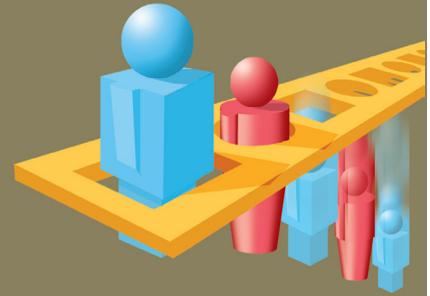


Chapter 4

Exploring the research process



By the end of this chapter you will be able to understand:

- the stages of sociological research
- the main principles of research methodology and design

Research design

This chapter focuses on the idea of sociological research as a *systematic process* — as something carefully *planned* and *organised* by the researcher — because this idea reflects a couple of significant observations.

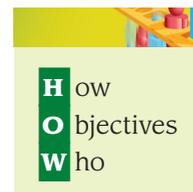
First, sociological research involves a range of factors — from what to study, through collecting data, to analysing and drawing conclusions from such data — that need to be addressed in a particular order. It would, for example, be extremely time-consuming for a researcher to start collecting data without initially deciding on the subject of such activity, mainly because the data collection would lack focus — it would not be clear:

- what data is needed
- how much data should be collected
- whether or not the data is going to be relevant

Therefore the researcher generally starts to collect data after deciding what to study — which gives a kind of logical flow and structure to the research process (and design).

Second, although any research process will have a basic design structure, important choices still have to be made by the researcher — choices that reflect their particular values and beliefs and also produce very different forms of sociological research. A researcher must make decisions about the following:

- Objectives — is the aim to test a theory or simply describe a situation?
- Who to study — will this involve everyone in a particular group or just a selection (sample) of these people?
- How to study them — what method or methods will be used?



These questions are an integral part of doing sociological research because the answers will determine the direction and scope of the research.

The research process

The knowledge produced by sociologists is different from 'common sense' or 'everyday knowledge'. A claim is not just the expression of someone's opinion; rather it represents data that has been systematically collected, analysed and interpreted through a research process. The key difference between sociological and common-sense knowledge, therefore, is that with the former some attempt has been made to verify (or check) its accuracy. This difference means that we need to explore the sociological research process in more detail. We can initially note that it has two main components:

- **Research methods:** These are the various ways sociologists go about the process of collecting data.
- **Methodology:** The systematic collection of data, although a necessary part of the research process, is not the full story. The decision to use certain methods (but not others) or collect certain types of data (but not others) is surrounded by beliefs — and these involve, for example, ideas about the nature of the social world, the ability of different research methods to capture that world and the capacity for different types of data to accurately reflect that world. In other words, sociological research and data collection are always surrounded by methodological questions that must be posed and answered by the researcher.

In terms of the general design structure of sociological research, Oberg (1999) suggests it generally follows an overall design blueprint that involves four distinct, but interconnected, stages:

- 1 Planning:** This is the initial decision-making stage where the researcher decides the basic format of the research (such as what to research and how to research it).
- 2 Information gathering:** This is a data-collection stage where research methods are applied (people are questioned, observed and so forth).
- 3 Information processing:** Once data has been successfully gathered it has to be analysed and interpreted.
- 4 Evaluation:** This normally involves two processes:
 - **Internal analysis** asks such questions as: Was the hypothesis addressed and tested properly? Was the data-collection method appropriate?
 - **External analysis** is the process whereby conclusions are presented to a wider public audience for their analysis and criticism.

In the remainder of the chapter, this broad outline is developed into a more detailed representation of the research process under the following headings:

- research problem
- research review
- research hypothesis or question
- respondents
- data collection
- data analysis
- presentation of completed research

Research problem

This is the initial stage at which the sociologist decides such things as:

- what topic to study (education? health? family life?)
- what aspect to study (having decided on, for example, education, decisions have to be made about what exactly will be researched – ‘attitudes to education’, say)

Research review

Depending on the aims of the researcher, the research review may serve a number of purposes, such as:

- generating ideas about what to study
- replicating (‘exactly repeating’) previous research
- avoiding errors made in previous research
- becoming more familiar with research on a topic

Research hypothesis or question

The hypothesis or question is the focus of the research and will set the basic theme for that research. For example, if a **research hypothesis** is used it will have to be **tested** and this means using research methods suitable for this purpose. If, on the other hand, a **research question** is used this may be sufficient as a means of guiding the research in a particular (possibly descriptive) direction.

