use a barometer to measure pressure or a stopwatch to measure time, so one needs some instrument to measure whatever it is one is studying in people. The most common instruments used in this regard are tests, interviews and questionnaires.

The design of a reliable instrument for measuring people’s attitudes or capabilities requires careful planning. The reliability can be checked by using one of three approaches:

- The test–retest approach: administer the same instrument at a later time and see if you get the same results. Many respondents (people answering the questions) are not happy answering the same questions twice, however, and the time lag needed between test and retest can also cause responses to change and respondents to be ‘lost’.

- The equivalent form approach: each question on the original test, interview or questionnaire is rephrased so that one winds up with two tests that ‘look different’ but effectively ask the same questions. If there is a high correlation between people’s responses to the original questions and to the rephrased questions, then one has an indication that the test is reliable (and that people aren’t just ‘answering at random’).

- The split-half approach: this is a modification of the equivalent form method where the two tests (original and equivalent form) are combined into one. The fact that question 14 and question 87 are differently worded versions of the same question will escape most respondents’ notice, and this allows the test to be given at a single sitting rather than at two.

Ease-of-use has a major impact on the reliability of instruments—as you can imagine, frustrated, bored or confused people cannot be relied on to answer a set of questions consistently—and it ensures higher participation.
An instrument is reliable if it consistently gives the same results—so a calculator consistently getting an answer of 3 when it adds 2 and 2 is reliable. This does not however make the answer valid (right) however! Testing the validity of a new instrument used for people is not easy. Three approaches are used; in order of preference these are:

- **Criterion-related validity**: this measures whether an instrument accurately predicts (predictive validity) or diagnoses (concurrent validity) some particular variable (criterion). For example, an aptitude test for trainee mechanics could be shown to be valid if the test scores correlate highly with eventual success or failure as a mechanic.

  Of course, if you have an existing instrument that measured the same thing and you know that it is valid, you could simply compare the results obtained from the new instrument with those of the old. (A new test might be desirable if, for example, the current instrument is too long.)

- **Construct validity**: If you have an existing instrument which measures something which is known to be closely related to the thing you want to measure, compare the results obtained by the new instrument with that of the old, and check that there is a high correlation. Say you have an existing way to measure writing skills, and your new instrument seeks to measure reading skills. If you know that reading skills are correlated with writing skills, then a high correlation between the results obtained by the new test on reading skills and the existing one on writing skills would suggest that the new test is valid.

- **Content validity**: If no related instruments exist, then gather expert opinion on each question on the instrument to determine whether or not each question actually tests what it is supposed to. In addition, the experts must agree that the questions as a whole constitute a valid and representative test of the variable being measured.

A more detailed treatment of reliability and validity of instruments for people is available in Huy87 and Mul87.

### 7.2 QUESTIONNAIRES

A questionnaire is a printed list of questions which respondents are asked to answer. These instruments are commonly used and commonly abused. It is easy to compile a questionnaire; it is not easy to compile an effective one. Effectiveness requires planning beforehand to ensure that the data can be objectively analysed afterwards.
Open (or unstructured) questions can be used in a preliminary survey or to get a feel for the subject. Here respondents answer in their own words to questions. Closed (or structured) questions are used in large-scale data collection. Here respondents choose from a collection of alternatives (e.g. true/false) or assign a numerical score or ranking. Closed questions often use four-point scales, for example:

Q12. Bart Simpson is a good role-model:
Strongly disagree Disagree Agree Strongly agree

A 4-point scale forces a decision, a 5-point scale allows a neutral answer.

A good questionnaire:

- is complete: gets all the data you need;
- is short: don’t abuse the respondent’s time or concentration;
- asks only relevant questions;
- gives clear instructions;
- has precise, unambiguous and understandable questions;
- has objective questions: don’t suggest answers;
- starts with general questions;
- has appropriate questions: if you have to ask sensitive questions, put these at the end;
- uses mostly closed questions, often with a 4-point scale.

You have no right to expect total honesty. Some participants will not care that much about the answer or will try to give the ‘socially correct’ answer. These problems can be reduced if you make the instrument easy to use, and explain the importance of the research to your respondents. Be courteous, thank people for their help, offer to share the conclusions, provide a stamped self-addressed envelope, etc.

A particular problem with questionnaires is that of non-returns. If you get 800 returned questionnaires from 1000 you’ve distributed (a very high return rate), what do you do about the ‘missing 200’ responses? Unfortunately you cannot just rely on 800 forming a big enough sample, as there is a problem of bias. The particular bias involved is determined by people’s reasons for not returning the questionnaire (too disorganised? too lazy? offended by the questionnaire? didn’t understand it? their area didn’t receive it?).
This problem with questionnaires is a specific case of the general problem with human volunteers. The fact that volunteers are unlikely to be representative of a population (more adventurous? poorer?) means that studies using volunteers are open to criticism. Studies on human sexuality have been particularly criticised in this regard (Lou91).

7.3 INTERVIEWS

An interview involves a one-on-one verbal interaction between the researcher and the respondent. Much of what was said above for questionnaires is true for interviews. An interview should have a plan. The researcher must not direct the person’s answers through his tone of voice or through the way he phrases the question—you agree with us that this is right, don’t you?

One area where you would need to use interviews rather than questionnaires would be in getting information from people who can’t read. Other advantages of an interview over a questionnaire are that one can clarify answers and one can follow-up on interesting answers. Some advantages of a questionnaire over an interview are that the answering of the questionnaire can be done at a time suitable to the respondent, and the respondent may not be as inhibited in answering. A questionnaire is also the only practical approach when dealing with large numbers of respondents, though.

7.4 ETHICAL CONSIDERATIONS

Apart from instrumentation and procedural concerns, collecting data from people raises ethical concerns. These include avoiding harm to people, having due regard for people’s privacy, respecting people as individuals, and not subjecting people to unnecessary research.

In order to avoid doing harm to people, one must guard against both physical damage and psychological damage. People have a right to privacy and the researcher must keep data collected confidential. This implies that the subject(s) should not be identifiable to anyone reading the eventual report. Importantly, the researcher must remember that the subjects are individual human beings, and so treat them with respect. These and related questions are discussed further in Chapter 15.

Discussion questions and exercises

1. How would you design an instrument to test:
(a) the height of adults?

(b) the intelligence of adults?

(c) the intelligence of babies?

(d) the attitude of elderly people to technology changes?

(e) the racial attitudes of urban dwellers?

2. Suppose one wanted to use an aptitude test in South Africa that had been produced in the United States of America. Suggest some changes that might be needed.