Overview

Collecting data: the variety of methods of collection

Primary and secondary sources:
- Primary data – you collect yourself
- Secondary data – you use data collected by someone else

Surveys:
- Using secondary sources
- Doing your own surveys
- Sampling:
  - Random
  - Nth number
  - Stratified
  - Convenience

Administering surveys by post and phone: the effects on the sample

The general problem of bias

Research instruments: the questionnaire:
- Setting questions
- Re-using questions from other sources
- Design: physical appearance

Tally sheets: unobtrusive data collection

This chapter focuses on quantitative methods. The larger part of the chapter deals with surveys, including using secondary sources as well as conducting your own surveys. There is a discussion about how different mechanisms for
administering surveys such as by post or phone will affect the sample and findings. Following this, we move on to designing questionnaires, including a detailed section on setting questions and some tips on making your questionnaire attractive and easy to use. Finally other quantitative methods are considered, particularly tally or score sheets, which are a quick and easy way to collect data in an unobtrusive way.

**Collecting Data: A Variety of Methods**

How you will record data is a key issue: get it right and this will be enormously useful when it comes to making sense of your data when you bring all of your records together. Getting it wrong may lead to incomplete datasets and difficulties in reconciling their apparently discrete meanings.

There are two kinds of data which you will collect and use:

- **Primary data** – you collect this
- **Secondary data** – you use data collected by other people

**Primary Data**

You should discuss with your tutor and class the different ways in which you can collect your data; however, some common approaches to primary data collection are:

- Field notes which include your own observations, thoughts, feelings, interaction including dialogue, what you sense etc.
- Maps and diagrams: drawn on site, these could include the placement of furniture, office layouts, streets and public spaces, the journeys of people, animals, vehicles, objects
- Tally or scoring sheets, a means of recording observed behaviour from which it is possible to work out patterns
- Surveys: typically using questionnaires with a prescribed set of questions, which can take different forms and will elicit different responses
- Interviews: recorded (for example by digital recorder), photograph, film etc., and with materials generated in the interview process by the research participants
- Time logs, diaries, budgets etc.: the research participant collects their own data which reveals a high level of detail, micro decision-making, and allocations of time to different kind of tasks
- Documents: letters, diaries, photographs, postcards, newspapers, ‘official’ documents such as annual reports and policy statements

**Secondary Data**

Secondary data refers to your use (or re-use) of data collected by other people.
The common source of secondary data are:

- Government sources of published statistics
- Associations, for example, trade unions, professional associations and membership groups, political parties, interest groups and agencies
- Statutory and charitable bodies
- Large, reputable, well-established global organisations such as the EU or UNESCO
- Archive organisations and documents

As you can see, the diversity of data collection which you might obtain lends itself to a range of quantitative and qualitative techniques. This chapter will concentrate on quantitative techniques, and Chapter 9 will concentrate on qualitative techniques.

Quantitative Techniques

Secondary Sources: Online Surveys and Compendiums of Data

In practice, much of the data that you can access nowadays will meet the needs of at least part of your research project. You might well find that it will suffice for all of it. There is a wealth of data sets available, and many useful compendiums such as bulletins, fact sheets, and yearbooks, news updates and so on. Data is easier to access than ever before, with more sophisticated searches and ready-made inquiries and reporting available which you might well be able to use quite successfully.

The advantages of using secondary sources include:

- Learning from others about good question formats, sampling strategies which have worked, possible response rates
- You can frame and complement your research with data which has already been collected
- Being able to use them as your main method: this might be the only alternative in some cases, especially for practical reasons of time and cost as well as methodological reasons, for example, historical timeframes

Some of the disadvantages of using secondary sources include:

- The social construction of reality by the choices of categories, measurement techniques, and the write-up is out of your hands
- The data was collected to meet the aims of research projects which do not share the aim of yours
- Changing definitions over time and between places might present issues of comparability
- The margin of error and quality of errors might be unknown or misjudged

Even so, before setting off to do your own project, do take time to familiarise yourself with the range and quality of the surveys and datasets readily on
offer. Table 8.1 shows some examples of the range of accessible surveys and compendiums of data which are generally regarded as robust and dependable.

**Remember:** It is always worth checking for secondary sources first, not only to frame and contextualise your research but also with a weather eye to the potential usefulness of the datasets in fulfilling much of the needs of your research project. Both undergraduate and postgraduate students may find that there is plenty of good data around that is fit for purpose – and most websites today have improved with user-friendly, online search and reporting facilities.

### Task 8.1

**Using NOMIS**

NOMIS provides data about labour markets in Britain since 1971, and can be found via the Office for National Statistics website. The NOMIS site will collate and present data in a range of ways, including maps, graphs and diagrams, and spreadsheets. You can select sets of summary statistics which package data by geographical area, presenting labour market profiles either by ward, local authority, or parliamentary constituency. These are a good, quick, easy starting point. Alternatively you can go freelance using the Wizard enquiries to obtain more detailed data defined by your specific choices. Using Wizard allows you to make comparisons between different areas of the dataset and so on. There is an Advanced Query option. Once you have become familiar with NOMIS and what it can do, this is quite useful.

