

Chapter 7

Evaluating Information: Validity, Reliability, Accuracy, Triangulation

Teaching and learning objectives:

1. To consider why information should be assessed
2. To understand the distinction between 'primary' and 'secondary sources' of information
3. To learn what is meant by the validity, reliability, and accuracy of information
4. To consider some warnings about 'official data'
5. To consider further the distinction between 'facts' and 'truth'
6. To understand the origin of triangulation and its application to research
7. To consider methods of sampling which can be used to collect data.

'Do not feel absolutely certain of anything'.

Bertrand Russell, 1951.¹

Introduction

In Chapter 6, you read how published research reports can be assessed. The research component analysis and Rose's ABCDE model examined the completeness and coherence of the research process adopted. They also considered the validity or otherwise of the relationships between theory and hypothesis, concepts and indicators, empiric data and analysis, and conclusions. Research essentially involves the gathering or collection of data that addresses the research question and enables theory to be tested or developed. So the data from which answers to the research question are to be drawn must be appropriate in terms of its relevance and efficacy – 'fitness for purpose'. Much of this information will be drawn from published sources

that will be supplemented as necessary by new information specially collected for the research project. So this chapter therefore suggests how best you can assess existing data and seek additional material.

Many textbooks use *information* and *data* interchangeably. Some complicate matters by treating 'data' as a plural noun and therefore writing 'the data are ...' While this is grammatically correct (for *data* is indeed the plural of the *datum*), it can sound odd to students untrained in Latin conjugation. The author, Kingsley Amis notably described such Latin correctness as the practice of 'wankers' as opposed to 'berks' who used slipshod English (1977)². So this book follows everyday practice of treating data as singular. Politicians also tend to use the word 'evidence' to describe what they would wish us to regard as 'conclusive, compelling information' which either proves or, in its absence, disproves allegation.

But is there a real difference between *data*, *information* and *evidence*? Certainly, the dictionary meanings are similar. But some distinction is useful. Researchers tend to speak of *data* as the mass of disordered, raw material from which *information* (knowledge) is abstracted to provide *evidence* to support argument and conclusions. (Information technologists adopt a similar distinction by defining information as processed data sets attaining meaning). **Information** *informs*. *Evidence* supports conclusions. So it is helpful to conceive of research as involving three stages. First, the raw data is gathered. Second, the data is organised and distilled into information. Thirdly, evidence is abstracted from the information through processes of analysis and testing. But neither information nor evidence is self-evident: the material seldom 'speaks for itself'. Some interpretation is required. However, when interpretation is re-interpreted, some distortion of the original is inevitable. So some distinctions, criteria and tests are useful to weed out distortions and 'untruths'. The distinctions adopted are between *primary and secondary sources of information*. The criteria used are *validity, reliability and accuracy*. The main test adopted is **triangulation**.

Primary and secondary sources

The value of this distinction depends on which of the different definitions of primary and secondary adopted. Some authorities adopt the definition that primary information is data generated specifically for the research project whilst secondary information is data collected for other research. But, in this book, the 'majority view' prevails: that data is distinguished at the outset by its *provenance* (source). **Primary** data is original, unedited and 'first-hand' whilst **secondary** data is 'second-hand', edited and interpreted material. However, the distinction between the information that you generate in the course of our research and that which you have abstracted from other sources is valuable. I will therefore term this (after Huxley) α **data**

and β **data**.³ Wherever possible, Politics researchers prefer to use primary, eye-witness data recorded at the time by participants or privileged observers. The main sources of primary data used by Politics researchers are fourfold:

1. contemporary documentary (written) records including minutes, letters, emails and diaries
2. your interviews with key individuals, 'agents' and 'actors'
3. numerical records, e.g. election results, census data
4. your own observation and records of interviews, etc. and other events.

Other sources are popular songs, poems, paintings and cartoons, photographs, graffiti, murals (e.g. N. Ireland), T-shirts and videos. But beware, all records, however 'primary' incorporate some degree of bias, perception, interpretation, and editing, whether contextual, cultural, curatorial or deliberate.

Written primary records include accounts of meetings, minutes, diaries, letters, reports, telephone transcripts, telegrams, emails, and newspaper reports, etc. But how reliable, accurate and truthful are they? Who actually prepared them and why? Arguably, all accounts are partial because they are *functional*, i.e. designed to fulfil a purpose. But *whose* purpose? Most public records reflect the interpretation of those holding power. Foucault argued that the victors write history. Alternatively, how *reliable* are the diaries of (former British Labour Cabinet ministers) Richard Crossman, Barbara Castle or Tony Benn? Did their cabinet colleagues know that they were keeping diaries and change their behaviour accordingly?

Crossman offers telling insights into official records:

'Thursday, 28 July 1966

One of the disconcerting features of the recent crisis has been the Cabinet Secretariat's habit of suppressing whole sections of the minutes on the grounds that they are too secret to circulate. But this morning they didn't do that. The section on prices and incomes was reported at enormous length and most of what we said has been very adequately summarised. Of course, this means that the Cabinet Secretariat regards the whole subject as fraught with danger and was careful to record the arguments of the opponents. Cabinet minutes are highly political and the way they are written has enormous effect. By eliminating whole sections from the discussion and reporting other sections in full, the Secretariat can greatly affect the way a decision is interpreted in Whitehall'.