Hypothesis

For many sociologists the hypothesis is the starting point and it’s easiest to think of a hypothesis as a question we want to answer or statement we want to explore. A hypothesis has one very important characteristic: we must be able to test it (to discover if it’s true or false) and to do this we have to look at the relationship between two or more things. For example, if we’re interested in researching ‘Why do people steal?’ a simple hypothesis could be: ‘People steal because they are poor’. In this way we can eliminate or confirm ‘reasons for stealing’ because we’ve tested a possible relationship.

Ginn and Arber’s (2002) analysis of how motherhood impacts on the lives of graduate women was based on the following research hypothesis: ‘Is the effect of motherhood on full time employment minimal for graduate women?’



Suggest a possible hypothesis for any of the following research topics:

- reasons for male underachievement in our education system
- the most common type of family structure in our society
- why people believe in God

Research question

Other sociologists begin with a research question they want to answer by collecting evidence. Although not directly tested, the answer to a research question can be supported (or not as the case may be) through research. An example could be: 'What are people's attitudes to theft?' Here the researcher would be gathering evidence about peoples' views on a particular form of behaviour.

Conway's (1997) examination of parental choice in secondary education was based around the question of whether such choice 'helped to strengthen the advantage of the middle classes over the working class'.

Respondents

These are the people who 'respond' to the sociologist's research (by answering questions, agreeing to be observed and so forth). Although it would be ideal to select and study everyone who might have a contribution to make to a piece of research (the '**target**' or '**general** population'), this isn't always possible or desirable. For example, if the target population was 'every secondary school teacher in England and Wales' (around 450,000 people) the size and geographic distribution of such a group would make it difficult to observe or question everyone personally. For this reason sociologists frequently take a sample of a target population.

A **sample** is a relatively small proportion of the people who belong to the target population. In the above case the researcher might choose 100 teachers and, by studying their behaviour, try to say something about all teachers. This involves the issue of **representativeness**, which is arguably more important than the size of a sample because it concerns whether the sample characteristics accurately reflect those of the target population. For example, if 60% of the target population are female then 60% of our sample should also be female.

If the sample group is representative then anything discovered about them can also be applied to the target population — an idea called **generalisation**. A representative sample allows the researcher to generalise the behaviour and opinions of this group to a target population — they can, in other words, make statements about a group they haven't studied (the target population) based on the behaviour of a group they have studied.

To construct a representative sample a researcher needs some way of identifying the people in a target population, for two reasons:

- If a researcher can't identify everyone in their target population their sample may not be representative; it will not accurately reflect the characteristics of the target population.

- To contact people in their sample (to interview them, for example), the researcher needs to know who they are.

To do this a researcher uses a **sampling frame** which, at its most basic, is a list of everyone in a target population; examples include:

- electoral registers — lists of everyone eligible to vote
- school registers — lists of children attending school
- professional membership lists (e.g. the register of all doctors in Britain kept by the British Medical Association)
- company payrolls — lists of employees

However, a researcher will not always have access to such sampling frames. Access may be denied for various reasons:

- **Legal reasons:** A school, for example, may not give unauthorised individuals access to their registers.
- **Confidentiality:** A business organisation may not give access to their payroll records.
- **Secrecy:** Some groups (such as religious organisations, political parties or criminal gangs) may not want to be studied.



Suggest one reason why a researcher may not be able to get access to a sampling frame.

Types of representative sampling

Simple random sampling

Simple random sampling is based on the *probability* that the random selection of names from a sampling frame will produce a sample representative of a target population. An important characteristic of this type of sampling is that for it to be truly random everyone in the target population must have an equal chance of being chosen. A simple random sample, therefore, is similar to a lottery:

- Everyone in the target population is identified on a sampling frame.
- The sample is selected by randomly choosing from the frame until the sample is complete.

A 30% sample of a target population of 100 people, for example, would involve the random selection of 30 people.