The first task:

- Find NOMIS
- Choose the Summary Statistics tab
- Find Local Authority
- Enter your postcode
  
  a) What percentage of people are economically inactive?
  b) What percentage of the population have no qualifications?
  c) What is the hourly pay of full-time workers? How does it compare with the GB average?

The second task:

- Find NOMIS
- Choose the Detailed Statistics tab
- Find Wizard Query
- Find Jobs Density
- Make choices of selection, clicking on the Next tab when a selection has been made, until a table appears. You can also request a map, which is interesting
# Table 8.1 Secondary Sources: Examples of Where to Find Good Online Data

<table>
<thead>
<tr>
<th>Source</th>
<th>Various sources including British Social Attitudes (see Box 8.1) and the Health Survey for England. Look under topic areas, e.g. ageing, disability etc. Also look for the Survey Methods Unit which deals with methodological issues.</th>
<th><a href="http://www.natcen.ac.uk">www.natcen.ac.uk</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>The Office for National Statistics (now governed by the UK Statistics Authority)</td>
<td>General Household Survey; Labour Force Survey; Social Trends; Regional Trends.</td>
<td><a href="http://www.statistics.gov.uk">www.statistics.gov.uk</a></td>
</tr>
<tr>
<td>The Centre for Longitudinal Studies</td>
<td>National Childhood Development Study, 1958 cohort; British Cohort Study 1970; Millennium Cohort Study.</td>
<td><a href="http://www.cls.ioe.ac.uk">www.cls.ioe.ac.uk</a></td>
</tr>
<tr>
<td>Scottish Government</td>
<td>Scottish Health Survey; Scottish Household Survey; Scottish Crime and Victimisation Survey; Scottish House Condition Survey; Public Attitudes to House Condition Survey. Also look for the link to Scottish Executive Research which has its own pages, and the Index of Multiple Deprivation reports.</td>
<td><a href="http://www.scotland.gov.uk">www.scotland.gov.uk</a></td>
</tr>
<tr>
<td>Northern Ireland Statistics Agency</td>
<td>Central Survey Unit; Census; wide range of topics is considered here etc. Look for the Northern Ireland Longitudinal Study (NILS) and the Northern Ireland Neighbourhood Information Service (NINIS).</td>
<td><a href="http://www.nisra.gov.uk">www.nisra.gov.uk</a></td>
</tr>
<tr>
<td>Welsh Assembly Government</td>
<td>A wide range of topics are covered including demography, health, housing, education, training, agriculture, industry, the economy, local government, transport, the environment and the Welsh language. Look for: 'statistics' under Topics. Also look for useful compendiums and bulletins e.g. StatsWales.</td>
<td><a href="http://www.wales.gov.uk">www.wales.gov.uk</a></td>
</tr>
<tr>
<td>World Health Organization</td>
<td>A wealth of data on matters including mortality; health.</td>
<td><a href="http://www.who.int">www.who.int</a></td>
</tr>
</tbody>
</table>
### Collecting Data: Quantitative Methods

<table>
<thead>
<tr>
<th><strong>Table 8.1</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The Social Sciences Data Collection based at the University of California at San Diego (SSDC)</strong></td>
<td>Portal collating many rich sources of data on the USA and the world. Includes the US Census, election and polling studies. Some sources are unavailable to non-SSDC subscribers.</td>
</tr>
<tr>
<td><strong>General Social Survey</strong></td>
<td>American attitude data. The main areas covered in the GSS include socioeconomic status, social mobility, social control, the family, race relations, sex relations, civil liberties, and morality. Also look for the Roper Centre.</td>
</tr>
<tr>
<td><strong>Inter-University Consortium for Political and Social Research (ICPSR) at Michigan</strong></td>
<td>The world's largest archive of social sciences data.</td>
</tr>
<tr>
<td><strong>Council of European Social Science Data Archives (CEESSDA)</strong></td>
<td>Portal to research in Europe including links to many key databases; data surveys, election studies, longitudinal studies, opinion polls, and census data. Look for the European Social Survey; Eurobarometers; and the International Social Survey Programme.</td>
</tr>
<tr>
<td><strong>European Union website</strong></td>
<td>A dense website with data on Euro themes: general and regional stats; the economy and finance; population and social conditions; industry; agriculture and fisheries; external trade; transport, environment and energy; science and technology.</td>
</tr>
</tbody>
</table>

*Continued*
Table 8.1 (Continued)

- Look also for key indicators;
- also useful annual publication,
- search under Eurostat Year
- Book – this provides access
to very useful smaller
- compendiums of data including
- The Pocketbook which contains
- the Key Figures on Europe.