(Crossman, 1976: 590)⁴

But the BBC's former Political Editor, John Cole wrote of Crossman that:

'... [Crossman] had a brilliant mind, was a great polemicist, and a subtle – though sometimes self-defeating – operator. But if you were Constable Plod seeking a reliable

witness, he would not be your first choice. I sometimes wondered if he knew how to distinguish what he said to the Prime Minister from what the Prime Minister said to him’.

(Cole, 1996: 64)⁵

But even PC Plod can be an unreliable witness: Churchill’s bodyguard, Detective Inspector Walter Thompson was criticised by Churchill’s biographer, Roy Jenkins for exaggerating his importance in events (Jenkins, 2001: 552).⁶

‘Inspector Thompson in his two volumes of reminiscences is good at capturing the heart of the matter but less reliable on exact dates, times and places than might have been hoped for a meticulous detective’.

(Jenkins, 2001: 562).

As a general principle, all primary information in the form of records – other than those that you make through your own observations – should be treated with caution. A ‘health warning’ is necessary. You should always ask yourself:

1. who prepared the record?
2. why?
3. for whom was it prepared?
4. for whom was it intended?
5. for what purpose was it made?
6. who would have ‘corrected’ or otherwise altered the record before it was finalised?

A common misconception is to believe that numeric information is more trustworthy than other formats because it is less vulnerable to ‘spin’. But, because numeric records are generally regarded as trustworthy, they attract manipulation. For example, the TUC and ILO accused the Thatcher government of changing the definition of unemployment twenty-three times (between 1979 and 1991) to reduce the headline figure and therefore conceal the true extent of unemployment. The government responded to the criticism by saying that each new definition distinguished further between genuinely unemployed people and others claiming to be unemployed to obtain benefits. A similar charge was levied later against the New Labour government that the lower levels of unemployment recorded and reported had been achieved by accepting more readily claims (on mainly health grounds) for the (higher) incapacity benefit. You should therefore check numerical records for any changes of definition and any selective use of periods to enable worst records to be omitted.

By implication, data that is not primary must be secondary – *after the event, second-hand*. But it should not be discarded. Secondary information will include records

gathered from a number of separate, primary sources and may contain authoritative commentary and analysis. The source's interpretations and bias are important – especially of evidence of how events were interpreted at the time and later, and the moral relativism of value-judgements.

Validity, reliability and accuracy

As you learned in Chapter 6, social science research confers a special meaning to *validity*:

'the extent to which a measure, indicator or method of data collection possesses the quality of being sound or true as far as can be judged. ... in the social sciences generally, the relationship between indicators and measures and the underlying concepts they are taken to measure is often contested'

(Jary & Jary, 1995: 714).⁷

In effect, the validity of information is its relevance and appropriateness to your research question and the directness and strength of its association with the concepts under scrutiny. Often you will have to use best available information whose validity may be weak. For example, to what extent, if any, does the decline in 'sectarian violence' in N. Ireland post-2001 reflect a lessening of antagonisms between conflicting groups? Does the election of an opposition party reflect popular support for its manifesto or criticism of the outgoing government? Do declining rates of party membership reflect a lessening of interest in health and education? One measure that intrigues Politics researchers is the *counterfactual* – events that don't happen – as evidence of hegemonic domination.⁸ But how can researchers be confident that the absence of an event can be attributed to the omnipresence of another? One solution to this particular problem of problematic validity is for you to adopt a wider range of measures to reduce dependence on any one.

Reliability is, literally, the extent to which we can rely on the source of the data and, therefore, the data itself. Reliable data is dependable, trustworthy, unflinching, sure, authentic, genuine, reputable. *Consistency is the main measure of reliability*. So, in literary accounts, the reputation of the source is critical. In John Cole's view, Richard Crossman was not a reliable diarist. Indicators of reliability will include proximity to events, (whether the writer was a participant or observer,) likely impartiality, and whether, as the police say, the record was really contemporaneous or an eventide reflection on the day's events. Very few politicians admit to real failings: all too often, their own agenda appears to justify their actions or to criticise others. Tony Benn's diaries seek to portray the inner workings of

cabinet government. But Dennis Healey claimed (playfully) that Tony Benn ‘always seemed to be on the toilet every time a difficult decision had to be made’ (BBC2 interview). Accounts may have been ‘sexed up’ to promote sales. Biographies may be hagiographic. For example, Michael Foot’s biography of Aneurin Bevan uncritically portrays the Welshman as a wholly heroic figure, whereas my father – a fellow native of Blaenau-Gwent – told me how, after 1948, some local trade unionists called the Ebbw Vale MP ‘Urinal Bevan’. This epithet followed Bevan’s assertion that:

‘it is for the [Party] Conference to lay down the policies of the Parliamentary Party, and for the Parliamentary Party to interpret those policies in the light of the parliamentary system’

(Foot, 1973: 236).⁹

In contrast, Grigg’s biography of another, Welsh hero, Lloyd George, provided a ‘warts and all’ portrait (Grigg, 1978).¹⁰

Numeric data need not necessarily be reliable. The source – even official statistics – may not be wholly impartial. Populations may be undercounted (e.g. 2001 census). The samples used may be insufficient or not randomly selected. Confidence limits (margin of error) may be omitted. The rate of non-responses to questionnaires may be disguised. Respondents may not have been wholly truthful in their replies. For example, on the basis of replies to their questions, most opinion polls (wrongly) predicted a Labour victory in the 1993 general election. Inappropriate statistical techniques may have been used. But reliable witnesses may also be inaccurate on occasions.