Systematic sampling

A variation on simple random sampling — often used when the target population is very large — is to select names *systematically* by taking the sample directly from a sampling frame. For a 25% sample of a target population containing 100 names, a systematic sample would involve choosing every fourth name from the sampling frame.



Random samples are based on chance distributions

Although this technique is not truly random (in this example the fifth name on the list could never be included in the sample so not everyone actually has an equal chance of inclusion), it is 'random enough' for most samples.

Advantages of simple random and systematic sampling

- **Time:** They are relatively quick and easy ways of selecting samples, especially when the target population is very large.
- **Expense:** Both are reasonably inexpensive to create using a sampling frame accurate for the target population.
- **Information:** Other than some way of identifying people in the target population (a name for example), the researcher doesn't require any knowledge about this population.

Disadvantages of simple random and systematic sampling

- **Sampling frame:** This may not be available for a target population.
- **Unrepresentative:** Chance doesn't always produce a representative sample.

? Identify and explain one advantage and one disadvantage of simple random sampling.

Stratified random sampling

A problem with simple random or systematic samples occurs when the target population is **diverse**, consisting of lots of small groups whose views are all important to the research (for example, different age groups within the target population). A **biased sample** can easily occur because with random samples some groups may be **over-represented** and others **under-represented**.

Stratified random sampling avoids these problems, while keeping the idea of sample selection based on chance, by dividing (**stratifying**) the target population into groups whose characteristics are known to the researcher (such as males and females, or different age groups). Each group is then treated as a random sample in its own right.

For example, imagine a target population of 100 people (80 females and 20 males) where the researcher wants a 10% sample. To exactly represent the gender balance of the target population the researcher needs a sample of 8 females and 2 males. This might be achieved by chance, but it's easier to give chance a helping hand by splitting the target population into two groups — the 80 females and the 20 males — and then selecting 10% of each (8 females from the female-only group and 2 males from the male-only group). If we then combine the two samples we get a final random sample fully representative of our target population.

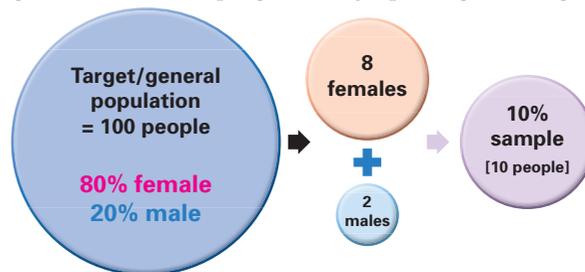


Figure 4.1 A simple worked example of stratified random sampling

Stratified quota sampling

This variation involves the same technique as stratified random sampling, with two main differences:

- **Sampling frame:** Although a sampling frame is always useful, it's not strictly necessary here. It's usually enough just to know the characteristics of respondents — and their associated quotas — in order to construct a sample.
- **Non-random selection:** Using the previous gender example, the selection of 8 females from the female-only group and 2 males from the male-only group is done on an 'opportunity basis'; the researcher, for example, works through the group of 20 males, asking each in turn to be part of the sample. Once 2 males have agreed, the '**quota**' for the male-only sample is complete and no further males can be selected — which means this technique isn't truly random in its sample selection because not everyone in the target population has an equal chance of being selected (the first person asked has a greater chance than the last).

Advantages of stratified sampling

- **Focus:** The researcher can focus the sample on relevant distinctions in the target population (age, gender, class, ethnicity, etc.) and ignore irrelevant factors.
- **Representativeness:** Known differences in the target population will be accurately reflected in the sample; the researcher can, therefore, be sure the sample will be broadly representative.
- **Assets:** In terms of resources, quota samples are usually cheap and quick to construct accurately (which is partly why they're used by political polling organisations).
- **Generalisation:** It's possible to generalise from the sample to the target population, even when the sample is relatively small in relation to the target population.
- **Size:** Stratified samples can be relatively small, since it's possible to make certain that the sample has accurately reflected the target population.

Disadvantages of stratified sampling

- **Information:** Accurate information about the target population isn't always available (which means a representative sample can't be constructed).
- **Uncertainty:** When using a team of researchers to construct a quota sample you can't be certain they have correctly placed everyone in the right quota category. If, for example, your research assistant cannot find '100 men over the age of 65' to fill their quota, there may be a temptation to fill it using men under that age — which would affect the representativeness of the sample.