Box 8.1 Spotlight on the British
Social Attitudes Survey

The British Social Attitudes Survey is an example of a large-scale, long-term, regularly conducted government survey.

Since 1983, the National Centre for Social Research has conducted annual surveys into changing attitudes in Britain. The results are published annually in the report British Social Attitudes, which is also available online.

The British Social Attitudes Survey collects data on a very wide range of areas, from gender issues, healthy diets, and civil liberties, through to marriage, family and friends, local government, the countryside, education, transport etc.

Over time, topics change and are redefined, new ones are added and older lines of inquiry fall away. The shifts in definition, tone, discourses and so on makes the Survey itself an interesting source of data. However, an undoubted richness lies in the content of the responses as well, i.e. the data on offer. The website provides access to many of the lines of inquiry and can generate comparative data, which saves time when working with the reports in hard copy.

Below is an example of the kind of data which can be obtained, comparing responses to a line of inquiry pursued twice in the British Social Attitudes Survey, in 1995 and 2003.

Example

How much do you agree or disagree with each of the following statements?

‘Immigrants take jobs away from people who live in Britain.’
Challenging Numbers: Things to Know about Data Sets

While useful, data generated in these datasets cannot be taken at face value. There are questions to be asked by researchers when working with data. These include:

- **Who has produced it?**
  This raises issues about who paid for it, vested interests, credibility

- **Do the numbers stack up?**
  This raises issues about what is claimed and whether the evidence supports such claims. Avoid fallacies – either those of others, or making your own through an incorrect reading or application of data

- **How was the data collected and what are its limits?**
  It is good research practice to always discuss data collection procedures, problems encountered, and the limits of the study set in train by the research design. Know what the data is and where it comes from, before attempting to use it in your study. Consider what bias from the perspective of your study might be built in, and consider how categories are socially constructed
They can be structured in different ways in order to allow relatively spontaneous, open responses (semi-structured) or more directed, closed responses (structured). You can also use scales to gain views or perceptions of the action. See examples of the kinds of questions which can be asked in Figure 8.1.

Disadvantages of surveys include:

- Once you’ve started you must finish. Unless you treat your first set as a pilot study you are on a course of action which is inflexible; changing direction means that you risk making your study invalid.
- They are not good at uncovering meanings, feelings, awareness, motives etc.
- Generating a lot of data: this can often be coded (see next chapter) and managed using software.

---

**Box 8.2 Spotlight on Social Trends**

*Social Trends* is a useful compendium of data with pithy and useful commentaries. Various areas are examined including demography; households and families; education; the labour market; distribution of income; expenditure patterns; health; welfare; crime; housing; the environment; transport; changing lifestyles; inclusion.

**Compilation Sources**

Note that a list of all of the major government surveys, data sources, archives etc., used is contained at the back. Table 8.2 shows examples of a few entries in the compilation sources, which demonstrates the richness of their range, as well as providing useful information about their sampling frame, frequency of data collection, response rates and so on.

*Source: Social Trends* 38 (2008: 16)

**Website**

www.statistics.gov – look for the product *Social Trends*

---

**Doing Surveys: Practicalities**

Often, the data available is not the right kind, amount, or accessible. This means that you need to collect your own. For many people, surveys are synonymous with using quantitative methods, although as we will see later, observation techniques can also yield quantitative data.
The advantages of surveys are:

- They can be used to collect data from a large number of people
- They are good for asking questions requiring factual answers

**Designing Questionnaires: Setting Questions**

Ensure that all of the research instruments which you use are clear to readers, easy to use and can be used in the same way over and over again. Group questions together into themes (for example, have one section on car parking availability; a different section on charges) and put the most important questions first (in case the respondent doesn’t complete the whole questionnaire). Include a section for the respondent’s details, such as demographic information (for example, whether they are male or female) and their relationship to the survey (for example, in the case of a survey concerning car parks, is the respondent a driver or a passenger).
**Example: structured question**

> Did you find a car parking space within ten minutes?
> Yes / No / Don’t Know

**Example: semi-structured question**

> How did you go about trying to find a car parking space?

‘I drove up Porter Street but there wasn’t anything there, then I went around the back of Carport and there was a space but I was too late. In the end I went to the meters by the library.’

**Example: scaled questions**

Your questions will allow you to collect different kinds of quantitative data:

- **In categories**
  For example, patient, doctor, visitor, receptionist, cleaner. These are nominal categories which don’t have a mathematical relationship unless you assign one.

- **Ordered values**
  This is where people can select from categories such as ‘strongly agree’ or ‘disagree’; ‘tallest’ as opposed to ‘shortest’ etc. These are categories which do have a mathematical relationship, indicating for example whether people or things are heavier or lighter; taller or shorter; bigger or smaller; stronger or weaker.