Andrew Marr, John Cole’s successor as the BBC’s Political Editor and a former editor of *The Independent* is very sceptical of the reliability of modern-day news reporting by the newspapers and TV news services (Marr, 2004). He blames this on the competition to drive down costs, consequent reductions in the number of journalists, and their being confined to their desks where they must too readily accept the stories ‘fed’ them by professional press officers. He recommends readers (and researchers) to:

‘Know [which newspaper] you’re buying. Reporting is so contaminated by bias and campaigning, and general mischief, that no reader can hope to get a picture of what is happening without first knowing who owns the paper, and who it is being published for. The Mirror defines its politics as the opposite of the Sun’s, which in turn is defined by the geo-politics of Rupert Murdoch – hostile to European federation and the euro ... It is ferociously against Tony Blair, this is because Number Ten has been passing good stories to the Sun’.

(Marr, 2004: 251)

He also warns against news of research from:

'hundreds of dodgy academic departments put out ... to impress busy newspaper people and to win themselves cheap publicity which can in turn be used in their next funding applications'

(2004: 254).

Similarly, Marr explains that TV news editors are:

'biased towards exciting or unusual pictures; news that is refreshing or odd; and news that bears some relation to viewers' lives'

(2004: 291).

So anything that looks dull, 'stories about northern European countries, about buses, about old people, about infrastructure, banking, manufacturing, Whitehall and regeneration,' is unlikely to be televised. Marr argues that a task of TV news is to increase viewing figures – which means also retaining the viewers of the preceding programme – usually popular light entertainment of the 'soap' or 'chat show' genre.

Accuracy is sensitivity to change— especially of detail, e.g. dates, numbers, persons present, etc. Remember that some biographers deliberately add false detailed information to trap and sue plagiarisers.

Facts and truth

Once again, you will find that adopting a critical distinction between *facts* and *truth* is useful. **Facts** are the available data. They present incomplete snapshots of events. **Truth** is the reality behind the facts. Sometimes the facts may obscure the truth – perhaps deliberately so. A good example was provided to me by a leading academic. He privately described how he had critically reviewed a best-selling account of British rural life where that the author had misrepresented the facts by combining material from a number of interviews to represent a composite figure. The author had replied to the effect that his critic was unable to distinguish between the facts and truth.

Interviews

Interviews with **political elites** provide a major source of information in politics research. They may be undertaken by the researcher or, where personal access is not

possible, by watching video recordings of interviews in TV news and documentaries. But you must never assume that what you are told or hear is reliable and accurate. For example, a former Prime Minister told me that he strongly supported a specific White Paper. However, the Minister of State who claimed to have instigated the new policies told me that the Prime Minister had opposed the White Paper. We must always question (implicitly) the answers to our questions and look for signs of deception or self-deception by informants, e.g. the *coping strategy* of long-term prisoners who are guilty but believe that they are innocent, i.e. in denial. TV interviews (i.e. secondary sources) are highly edited – especially field interviews where a single camera is used or where the interviewee has been granted some editorial control.

Triangulation is the means adopted by researchers to secure effective corroboration. However, before this method is described, consider the case study below:

Case Study

Harold Nicholson provides a detailed narrative of the fall of the second Labour government in 1931 and its replacement by a National Government which was to last effectively until 1945 (Nicholson, 1953: 453–469)¹¹ The Labour Prime Minister, Ramsay MacDonald became the leader of the National Government in what became named by Labour party members as the ‘great betrayal’.

Nicholson describes the relevant background as the rapidly deteriorating public finances caused during the worst years of the Depression when the demand for public expenditure on unemployment benefit etc. grew whilst income from taxation fell. In response to demands by the Conservative and Liberal parties (amplified by the Tory press), the Government formed an independent committee under Lord May. On 31 July 1931, May recommended substantial cuts of up to 20% in public sector salaries, 20% cut in unemployment benefit and reduction in the pay of the armed services to 1925 levels. But two, of the six May members, issued a minority report dissenting from May’s recommendations on the basis that the costs of the cuts would fall mainly on the working classes. Nicholson wrote that ‘The rank and file of the Labour party agreed whole heartedly with [the Minority Report]; MacDonald and Snowden [Chancellor of the Exchequer] did not’ (1953: 455). Nicholson reports how, later that day, MacDonald formed a five-man, special, Cabinet Economy Committee to consider how May could be implemented. The ‘Big Five’ consisted of MacDonald, Snowden (Chancellor), J. H. Thomas, Arthur Henderson and William Graham. The likely continuing withdrawal of deposits held in London meant that the government would be unable to fund the public sector deficit without support from bankers in Paris and New York. The bankers were unwilling to lend the