Suggest one advantage and one disadvantage of stratified sampling.

Non-representative sampling

Although researchers generally find it useful to make their sample representative of the target population, there are times when they might not use a representative sample.

Choice

For some types of research the sociologist might not want to make generalisations about a large group based only on a sample of that group. They might, for example, simply be interested in the behaviour of the group itself, rather than what they may or may not represent. An example of this type of non-representative sampling is a **case study**. The objective here is to study, in great detail, the characteristics of a particular group (or case). Although a case study is technically an example of a **research method** (see below), it illustrates how a non-representative sample works.

The researcher doesn't care whether the group is representative of any other similar groups and the sample in this type of research *is* the target population. This is perfectly acceptable as long as the researcher doesn't try to **generalise** their findings (the group being studied is representative only of itself). An example of a case study is Ward's (2008) research into drug selling among 'rave' dance participants in London.

Necessity

Circumstances sometimes make it impossible to create a representative sample and the researcher may be forced to settle for **opportunity sampling**, a general type of non-representative sampling that has two main subdivisions: 'best' opportunity sampling and snowball sampling.

'Best opportunity' sampling involves deliberately choosing a sample that gives the best possible opportunity to test a hypothesis. If the hypothesis is false for this group, it will probably be false for any other similar groups. Goldthorpe et al. (1968), for example, wanted to test the claim that the working class in Britain was becoming indistinguishable from the middle class. Their best opportunity sample consisted of highly paid car assembly workers in Luton, chosen on the basis that if any working-class group was likely to show lifestyles similar to their middle-class peers it would be this group of 'affluent workers'.

Just as a snowball rolling downhill gets larger as it picks up more snow, a **snowball sample** picks up more and more people to include in the sample over time. A researcher, for example, identifies someone in the target population who is willing to be part of their research. This person then suggests another two or three (perhaps more) people who are also willing to participate. These then suggest further possible participants until the researcher has a usable sample.

Although this technique doesn't produce a representative sample, it may be the best that can be achieved in certain situations. Wallis (1977), for example, used this technique to contact (ex-) members of the Church of Scientology when his request to interview current members was rejected by the Church authorities. This technique is useful when:

- No sampling frame is available.
- The researcher knows little or nothing about the characteristics of their target population.

While this technique is not ideal (and runs a real risk of being unrepresentative), it may be the only way a researcher can construct a sample. Charlton et al.'s (2001) study of young children's mobile telephone ownership and use used an opportunity sample of schoolchildren in the absence of any available sampling frame. For Sappleton et al.'s (2006) analysis of gender segregation in the audio-visual industries, 'Respondents were enlisted through personal referrals, prior contacts and cold calls.'

Advantages of non-representative sampling

- **Availability:** A researcher can construct a sample in situations where it would be impossible to use any other sampling technique.
- **Resources:** It can be a cheap and quick method of sampling (although this depends on both sample size and the speed at which the researcher is able to contact respondents).

Disadvantages of non-representative sampling

- **Representativeness:** It's unlikely the sample will be truly representative and there is no way of checking its representativeness.
- **Resources:** It can be a relatively expensive and time-consuming sampling technique (especially if the sample is large and/or widely dispersed across a large area, and respondents are reluctant or unable to suggest further potential respondents).
- **A self-selected sample** is a distinct possibility because this type of sample effectively 'picks itself' rather than being selected by the researcher.

A good example of a self-selected sample is Hite's (1976) investigation into female sexuality that claimed to uncover a range of interesting sexualities and practices 'representative of the population of America'. Her respondents, however, were self-selected because she chose anyone who responded to her advertisements asking for people to talk openly about their sexual behaviour. In this respect, therefore, the responses of a small number of unrepresentative people who wanted the world to know about their sexual behaviour came to represent 'general public behaviour in America'.

This research also illustrates a further potential problem (not exclusively related to opportunity sampling): that of **statistically inadequate samples**. A sample that is too small to accurately represent a target population (such as a couple of people in the street who are asked what they think about the education system) will be inadequate for research purposes. As a general rule, the larger the sample as a proportion of a target population the greater the probability it will be statistically adequate. This may improve the chances of your sample being representative, but a large sample is no guarantee of a representative sample.