- **Scale values**
  A few examples of scale value categories are between 16–21 years old, 22–27 years old, 5–10 kg or 11–16 kg. These are categories which can be accurately put into an order.

**An Example of Using a Scale – the Likert Scale**

The Likert scale is a common device used to prompt responses, where the respondent selects a point on a continuum. Likert scales generally use five measurement points or choices, although some applications also use an additional category or two, for example ‘Don’t know’. The key quality is that they avoid the trap of using too few categories, such as three. When faced with three possible choices, there is a tendency for respondents to plump for the middle one. Using an uneven number and aiming for around five choices, i.e. providing a sufficient range, pushes respondents to make a choice while not overwhelming them with options. A Likert scale can be arranged in a list, such as the one below, or sometimes the selection points might be arranged as a horizontal bar with numbering attached to each choice, so that ‘Not at all successful’ might equal 1 and ‘Very successful’ would equal 5. The bar approach can speed up the data collation stage.

How successful were you at finding a car parking space in town today? Please tick one of the following:
Collecting Data: Quantitative Methods

<table>
<thead>
<tr>
<th>Not at all successful</th>
<th>☐</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not particularly successful /</td>
<td>☐</td>
</tr>
<tr>
<td>Fairly successful</td>
<td>☐</td>
</tr>
<tr>
<td>Quite successful</td>
<td>☐</td>
</tr>
<tr>
<td>Very successful</td>
<td>☐</td>
</tr>
</tbody>
</table>

Figure 8.1 Kinds of Questions: Structured; Semi-structured; Scaled

You will be able to do different things with the data achieved, either manually or by using a software package such as SPSS – it is well worth you building your coding design into your questionnaire design. You will need to work out what you will do with the data as part of designing the instrument to collect it.

It is important to spend time on devising the questions: see Table 8.3 on setting questions.

Online Resources for Accessing Surveys and Designing Excellent Questionnaires

In order to assist you to design excellent questionnaires, see the sources available online in Figure 8.2, Useful Sources for Working with Surveys.

1) The STILE questionnaire database and useful links available from the Higher Institute for Labour Studies

The STILE questionnaire database provides access to an international selection of organisational surveys and the questions that are currently in use. You can run queries to find the ones best suited to your purposes. For instance, surveys can be selected by use in particular countries and by their design criteria such as questions suited to longitudinal studies. In ‘advanced research’ you can specify additional criteria such as languages or type of scale. There is a link from each question to the survey where it comes from. You can also print your selection for your ease. (See: www.stile.be/HomeAndNews.htm)

2) Question Bank: University of Surrey

The Question Bank is an information resource in the field of social research, with a particular emphasis on quantitative survey methods.

The website has been designed to help users locate examples of specific research questions, and to see them in the context within which they have been used for data collection. It is intended to assist with the design of new survey questionnaires, the search for data for secondary analysis, and the teaching of survey research methods. It is funded by the Economic and Social Research Council in the UK. (See http://qb.soc.surrey.ac.uk)

Figure 8.2 Useful Sources for Working with Surveys and Designing Excellent Questionnaires
### Table 8.3 Setting Questions: Pitfalls and Fixes

<table>
<thead>
<tr>
<th>Example Question</th>
<th>Problem</th>
<th>Fix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you not go to the match with James who didn’t want you to go?</td>
<td>Too difficult to understand without further explanation</td>
<td>Avoid using negatives, and simplify and divide questions if you need that information: Did you go to the match? Was that with James? Did James want to go to the match?</td>
</tr>
<tr>
<td>Do you come here often?</td>
<td>Open to misinterpretation: invites a range of responses (comical, offence and so on) none of which are the kind of response that you are looking for</td>
<td>Clarify the question, use neutral language and be specific about what you want to find out, for example: How often do you visit this shopping centre?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>i) Every day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii) Twice a week</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii) Once a week</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv) Once a fortnight</td>
</tr>
<tr>
<td></td>
<td></td>
<td>v) Monthly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vi) Annually</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vii) Never been before</td>
</tr>
<tr>
<td></td>
<td></td>
<td>viii) Other</td>
</tr>
<tr>
<td>What would it take to persuade you to re-use your plastic bags?</td>
<td>Makes assumption: the respondent is not re-using bags</td>
<td>Be specific about what it is that you want to find out:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>i) Are these new bags or did you already have them?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii) Explore reasons</td>
</tr>
<tr>
<td>Are you flying to France and then changing for Vietnam when you get there?</td>
<td>Asks at least two questions in one; invites an answer to only one part or a confused response, possibly with accompanying explanations which may/not be clear</td>
<td>Again, simplify the questions, splitting them up into easy to understand and answer queries:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>i) Are you flying to France?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii) Will you be flying on to Vietnam?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii) When will you be flying on/how long will you stay in France?</td>
</tr>
</tbody>
</table>
### Table 8.3