Continued

Continued

money unless and until firm proposals were made to balance the UK budget. On 19 August, the Cabinet's Economy Committee proposed cuts similar to the May report. Half the deficit would be met by reductions in unemployment benefit and public sector pay and the other by increasing taxation etc. According to Nicholson, the proposals were reluctantly approved by a majority of the Cabinet with the exception of transitional unemployment pay (1953:457). They suggested an additional revenue tariff which MacDonald told them would not be accepted by the Liberals. The TUC – which had created and funded the Labour party – met MacDonald on 20th August. The Council refused to accept cuts in unemployment benefit or public sector pay. On 22 August, MacDonald proposed a modified version of the programme – including a 10% reduction in unemployment pay – to the Cabinet. He obtained Cabinet support to 'enquire' of the Opposition leaders whether the revised proposals were acceptable. The Opposition leaders responded that the overseas bankers support was critical to wider support for the package.

The King (George V) had been kept informed by MacDonald of the increasing crisis. He returned from Balmoral to London on 23 August when he was told by MacDonald that leading members of the Cabinet would not support the latest proposals (being considered by the bankers). The King decided that the 'correct constitutional course' would be to meet the leaders of the three main parties: MacDonald (Labour), Baldwin (Conservative) and Samuel (Liberal, as Lloyd George was in hospital). Nicholson describes how the King stated his preference for MacDonald and the Labour government to stay in office and to implement the cuts. If that were not practicable, then the best alternative would involve the formation of a National Government – headed by MacDonald – with a Cabinet drawn from three parties and commanding a sufficient majority in the House of Commons to approve the necessary legislation. The King met the three leaders separately who agreed to join a National Government if necessary. Once the crisis had been resolved, then new elections should be held.

Later that day, the Labour Cabinet met. After a long adjournment, they were advised by telegram from the Government's agents (the bankers, J. P. Morgan) that the necessary US public support for a public loan would be problematic until Parliament had approved the proposals. They suggested a short-term treasury loan. Finally, they enquired whether they were correct in assuming that the package proposed by the Cabinet had the support of The Treasury and City. But the Cabinet had not and would not approve the programme. MacDonald told them that he would report the divisions (eleven in favour: eight opposed) to the King whom he would ask to convene a meeting of the three party leaders. He would tell the King that the Cabinet had placed their resignations in his hands. He immediately reported to the King. Acting unilaterally, MacDonald began that night to plan the new Government

Continued

continued

with what Snowden described 'an enthusiasm which showed that the adventure was highly agreeable to him' (1953: 465). On 24 August, the King met the three party leaders who agreed to form a National Government. MacDonald tendered his resignation. He was then invited to lead the National Government. He asked them to prepare a communiqué saying that the formation of the new government was being considered. The party leaders agreed that the new Government would not be a coalition but a 'co-operation of individuals'. MacDonald described the proposals to his Cabinet and invited them to join a new 'Cabinet of Individuals'. With the exception of J H Thomas, Lord Sankey and Philip Snowden, they declined. Following the resignation of the Labour Cabinet, the new Cabinet was formed on 26th August with Baldwin as Vice-Premier.

There are a number of other accounts of this episode. One dispute among them is whether the King or MacDonald first raised the proposal for a National Government, i.e. whether MacDonald accepted the King's proposal to lead a national government out of patriotism and loyalty, or, proposed the arrangement as a means of continuing in office and increasing his power whilst appearing to follow the constitutional requirements of the King. For example, the celebrated Labour historian, G. D. H. Cole wrote that:

'The exact method of the split is vehemently disputed. The Labour Cabinet was still discussing the outrush of gold and the 'threat to the pound' under the influence of the deliberately exaggerated menaces of Philip Snowden, when it came. They had agreed to enormous concessions but jibbed (it is stated) at penalising the unemployed. Then it was put to them that arrangements had already been made, with the King's consent but clearly on MacDonald's initiative for the formation of a 'National Government' of Labour, Conservatives and Liberals'.

(Cole, 1938/66: 593)¹²

So how much importance should the researcher apply to Nicholson's account where it differs from others in critical respects? Obviously, the researcher will seek corroboration from other sources and assiduously compare the various clues. But just how *reliable* is Nicholson's account. Who is Nicholson? What was his book about? Why did he write it? What documents and witnesses did he have access to?