Suggest one reason why a self-selected sample may be unrepresentative.

Data collection

Once the researcher is ready to start collecting data, a range of **methodological considerations** arise. 'Methodology', Browne (2000) notes, 'is concerned with the choice of research methods for collecting, analysing and interpreting data.'

Pilot studies

It is sometimes useful to conduct a pilot study. Van Teijlingen and Hundley (2001) refer to this as a 'mini version of a full-scale study'; it is, in other words, a small-scale exploratory piece of research, carried out prior to the actual research, designed to test and tweak the research design. The purposes may include:

- convincing funding bodies of the research team's competence and demonstrating that a study is both possible and worth funding
- identifying any problems that might occur in a full study (such as establishing access to respondents)
- testing the research methods (e.g. by trialling questions to eliminate possible sources of bias, checking for misunderstood questions and collecting preliminary data)
- establishing whether the sampling frame, technique and size are appropriate
- determining the level and extent of resources (such as finance and staff training) needed for a study

Concept operationalisation

A researcher always needs to define, test or measure (or '**operationalise**') the various elements involved in the hypothesis or research question. In the poverty hypothesis we used previously, a researcher would need to be clear about how poverty is defined — otherwise it would be impossible to test the relationship between poverty and stealing. We can operationalise the concept of poverty by using one or more **indicators**. In this instance income would be a useful indicator: we could specify that, for the purpose of the research, anyone with an annual income of less than £10,000 would be considered poor.

Most research involves concepts that can't be easily defined or measured because they don't physically exist. We therefore need to use indicators of their existence (such as income in the case of poverty) which can be measured.



Suggest a possible indicator we could use to measure each of the following:

- 1 educational achievement
- 2 religious beliefs
- 3 the extent of crime in the UK

Choosing a research method

Although all research methods have their strengths, weaknesses, uses and limitations, we can identify two methodological concepts that also influence the researcher's choice of method in terms of data collection: reliability and validity.

Reliability

Reliability refers to the *consistency* of the data collected in any research process. We can check the accuracy of data by repeating the research to see if we get the same, or very similar, results (we may have to allow for possible changes over time). In general, data is reliable if similar results are gained by different researchers (or the same researcher at different times), for example when asking the same questions of similar people.

A simple example might be a researcher cross-checking the reliability of a response within a questionnaire by asking the same question in a different way:

- How old are you?
- When were you born?

If they get two different answers, the data is unreliable — and if data is unreliable, any conclusions we draw from it are going to be of limited use (if not useless).

We need, therefore, to be aware that data reliability is affected by such things as:

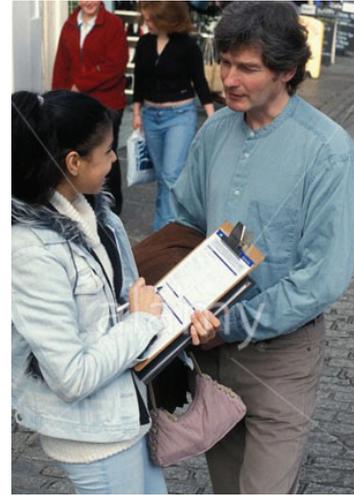
- **Researcher bias:** Are there opportunities for the researcher (consciously or unconsciously) to influence or distort the data collection process?
- **Definitions:** When comparing data over time, is the same definition used for whatever is being measured?
- **Replicability:** If another sociologist attempted to exactly repeat a piece of research, would similar results be achieved?

Validity

Validity refers to the extent to which data actually measures or describes what it claims to be measuring or describing. For example, if we were interested in the extent of crime in our society, we could use official crime statistics (a **secondary data** source published by the government — see below). We would need to be aware, however, that the validity of these statistics may be limited because they only record reported crimes — and people may not report crime for many reasons.

Validity is a useful concept because it reminds us to consider the accuracy of different data types (primary, secondary, qualitative and quantitative). While some forms of data (such as official statistics) may be *reliable*, their *validity* may be questionable for two reasons:

- **Representativeness:** They may not apply to everyone in a particular group. In Britain, for example, 'unemployment statistics' represent only those registered for unemployment benefit, not all those without a job. To achieve representativeness, the information we collect must be sufficiently comprehensive to accurately embody



Questionnaires are a popular research method

Angela Hampton Picture Library/Alamy

whatever the research claims to represent. Official crime statistics, for example, are unrepresentative of all crimes committed in our society so anything a piece of research says about crime on the basis of this data source needs to be *qualified* by saying that some types of criminal behaviour may not be fully represented in the statistics.