<table>
<thead>
<tr>
<th>Example Question</th>
<th>Problem</th>
<th>Fix</th>
</tr>
</thead>
<tbody>
<tr>
<td>So, you were 49 years old or was it 50?</td>
<td>A potentially sensitive question asked boldly; might also be too specific for respondent to answer easily or smoothly since if respondent does not view this as a sensitive question, they might be hard pressed to quickly work out the correct answer. In either case, the question invites some inaccuracy in the response.</td>
<td>Allow greater vagueness on the part of respondents; ask permission for information by showing how this would help you. For instance, phrase it: May I check my understanding so that I am clear about how old were you would you say at that point?</td>
</tr>
<tr>
<td>When you found the £10 note on the floor, did you do the right thing and hand it in to the shop assistant?</td>
<td>A leading question loaded with value judgements: invites inaccurate or confused reply unless affirmative</td>
<td>Decide what the true line of inquiry is about. There are two options here: (1) Rephrase the question to focus on the behaviour not the attitude in the first instance; and then use prompts. For example: What did you do when you saw the £10 note? Do you recall what went through your mind? (2) Pose the question as a vignette: People sometimes come across money lying on the ground. What do you think they should do when they find: i) £10 lying in the street ii) £10 lying on a shop floor</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Example Question</th>
<th>Problem</th>
<th>Fix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you enjoy the film?</td>
<td>An incorrect use of scales. Firstly use five rather than three choices, since there is a tendency for respondents to drift towards the centre or middle choice. Secondly use language which is less ambiguous – ‘Mainly’ is too vague and open to too many interpretations.</td>
<td>iii) £10 on the stairs at a birthday party in a friend’s house. I really enjoyed the film. Do you:</td>
</tr>
<tr>
<td>Yes quite a lot</td>
<td></td>
<td>Strongly agree</td>
</tr>
<tr>
<td>Mainly</td>
<td></td>
<td>Agree</td>
</tr>
<tr>
<td>Not really</td>
<td></td>
<td>Neither agree nor disagree</td>
</tr>
<tr>
<td>How much do you earn in salary a year? (excluding bonuses, savings interest or dividends, benefits and pensions):</td>
<td>This attempt at using scale values to generate factual data needs to be staged with intervals which are: • Consistent • Meaningful</td>
<td>Less than £5k £6–10k £11–15k £16–20k £21–25k £26–30k £30–35k £36–45k £46–55k £56–65k £66–75k £76–90k £91–105k £106–120k £120–150k £150+</td>
</tr>
<tr>
<td>£10–15k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>£15–25k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>£26–40k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>£41–70k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>£71–180k</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Task 8.2**

**Devising Questions**

Imagine that you are conducting an opinion poll investigating political beliefs. Devise five questions which would allow you to investigate public opinion relating to foreign policy.
Collecting Data: Quantitative Methods

Share them with your colleagues.

What different questions formats were used, and in what combinations (for example, what proportion of closed to open questions)?

What sources of inspiration were used?

What different scales were used?

How persuasive were the rationales for their choices presented by your colleagues?

Designing the Questionnaire: Appearance

When designing your questionnaires make sure that you have got your audience in mind. You need to make everything very straightforward for them. This isn’t patronising, but recognises that:

- You are very close to your research, and what seems like common sense to you might well be rather mysterious to others
- People aren’t always very good at following instructions

Tips on Physical Design

- Write for your audience. The audience are the research respondents – not your university tutors, friends or family
- Avoid jargon, sophisticated phrasing, and overly complicated questions; this will just bamboozle your respondents, especially if they are trying to complete the questionnaire quickly
- Don’t give your respondents ‘difficult’ tasks to do, for example, questions requiring mental arithmetic or exercises requiring spontaneous drawing, unless you have been able to pilot the study and know that this is presented in a way which definitely works
- Make sure that your questionnaire is quick to complete, to ensure that respondents do make it to the end
- Avoid ‘over-designed’ questionnaires with fancy fonts; lots of colour; a high level of decoration; borders; patterned paper etc., this will confuse your respondents
- Use plain white paper and a simple to read font size of at least 12 and perhaps 14 points
- Use as few sheets as possible, the less the better
- Test your questionnaire out on a few people from different social groups including acquaintances: a hall porter; a friend; someone at work. Take heed of their queries and hesitations – these are signs that they didn’t immediately understand the instructions or questions, and take on board any
other data or feedback on offer, for example, the length of time it took for
them to do it etc.

Remember: Beauty is in the eye of the beholder – and words are in the eye
of the reader.

Look at Appendix 2 to remind yourself of the ‘readability’ issue
which you will need to address in the design of your research
instruments and any covering documentation, such as information
sheets.