The answers are that Sir Harold Nicholson (1886–1968) was a distinguished diplomat, historian and biographer. He was educated at Wellington and Balliol

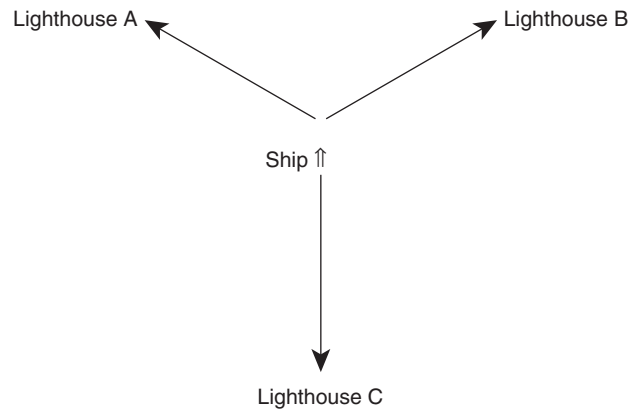
College. He attended the Paris Peace Conference (1919) as a diplomat. He retired from the service to become a writer. In 1931, he stood as an MP for Harold Mosley's New Party but left when Mosley formed the British Union of Fascists. He became a National Labour MP in 1935 but was defeated in 1945. He became a governor of the BBC. He was a 'man of independent means' who married Vita Sackville-West. They lived at Sissinghurst Castle. Both were bi-sexual and practised an 'open marriage'. However, their life-style was not unusual amongst the upper classes. We can therefore regard Nicholson as an Establishment figure albeit on its arts and literary wing. His book was entitled *King George V* and published in 1953, i.e. twenty-seven years after the King's death in 1936. He had been appointed by the Royal Family to write an 'authorised biography'. So Nicholson had access to the King's diaries and the official papers kept by his secretaries. He also interviewed the King's secretary, Lord Samuel, Herbert Morrison and other participants. However, we are unaware of the editorial control exercised by the Royal Family.

Clearly, the overall purpose of the biography was to make public King George's hand in the making of history. But Nicholson could not be accused of presenting a wholly flattering picture of the King. For example, he observed that George was neither very clever nor witty: he was a relatively dull man who was therefore entirely representative of the British people. Overall, we can probably conclude that Nicholson's account is probably very accurate in terms of the detail. However, we cannot be entirely sure about its *reliability*: the biography is more likely to portray King George V as a central figure, imposing wise, constitutional solutions, rather than acquiescing to what other commentators (like Cole) saw as MacDonald's devious scheme. No single account can ever be regarded as wholly reliable or accurate. Other sources must be sought and used. However, the number of competing sources may be huge. The most widely used method of selecting sources and materials from the range available is termed *triangulation*.

Triangulation

Triangulation is a method developed over the centuries for navigation and surveying. It provides the basis for satellite navigation. Its origins lie in geometry. A point can be precisely defined in space by the angle it subtends to a line joining two other points by the application of the Law of Sines. Any geographical area could be mapped by first selecting 'trig points' (triangulation points) that are a measured distance apart and then recording the position of any other point in terms of the angle subtended. In this way, a third trig point could be established and the exercise extended. In navigation, triangulation is used to establish a ship's precise position by taking bearings on three or more known landmarks. The position is

most accurately determined when the three points are equally located around the ship, namely



The same method is used in Politics research to obtain an efficient corroboration of any crucial account. *Triangulation* involves seeking accounts from three or more perspectives.

So, for example, a researcher investigating the General Strike of 1926 would seek to obtain accounts from the TUC, the Government and at least one source independent of the two adversaries. Politics researchers face a special challenge: the two main parties and their perspectives are often diametrically opposed to each so that any, third-party, independent views are gained from only one side. To overcome this difficulty, they seek as many independent sources as possible. Furthermore, Politics researchers seek to triangulate at each level of data media. So you should seek to triangulate between contemporary written records and news reports, autobiographies, personal interviews with participants, and other research narratives. You should also try to find new angles. However, given that each party will adopt different perspectives, then the 'truth of the matter' may be unique to each participant. You may be able to repudiate some accounts but you may find that you are unable to offer a definitive version of events. Indeed, the participants may be unsure of their real motivations or involvement. At the end of your triangulation, you may well know more about and understand better the particular event that the participants because you will have accessed records unavailable to them. But remember the adage that, whilst success has many parents, failure is an orphan. However, by demonstrating the application of triangulation, you will be able to show the reader the process by which corroboration has been sought. You should also be able to pinpoint both gaps in and inconsistencies between the accounts.

Sampling

You will quickly find that, even when you adopt triangulation, the volume and potential sources of data in terms of people and records can still remain vast. So a selective approach is essential. You simply cannot interview every member of a union or examine every council minute. The age-old solution to this particular problem is to concentrate your activity on a *sample* of the *population*. These everyday words have special meanings in academic research which warrant explanation. The *population* is the universe of all the subjects or cases under study. You must define your population. It may be all the members of a union, the residents of a city or town, states of Africa, a particular ethnic or age group and so on. The population is a *set* of individuals, cases, states, etc., which share a common characteristic. The *sample* is a selection of individuals, cases, states, etc., made from the population. The sample is, therefore, a *subset* of the population.

You must define your research population in a *sample frame*. Your **sample frame** is a list or schedule of the population from which the sample will be drawn. It may be a membership list or a directory of engineering firms.

You can use either *probability or non-probability samples*.