- **Depth:** They may lack the depth and detail required to accurately represent the views of a particular individual or group (and so to measure what they aim — or claim — to measure).

When evaluating the validity of a particular research method, data type or data source we always need to question whether they actually measure what they claim to be measuring. If they do (however limited their scope may be), they are valid. If they don't then validity is likely to be low.

Primary and secondary data

The actual mechanics of data collection is an important element in research design. There are two basic types of sociological data: primary and secondary.

Primary data

Primary data involves information collected personally by a researcher (or their research team). Sociologists use a range of research methods (such as questionnaires, interviews and observational studies) as sources of primary data.

Advantages

- **Data control:** The researcher is responsible for collecting data and this gives complete control over how it is collected, by whom and for what purpose (there is no need to rely on other people to collect the data accurately).
- **Reliability, validity and representativeness:** The ability to exercise a measure of control over how data is collected doesn't guarantee its reliability, validity or representativeness — a badly designed piece of research can be unreliable, invalid and unrepresentative. However, it's much easier for the researcher to deal with these issues when designing and carrying out the research themselves.

Disadvantages

- **Time:** It is time-consuming to design, construct and carry out research which collects primary data. If the group being studied is large and if individual interviewing is involved, the collection will take considerable time and resources.
- **Expense:** The cost of a researcher's time (among other things) may be a factor in the design of the research.
- **Access:** Having designed a piece of primary research, you need access to the people you want to study — and this may involve both material problems (your plan to interview the 100 richest people in the UK will come to nothing if they refuse to be interviewed) and immaterial problems (with historical research, for example, potential primary sources may be dead).

Secondary data

Secondary data involves information *not* personally collected by the researcher. Sources include newspapers, books, personal documents (such as letters and diaries), official documents (such as government reports and statistics) and even the research of other sociologists.

Advantages

- **Resources:** It may be unnecessary or impractical to create some forms of data (using primary methods) when it already exists. The British government, for example, collects and distributes a huge amount of statistical data each year. For the price of a book, a visit to library or an internet search, the researcher has immediate access to data — such as national crime or marriage statistics — that would cost an enormous amount of money, time and effort to collect personally. In addition there may be situations where secondary data is the only available resource (for example, the primary data sources may be dead, as in some historical research or research into suicide).
- **Comparisons:** Secondary data can be useful for historical and/or comparative purposes. Aries (1962), for example, used historical data (such as paintings and documents) to support his argument that childhood was a relatively recent invention.
- **Reliability:** Some (but not all) forms of secondary data can be highly reliable. Official statistics (those produced by the UK government, for example) are generally considered reliable because:
 - They are collected regularly and consistently in the same way from the same sources. Education statistics, for example, are regularly collated by the Office for National Statistics from a variety of government sources and surveys.
 - They generally measure, within reason, the same things each time they're collected; comparisons made between different years are, therefore, comparing 'like with like'. Official statistics measuring educational achievement at GCSE, for example, consistently use the same definition of achievement (grades A*–C).
- **Validity:** Some kinds of secondary source (biographies and personal documents such as diaries, for example) provide highly valid data because they give detailed, well-rounded insights into people's thoughts and behaviour — something that may be especially useful in circumstances where individuals are dead or have written contemporary accounts of long-past historical events.
- **Representativeness:** Where data is produced by the British government there is normally a high level of representativeness because the resources (e.g. funding, number of researchers) committed to collecting data mean that large samples can be accurately constructed. The census (a questionnaire distributed to every household in the UK every 10 years), for example, draws on a highly representative sample of the UK population (its reliability is also high because it must, by law, be completed by every household).