Task 8.3
Evaluating a Survey

- Find a survey: this might be a customer satisfaction survey; a
  public information survey; an online media inquiry; a local sur-
  vey relating to services
- Complete it
- Reflect upon its design from all points of view: the kinds of
  questions, layout, readability, how long it took you to do,
  etc.
- Identify two improvements which you would make. Give reasons
  for your choices

Administering the Survey: Finding a Sample

In theory, perhaps the best approach is to attempt to access the
whole population. In practice, this is extremely difficult, for issues of
time and cost, as well as generating huge and unmanageable data
sets. Thus, the need to draw a sample arises. The sample needs to be
large enough so that there is some security in terms of the breadth of
the sample which is likely to gain coverage and which allows gener-
alisations to be made. However, samples are by nature sub-groups,
so thought has to be given to working out where and how to find a
sample.

The first step is in deciding which segment(s) of the whole population
you will seek to access, i.e. your sampling frame, a methodological deci-
sion determined by the needs of your research project. But the second
task is to actually track people down, persuade them to participate, and
maximise the chances of them doing so. See Table 8.4, Methods for Finding a Sample, as the first step.

**Administering the Questionnaire: Maximising the Sample’s Response**

Surveys can produce a significant non-response. This is so even for surveys with government or corporate backing, substantial resources, a well-organised system to carry them out, and with incentives and penalties for respondents. Without resources and publicity it is impossible to achieve a whole population response, but even mobilising a sample to reply can be difficult.

Surveys can be conducted in different ways such as by post, email, in person or by telephone. These methods can be quick to carry out and relatively cheap in terms of the costs per unit of data. However, different methods used for different kinds of audiences will increase or decrease the chances of a good response rate. You need to match the approach with the likely behaviour of the group in which you are interested. For example, the received view is that conducting questionnaires by telephone can increase the response rate over alternatives such as the postal questionnaire, although neither produces such a high response rate as face-to-face interviews. Response rates to questionnaires carried out for the purposes of social research may be as low as 10–15% (Donscombe, 2003) while in market research this may be even lower, at less than 5%. Mobile phones might be better suited to some audiences than others, while ringing home phone numbers of established, mature households means that your call will be abruptly ended. In person research conducted on a railway station concourse full of hurrying passengers produces a worse response than in the station waiting room.

**Conducting Surveys by Post**

Postal surveys can produce good quality responses, since those that commit to completing and returning the questionnaire may well spend time actually filling it out. However, notoriously, postal questionnaires are generally returned by relatively small percentages of the population, as low as five per cent is not uncommon especially if the questionnaire is complicated. Thus, to gain an adequate number of responses, the researcher generally has to build large samples into the design, increasing the cost and complexity of carrying out the exercise, and might have to send out the questionnaire more than once, or send a
### Table 8.4 Methods for Finding a Sample

<table>
<thead>
<tr>
<th>Sampling Method</th>
<th>Description</th>
<th>Best Uses</th>
<th>Potential Drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snowball sampling</td>
<td>Each new subject refers the researcher on to further subjects</td>
<td>To access hard to reach</td>
<td>Relies on the judgements of various participants which may not be consistent and thus present groups which are unsystematically defined or consistently biased in a particular direction</td>
</tr>
<tr>
<td>Random sampling</td>
<td>All subjects have an equal chance of being selected</td>
<td>Increases representativeness of the sample</td>
<td>Difficult to define the parameters of the whole population and difficult to definitely include all possible candidates in the whole population from which to select, therefore may produce bias</td>
</tr>
<tr>
<td>Nth number sampling</td>
<td>Every nth name is selected for inclusion in the sample</td>
<td>Can be a quick and easy selection method to use when handling large amounts of data, for example, when using a database</td>
<td>Need to be sure that the way that the data is sorted and presented contains no particular order, i.e. that its order is truly random</td>
</tr>
<tr>
<td>Stratified sampling</td>
<td>Sub-groups are defined in the data, and the number of subjects selected from each group is weighted according to the overall importance of the sub-group to the population as a whole</td>
<td>A good way of checking the power of sub-groups which otherwise distort readings due to their increased probability, for example, of being selected, or their propensity to be selected</td>
<td>Need to be sure that the definition of each sub-group does not lead to bias in terms of creating or overstating significance; also need to ensure the allocation to each sub-group is accurate</td>
</tr>
</tbody>
</table>
remind for completion. This is also the case when using email. A quick and easy-to-complete online questionnaire (to the right audience) can achieve a higher response than traditional snail mail. However, a long questionnaire which crashes halfway through does not produce any favours.