Non-probability samples

Non-probability samples are samples where members of the population do not have an equal chance of being selected. They are not statistically reliable. They cannot generate generalisable data. You select the members of the sample. Non-probability samples include *nomination*, *snowballs*, *volunteers*, *case studies* and *theoretical samples* (used in grounded research). Non-probability samples will not generate generalisable information. You have already been introduced to case studies and theoretical samples in Chapter 5. Non-probability samples are very small. They are used for qualitative research.

Nomination is the most widely-used form of recruiting a non-probability sample. Essentially, you ask a local *social gatekeeper* or intermediary to nominate (name) a group of people who meet your requirements for a research sample. A *social gatekeeper* exercises control over who enters a community. They may be head teachers, village headmen, employers or heads of family (if you want to interview children). Essentially, you search them out, write to them explaining who you are and the nature of your research, and ask them to nominate a cross-section of local people to interview. Often, they will offer to arrange the meetings. But this offer of assistance is incompatible with the ethical ideal of voluntary consent. The people nominated may feel as if they have been commanded to attend. This will inhibit your meeting: they will feel obliged to say what the gatekeeper would like them to say. It is better if you contact the named people individually (preferably in writing). Again you should introduce yourself, explain the nature of your research and its

benefits to them. You should say why you want to speak to them individually or at a (focus) group meeting, and add that *X* (the local social gatekeeper) has suggested that they might be able to help. *Intermediaries* are not social gatekeepers but trusted outsiders or *thresholders* who are respected by the community. Typically, they are ministers of religion, teachers, nurses or local leaders of voluntary agencies or NGOs (Non-Governmental Organisations) like Oxfam.

Snowballs grow larger as they are rolled across snowfields. The term **snowball sample** is therefore used to describe samples which become larger as each contact suggests more people to contact. They are particularly useful for researching groups whose identity is concealed. For example, say you wish to research the likely impact on crime of a greater liberalisation of laws against illicit ‘hard’ drug use (crack, meths, etc.). You would be unlikely to make contact with drug-using criminals through a probability sample of, say, 1000 from the UK population. First, the incidence of drug-using criminals is relatively small. And, second, they would be unlikely to reveal themselves to you. The two problems you face here are *access* and *trust*. These can be overcome by using a snowball sample. In the example above, you could begin by contacting a person who is publicly known to have had some contact with these criminals. They might be a specialist doctor, a prison visitor, shelter manager or journalist. Your first task would be to meet them to establish your *bona fides* (Latin, meaning ‘good faith’) as an academic researcher and, therefore, trustworthy. Several meetings might be required. You would ask them to refer you to other people who might be able to assist you. They might refer you on to recovering addicts, their relatives, or organisers of self-aid groups. In turn, you would hope that they would find you sufficiently genuine and trustworthy to provide further contacts. You would hope that your research trail would end in clandestine meetings with practising addicts who funded their drug purchases through petty crime or wider involvement in criminal networks. At the end of your research, you will not be able to claim that your findings applied to addicts beyond those you had met: however, you may be able to claim that they apply to all the subjects in your sample.

A *volunteer sample* is one where members of the research population volunteer to take part in your research. You are most likely to seek volunteers where your sample is likely to undergo a period of discomfort, pain or financial cost. You will have received emails from university departments – especially Psychology – asking for volunteers for lab-tests. They may offer a small cash payment to volunteers. For example, you might seek volunteers from colleagues if you wished to pre-test and compare the possible impact of positive and negative political advertising in the UK. But appeals for volunteers need not necessarily be made only to people. You can also ask for volunteers from organisations. For example, to pursue research on political networks, you could write directly to each unitary authority asking them to take part in your research. The great advantage of seeking volunteer samples is that they can be relied upon to co-operate fully. Conversely, because the volunteers

are self-selecting, then they are more likely to be especially interested in the topic and therefore may be less likely to be representative of the population as a whole.

Probability Samples

Probability samples are samples in which every member of the defined population has an equal likelihood of being selected for inclusion. So, in a population of 1,000, each person has a 1/1000 probability (also expressed as 0.1% or $p = 0.001$) of being selected. Probability samples are *statistically reliable*. This means that they are capable of generating data which is representative of the population. In other words, if the average age of members of a probability sample is 30, then you can be confident that the average age of the population as a whole is also 30. This capacity to generate accurate representation is called *generalisability*. Probability samples are large. They are used in all quantitative research where the population is very large.

The reliability of the data obtained from a sample will increase as the sample increases in size towards that of the whole population. But it is the size of the sample which determines its accuracy: the size of the population is less relevant. But doubling the size of the sample will not double the reliability of the information. *Accuracy is proportional to the square root of the sample size*. So, to double the accuracy, the sample size must be increased fourfold – which will greatly increase the cost of the sample survey.

There is, therefore, a trade-off between cost and reliability – and time. This is an example of the so-called *law of diminishing returns* (or *diminishing marginal utility*). It explains why most samples are relatively small, e.g. national opinion polls rarely use more than 1,500 people, whilst even national, life-or-death medical surveys rarely exceed 60,000 i.e. a 0.1% sample.