Disadvantages

- **Data control:** This may be difficult because secondary data is not always produced with the needs of sociologists in mind. For example, the way governments measure social class may differ from sociological ways of measuring class.
- **Reliability:** The range and variety of secondary data make generalisations about reliability difficult — some sources, such as official statistics, may be reliable whereas others, such as a diary or newspaper article, may be unreliable. To assess the reliability of secondary data we always need to ask questions about:
 - **who** produced it
 - **how** it was produced
 - the **objectives** behind its production
- **Validity and representativeness:** Does the data represent the viewpoint of just one individual or a much wider range of views? A newspaper article, for example, can express the personal and unrepresentative view of a single journalist. Similarly, historical documents may reflect the views of particular social classes (because it was generally the upper classes in Britain who, until quite recently, recorded their particular view of the world). Conversely, the only surviving record of something may provide a valid insight into that event, but without supporting evidence (a question of *reliability*) we can't be certain of either its validity or representativeness. As with reliability, the range and scope of secondary data makes it difficult to generalise about its validity — some forms (such as eyewitness descriptions of an event) may have greater validity than official statistics that simply focus on quantifying something.



Identify and explain two reasons why a sociologist might collect secondary rather than primary data.

Quantitative and qualitative data and methods

Primary and secondary data can be subdivided into two types: quantitative and qualitative.

Quantitative data

Quantitative data is data expressed *statistically* or *numerically*. We could, for example, count the number of people in Britain who live in poverty or vote Conservative; such data is usually expressed in one of three ways:

- **Number:** For example, the total number of people who live in poverty could be stated.
- **Percentage** (the number of people per 100 in a population): For example, 30% of British voters regularly vote Conservative.
- **Rate** (defined here as the number of people per 1000 in a population): For example, if the birth rate in a particular country is 1, this means that one baby is born each year for every 1000 people in the population.

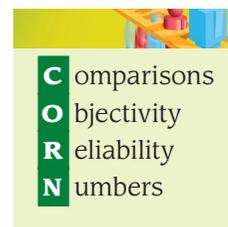
Although raw numbers can be useful (e.g. knowing the number of children who will be starting school in 10 years' time allows the government to plan for the number of schoolteachers, school places and so on), data is often expressed as a rate or percentage because it allows **comparisons** between and within groups and societies. For example, when comparing unemployment between Britain and America, expressing unemployment as a raw number wouldn't tell us very much, since the population of America is roughly five times larger than that of Britain. Expressing unemployment as a percentage or rate allows us to compare 'like with like' by taking into account the fact that one society has more people than the other (so we might expect the larger society to, numerically, have more people unemployed, even though unemployment rates might be broadly similar).



Briefly explain, using an example to illustrate your answer, what is meant by quantitative data.

Strengths

- **Comparisons:** Where the researcher wants to test a hypothesis, quantitative data allows for relatively simple 'True/False' distinctions to be made on the basis of statistical comparisons (the hypothesis will be either true or false). Similarly, cross-cultural comparisons (crime rates in different countries, for example) are made possible through the use of quantitative data. **Longitudinal studies** (where the same group may be questioned at different times to track changes in their behaviour) can also exploit this comparative feature of quantitative data. As Kruger (2003) argues, quantitative methods and data 'allow us to summarize vast sources of information and make comparisons across categories and over time'.
- **Objectivity:** Where the researcher has no direct, necessary and personal involvement with the generation of data, it is less likely (but not, of course, impossible) that personal biases will intrude into the process. This gives quantitative data collection what Kealey and Protheroe (1996) term the ability to 'eliminate or minimize subjective judgments'.
- **Reliability:** Quantitative data tends to be more reliable than qualitative data because it's easier to **replicate** (repeat) the data collection process. This is because **standardised questions** (questions that don't change) can be asked of different groups (or the same group at different times). Matveev (2002) notes that the ability to control the conditions under which data is collected (such as using standardised questionnaires) makes quantitative data more reliable.
- **Numbers:** The ability to express relationships statistically can be advantageous if the researcher doesn't particularly need or want to explore the reasons for people's behaviour. For example, if you simply need to know the number of murders committed each year or the number of students absent from the classroom in any given month then quantitative data satisfies this purpose more than adequately.