### Conducting Surveys by Telephone

Telephone surveys tend to produce a better response rate, especially if the researcher has been able to use a gatekeeper to obtain lists of potentially interested respondents. For example, a researcher working through a list of members of a club or association which has given consent to the research and has forewarned its members may get a much better response than when cold calling names from a telephone book. In both cases, it would be possible to attend to the socio-geographical dispersal of the sample, this being much more difficult with face-to-face interviewing which may be affected by the constraints of the interviewer's capacity for travel and availability. Yet Figure 8.3 demonstrates some of the issues around sampling which arise for both telephone and postal questionnaires.
Telephone interviewing:
Some groups are automatically excluded:
1) Those without a phone (social exclusion bias)
2) Those who are ex-directory (creating gender and middle class bias)

Some groups are less likely to respond at certain times, days, dates, seasons, for example:
1) Faith communities
2) Shift workers
3) Carers (gender bias)
4) During major holiday periods

Postal Questionnaires:
Groups are more likely to respond:
1) Where respondents feel strongly about the topic, for example, when there is a local opinion survey of proposed planning permission in the high street
2) If they are well-organised; with time to complete the questionnaire
3) When they have a good level of literacy; they can cope with the questionnaire without further explanation

Some groups are less likely to respond:
1) The busy and generally stressed may not see questionnaires as a priority
2) If the questionnaire touches on subjects which are considered sensitive for the population on the basis of, for example, invasion of privacy; embarrassment or shame; sex, religion, politics; financial matters, especially income, debt or tax related

Further decisions by the researcher can affect the sample and outcomes: what kind of telephone to use – mobile or land line? Should we use snail mail or email? But note that while these choices will influence the social characteristics of the sample selected and thus the kinds of data obtained, decision-making such as this can be used to specifically target populations.

Figure 8.3 Social Factors Which Affect Sampling

Box 8.3 Spotlight on the Census

The Census dates from 1801, and has occurred every ten years apart from 1941 when it was disrupted by war. The Census Act of 1800 was passed amidst fears that the growing population would outlive the food supply, prompted by the debate following the publication of Thomas Malthus' essay on the 'Principle of Population' in the late eighteenth century. The original data collection was carried out by the Overseers of the Poor in England and Wales, and schoolmasters in Scotland. In 1841 the Registrar General's post took control of the Census.
The collection aims to collect data from 100% of the population, with penalties for non-completion. In order to achieve maximum returns and to avoid imposing penalties, a system of door-to-door visits and postal returns has been devised and perfected using Area Monitors and other workers.

The questions posed, their shifting definitions and the sometimes controversial public debate surrounding them are themselves interesting data. Take for example the question in 1871 which asked respondents to declare if they were ‘lunatics, imbeciles or idiots’ (this was abandoned in 1881). Today, current debates surround how best to define and count various categories, what ethical issues are raised in doing so, and whether doing the Census is even necessary or within the principles of human rights. Controversies take place around specific lines of inquiry including religious belief and practice, and ethnicity and nationality. Censuses are found not just in the UK and USA but all over the world. Google, for example, the Census in Niger in 2006.

**Samples of Anonymised Records (SARs)**

The SARs are a recent innovation. They allow researchers to access samples of survey data (between 1% and 5% of the population) without needing to access or deal with the whole of the census. Drawn from the 1991 and 2001 Census, the SARs allow multivariate analysis of the full range of categories found in the Census.

Further information relating to SARs in particular can be found at the Cathie Marsh Centre for Census and Survey Research (CCSR)

**Websites**

www.ccsr.ac.uk
www.census.ac.uk

**Bias in the Response**

Bias in the response is generally caused by:

- Decision making over administering the questionnaire which causes a concentration of responses to come from some groups rather than others
- Non-responses as a result of refusal: this is itself an important source of data, indicating taboo subjects, obstruction, and political resistance
- Miscommunication of the research question: this can be systematic, indicated by significant numbers of respondents apparently misunderstanding the question
- Demographic bias which is built into the questionnaire or the process by which it is administered
Uncovering Demographic Bias

When designing your questionnaire, decide what demographic factors it is useful to know: age, gender, race, religious identity, gender, postcode and so on. When looking at datasets, do checks on these demographic groupings. Look for:

- Over representation, or heavy clustering: a large percentage of respondents were from a particular group
- Under representation – the absence of groups including hard to reach groups
- Selective responses – a particular group ‘refused’ a particular question or set of questions

Dealing with Bias

In order to deal with bias, a few steps can be taken:

- Identify possibilities of bias and either take steps to mitigate it, or take steps to use it as a vehicle to obtain certain data from certain respondents. For example, keep a rough tally sheet of respondents so that you can seek out those who are at risk of becoming under represented
- Reflect upon whether using a split method for conducting surveys would be useful, for example, using mobile phones for 14–19 year olds, and postal surveys for 40–55 year olds. If a split method is used, then the consequences of this for the integrity of the data and bias which might arise need to be evaluated in the data analysis and the findings reached:
  - Ensure that all research instruments used are geared to levels of reading and speech which are lower than university undergraduates might assume. This increases accessibility. See Appendix 2 for some information about ‘readability’ including a couple of examples of measurements of reading ages. The likelihood that more respondents will interpret the question in the same way.
- Weight the data during data analysis. This is to compensate for the influence of under- or over-represented groups