There are many types of probability samples: *simple random samples* (entirely random); *systematic samples* (every *n*th person); *stratified samples* (e.g. 50:50, men: women, etc.); *multi-stage cluster samples*, and, *probability proportionate to size (PPS) samples*. Essentially, stratified, multi-stage cluster and PPS involve pre-designing the sample to reflect the known characteristics of the population under study, e.g. by gender, age group, ethnicity, social group, residence. They are also used to make the sample more readily contactable and thereby reduce costs. For example, if you were proposing to carry out a face-to-face survey of 10,000 people in England, then you could choose 1,000 electoral wards randomly and then 10 addresses within each (or 200 wards and 50 addresses in each). A further refinement would be to select randomly 125 wards within each of the 8 standard regions and 10 residents in each ward. But some regions have larger populations than others. So you could vary the number of wards *pro rata* so the likelihood of any ward being chosen was more equal. But wards also vary significantly in size (depending on whether they elect one, two, or three members). So the list of wards from which the sample is to be chosen should reproduce multi-member wards twice or three times, and so on.

In this way, you are constructing a sample with a probability proportionate to size. In this case, you would have constructed your sample frame in three layers: region, ward, population. These layers are termed *strata*.

The more strata employed, the smaller the size of each group and the lower the reliability of the data. One statistical survey rule – rarely followed – is that the smallest group should not be lower than 1,000. In practice, a ‘rule-of-thumb’ is adopted in which the group should be at least 50–100 (Hoinville, 1977: 61).¹³ There is a particular problem with very small minorities, e.g. non-whites in UK rural county towns. A group of, say, 100 educated males 26–35 may well include less than five members of minority ethnic communities. But five people are unlikely to be representative of the many minority ethnic communities. In this case, you should seek a higher number of participants from these communities and scale down the data accordingly. Alternatively, if ethnicity is a critical variable, then a very much higher stratified sample should be sought.

Tables have been developed which relate sample size to the degrees of acceptable *sampling error* and *levels of confidence* (CL). Most Politics research adopts *levels of confidence* of 95%. This means that you are confident that, in 95 out of every 100 cases, the characteristic (e.g. party preference) shown by the sample will be shared by the research population. The **sampling error** is the inaccuracy arising from the use of a sample. So, as Table 7.1 shows, if you are willing to accept a sampling error of 5% either way, you can use a sample of 400. But, if you insist on a sample error as low as 1% either way, then you must use a sample of 10,000.

For example, say you have adopted a random sample of 2,500 people whether they support or oppose NATO forces involvement in Afghanistan. 36% say they are supportive. What the table above tells you is that, in this case, you can be confident that, in 95 out of every 100 members of the population, 36% will be supportive \pm 2%. So, at 95 confidence levels, support will lie between 34% and 38%. So, if you

Table 7.1 Sampling errors/sample size of random samples at 95% confidence levels.

Sampling Error %	Sample Size
1	10,000
2	2,500
3	1,100
4	625
5	400
6	277
7	204
8	156
9	123

(abstracted from De Vaus, 2001: 71, Table 5.4)

want to reduce the sampling error by half (to $\pm 1\%$), then you will have to increase your sample to 10,000.

The greatest difficulty in using probability samples is designing the *sample frame*. The *sample frame* is the list of the population from which the sample will be drawn. Say, for example, you want to carry out questionnaire research of residents of a city. Twenty years ago, you could have used the electoral roll or telephone directory. However, to protect the privacy of the public and to encourage more people to register and therefore to vote, the full electoral roll is no longer publicly available. The telephone directory is now much less representative than previously now that a third of 'subscribers' choose to go 'ex-directory'. In any event, the telephone directory lists heads of household and not their partners. Additionally, young people are more likely to use only mobile telephones which are unlisted.

Large, large-scale probability sampling is very expensive. One practicable means available to (fully-funded) researchers to obtain very large samples is to buy into one of the very large (100,000) sample *omnibus* questionnaire surveys undertaken by market researchers acting on behalf of commercial clients – notably superstore grocers and financial services. They use random samples obtained from a sample frame of a national gazetteer of postal addresses. The disadvantage is the relatively low completion rates and the bias arising from the use of prize incentives.

For these and other reasons (especially costs and the lack of life-or-death consequences for the population), Politics researchers rarely use probability samples. Instead, like most private firms, you are most likely to use *quota samples*.

Quota sampling

Despite being the most widely-used, quantitative sampling technique, *quota sampling* is non-probabilistic. However, for calculation purposes, custom allows you to analyse the data using the same statistical techniques as if it had been obtained using simple random sampling. A *quota sample* is a sample of the population which is pre-designed to be representative. So, for example, if you know that 69% of your population of UK electors voted in the last general election, then you will design your sample to have a quota of 69% voters. How do you know whether a person voted? You ask them. Quota samples are usually recruited in town centres by researchers who select passers-by to complete their quota of interviews. It is non-probabilistic because the interviewers select individuals to meet their quota. So each passer-by does not have the same probability of being selected. Furthermore, because the weekday city centre contains higher proportions of some groups of the population and less of others, then the probability of each member of the research population of being selected is unequal. This also explains why you may not have been interviewed by the ever-present pollsters in your city centre. Their quota of people like you had already been filled. Or, alternatively, your type of person did not form part of the sample frame of, say, pensioners. Or, importantly, they may have felt uncomfortable

by your appearance. You will find out when you carry out your own quota sampling that you are likely to select people with whom you believe that you are likely to develop rapport. You will find more detailed guidance on *administering* (carrying out) an on-street survey in Chapter 8.