Data Collection by Observation: Record Cards and Tally Sheets

There are alternatives to collecting data by questionnaires. Alternative techniques do different things, for example, tally sheets allow researchers to collect data relating to behaviour by observation, with limited or no interaction with the (unwitting) research participant. This allows a system
of scoring which has been used in various ways, for example, recording children’s play behaviour; hospital visits; traffic flow; use of shopping centres; animal–human interaction.

Tally sheets typically contain a system of categories and tick boxes, which can be speedily completed in a systematic way. An example can be seen in Table 8.5, Tally Sheet: The Dental Surgery Waiting Room.

The advantages of tally sheets are that they:

- Allow the researcher to capture events while in mid-flow
- Allow the researcher to allow speedy data capture
- Are non-invasive: don’t interrupt the flow of events
- Are not influenced by interaction with the research participants
- Can be used unobtrusively

The main disadvantages with tally sheets are that they:

- Don’t access meanings and can’t explain ‘Why?’
- Are carried out without the research participants’ knowledge or consent
- Allow certain events and objects to be captured but not others: are somewhat inflexible

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read a newspaper or magazines</td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>111</td>
</tr>
<tr>
<td>Used mobile phone</td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>1111</td>
</tr>
<tr>
<td>Browsed at surgery leaflets, treatments, costs etc.</td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Made further enquiry of receptionist after initial wait</td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Returned to car park</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>1111</td>
</tr>
<tr>
<td>Ate and drank</td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Went outside for a cigarette</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
</tbody>
</table>
Nonetheless, tally sheets can be a useful technique to use in a dissertation, even if they are not the only method used. They are quick and easy to use and produce a solid data set which can be pondered over later.

Summary

- There is a wide and critically engaged literature about quantitative and qualitative research techniques with which you should be familiar.
- The most common quantitative technique for data collection used by students is surveys although sometimes it is possible and preferable to use survey data from secondary sources (provided that they are reliable), for example, national surveys. There are a number of good secondary sources of survey data, and nowadays these are more easily accessible online.
- There are alternative strategies for recording data including the use of tally sheets or logs. These allow the researcher to record observed behaviour in an unobtrusive way. They are quick and easy to use, although they do not provide meanings or allow research participants to contribute, making them research subjects rather than participants.
- A key challenge in doing successful surveys is linking your questionnaire to your sample. Sampling techniques vary, each has its own strengths and weaknesses in terms of bias and effectiveness. Ensure that you choose the sample that best fits your research question:
  - Random: all subjects have an equal chance of selection
  - Nth number: every n-th number (e.g. every 25th name) is selected
  - Stratified: samples are divided between sub-groups and a proportion selected from each sub-group depending on their importance in the sample as a whole
  - Convenience: respondents are easy to locate and access; they are ‘to hand’
- Having selected your sample, you then need to work out the best way of reaching them. Different approaches might mobilise some social groups more effectively than others. How you administer the survey (by post, in person, by phone, by email and so on) will affect the data which you collect. For example, doing so in person at the school gates will allow you access to waiting parents but not to parents at work.
- The general problem of bias can be created by a number of factors, including where:
  - The sampling technique itself is wrong
  - The method of administering the questionnaire affects the sample
  - Non-responses are too numerous and invalidate the study
  - The questions are ambiguous or the ‘wrong’ kind (for example, closed instead of open) and produce poor quality responses.
• You need to design your own research instruments. The most frequently used one is the questionnaire. Be careful how you set questions, and consider re-using and adapting questions from other sources such as Question Bank.

Further Reading


Chapter 9

Collecting Data: Qualitative Methods

Overview

The variety of qualitative techniques:

- Interviews of various kinds
- Observational inquiries, for example, of interaction and problem solving
- Case studies, which explored a particular setting, situation, organisation, institution, unit, arena, network, system in-depth
- Visual methods including the analysis of images, and tracking the production, consumption, and journey of objects through material culture
- Ethnographies

The biographical turn

Using documents

Oral history; life history

Doing interviews

Time logs and diaries

Visual methods

This chapter focuses on mainstream qualitative techniques that are located in a framework of biographical inquiry and lived experience. Examples of good practice are suggested in each section which explore, using documents: life histories; oral history; and data collection for these by interviews. Guidance on doing both one-to-one and group interviews is provided alongside discussions about structured, semi-structured and focused techniques. Another form of data collection is considered, that of time budgets, logs, and diaries which allow for detailed record keeping, often by the research participants themselves. Finally, the chapter considers visual methods.