Designing a quota sample

You will find that the simplest way to design and use a quota sample is to start by setting up a 10 by 10 matrix of 100 cells.

For example, suppose you require a representative sample of the research population in which age, gender and socio-economic group are considered relevant – as potential independent variables – to your research question. You find from published sources that:

1. the breakdown of population between males and females is 50:50%
2. between 18–34, 35–59, 60+ years old, the relative distribution is 20:50:30%
3. as a proxy indicator of socio-economic group, the percentage ratio of owner-occupations to tenant is 60:40%.

Then you can sub-divide the 100-cell matrix into columns for sex and socio-economic group (assuming the owner/tenant distribution is uniform between genders and ages) and into rows for age bands, namely:

In this matrix, the highest, right-hand cell (marked ‘X’) will be a woman, aged 18–34 and living in rented accommodation whereas ‘P’ will be a man, aged 35–49 who is an owner-occupier.

You can then ‘scale-up’ the matrix to provide the optimum sample size on the basis of weighing the advantages of accuracy, reliability and representativeness, against the resource costs. You can, of course, add further sub-divisions, for example, of ethnic origin. But remember that, each time you subdivide the sample further, the sub-groups become smaller and potentially less representative.

So armed with your quota matrix and questionnaire, you can begin your quota sampling. As you recruit each member of the quota sub-group, you ‘tick them off’ from the matrix. Beginning is easy. However, the technique becomes more difficult as the number of vacant cells in the matrix reduces.

Questions for discussion or assignments

1. What do you understand by the essential difference between validity, reliability and accuracy, in terms of data? Why are the distinctions useful?

relationship between 'green practices' and age, class and sex. You have learned that 60% of households separate and sort recyclables from their household refuse. Find the data on age, class and sex from the census data for your city or town. What are the advantages and disadvantages of this method?

FURTHER READING

- Harrison, L. (2001) *Political Research: An Introduction*. London: Routledge. pp. 25–29, 106–12. In the first extract, Harrison discusses and distinguishes between validity and reliability, and provides additional material on types of validity including construct validity and content validity. In the second extract, she examines and discusses alternative sources on existing political data including the 'mass media', party resources, biographical, autobiographical and political memoirs and the internet.
- Neuman, W.L. (2003) *Social Research Methods*. London: Pearson. pp. 178–87, 137–38. The first extract discusses reliability and validity in the separate contexts of quantitative and qualitative research. Table 7.1 on p. 183 summarises the measurement reliability and validity types identified. The second extract provides a very readable discussion of triangulation of which he provides an example of four types in Box 6.1 on p. 138.
- De Vaus, D.A. (2001) *Surveys in Social Research*. London: Routledge. pp. 54–79. This extract begins with a discussion of reliability and validity. The author introduces three means of assessing validity: criterion validity; content validity; and construct validity. He also discusses the special problem of how people may interpret indicators in different ways. The second part of the extract provides good, practical advice on the separate types of probability and non-probability samples.

Notes

- 1 Russell, B. (1971) A Liberal Decalogue In *The Autobiography of Bertrand Russell*, 3. 1944–67, London: George Allen & Unwin pp. 60–01.
- 2 Amis, K. (1997) *The King's English – A Guide to Modern Usage*, cited by Mount, H. (2006) *Amo, Amas, Amat and All That: How to Become a Latin Love.*, London: Short Books. p. 40.
- 3 Huxley, A. (1932) *Brave New World*.
- 4 Crossman, R.H. (1976) *The Diaries of a Cabinet Minister 1 Minister of Housing 1964–66*. London: Hamish Hamilton and Jonathan Cape. p. 590.
- 5 Cole, J. (1996) *As It Seemed to Me: Political Memoirs*. London. Weidenfield & Nicholson. p. 64.
- 6 Jenkins, R. (2001) *Churchill: A Biography*. London: Macmillan.

- 7 Jary, D. and Jary, J. (1995) *Sociology*. Glasgow: HarperCollins. p. 714.
- 8 Lukes, S. (1974) *Power: A Radical View*. London: Macmillan. p. 51.
- 9 Foot, M. (1973) *Aneurin Bevan: A Biography, 2. 1945–60*, London: Davis-Poynter.
- 10 Grigg, J. (1978) *Lloyd-George: The People's Champion*. London. Eyre Methuen.
- 11 Nicolson, N.(1953)*King George V*. London: Routledge. Chapter XXVII. pp 453–469
- 12 Cole, G.D.H. (1938/66) *The Common People*. London: Methuen.
- 13 Hoinville, G., Jowell, R. et al (1977) *Survey Research Practic.*, London: Heinemann. p.61 cited in De Vaus, D.A. (2001) *Surveys in Social Research*. London: Routledge. p. 73