Weaknesses

- **Control:** Although the ability to quantify behaviour can be a significant plus factor for researchers, this situation is frequently achieved by placing the respondent in an artificial social setting. In other words, realism is sacrificed for control. In their everyday lives, for example, people rarely, if ever, encounter situations where they're asked to respond to a list of questions from a stranger; similarly, people are rarely placed in situations where their behaviour is observed (secretly or otherwise). The main question here, therefore, is the extent to which a researcher can capture people's 'normal behaviour' or 'real opinions' when they place respondents in a situation that is neither normal nor real.
- **Validity:** Quantitative methods only capture a relatively narrow range of data about people's behaviour — what Day (1998) calls the 'Who What When Where'. Although these can be interesting and informative questions, quantitative methods are relatively poor at capturing the *reasons* for such behaviour — something related to the problem of depth (see next point).
- **Depth:** The more detailed the data about behaviour, the more difficult it is to meaningfully quantify. One criticism of quantitative methods, therefore, is that they focus on relatively superficial aspects of behaviour while failing to address the complexities involved in even very simple forms of behaviour.
- **Meaning:** A lack of depth and detail leads to a further limitation — one noted by Kruger (2003) when he suggests it is 'difficult to get the real meaning of an issue by looking at numbers'. Although quantitative methods can ask people why they commit crimes or why they truant from school, their lack of depth and detail means they can't easily express the 'richness of meaning' behind their behaviour.
- **Prejudgements:** McCullough (1988) argues a significant methodological limitation is the fact that 'issues are only measured if they are known prior to the beginning of the survey (and, therefore, have been incorporated into the questionnaire)'. In other words, in order to quantify behaviour the researcher must decide, in advance of their research, what is and what is not significant in relation to the behaviour being studied. There is little or no opportunity to develop the research outside of the original parameters decided by the researcher.
- **Reliability:** Although, as a general principle, quantitative data is usually considered both 'highly reliable' and 'more reliable' than qualitative data, this is not necessarily the case (reliability is not an automatic quality of any one particular research method). As Harvey (2002) argues, 'Many apparently quantitative data depend critically on the way in which they were collected, who collected them, where they were collected, when they were collected and from whom they were collected.'

We can also note the risk of **quantophrenia** in the choice of quantitative data. This term refers to what Sorokin (1956), partly tongue-in-cheek, terms a 'psychological compulsion to grasp for the numeric' — a 'condition' that leads to the use of quantification for its

own sake, regardless of whether or not it tells us anything useful or interesting about the behaviour being quantified. As Eberstadt (2006) puts it, the ‘victims’ of this condition ‘obsess over numbers as descriptors, no matter how dubious their basis or questionable their provenance’.



Assess the strengths and weaknesses of quantitative data as a way of understanding people’s behaviour.

Qualitative data

With qualitative data the aim is to capture the *quality* of people’s behaviour (what they feel, for example, about a sociologist asking them about crimes they may have committed). Such data says something about how people *experience* the social world and can be used to understand the **meanings** they give to behaviour. Boyle (1977), for example, studied the behaviour of a juvenile gang from the viewpoint of its members while Goffman (1961) tried to understand the experiences of patients in an American mental institution. Both were trying to capture and express the quality of people’s behaviour — how they feel about and react to different situations.



Briefly explain what is meant by qualitative data.

Strengths

- **Imposition:** If your research objective is to understand the meaning of people’s behaviour, you must allow people the scope to talk freely. Qualitative data allows for this because the researcher doesn’t *impose* their interpretation on a situation (by asking direct, quantifiable questions for example).
- **Depth:** Qualitative data provides greater depth and detail about behaviour since, as Day (1998) suggests, it is concerned with discovering ‘the Why’ — the complex reasons for such behaviour.
- **Prejudgements:** Qualitative methods avoid the problem of the researcher prejudging what is and what is not significant data prior to starting their research. Where the research objective is to describe or draw out people’s opinions and the reasons for their behaviour, the respondent, rather than the researcher, is effectively the driving force — they lead and the researcher follows. The researcher can’t tightly control the research process because respondents may start to talk about things they see as significant and take the research in directions the researcher had not originally thought about when the research was planned.
- **Rapport:** Many qualitative methods (such as observational methods where a researcher may live closely with those being researched) demand the researcher establish a close relationship with the people being researched (which doesn’t mean they have to like them, only that they come to understand their situation). This has a couple of advantages:

- It means that everyone involved in the research is free to suggest new ideas and directions — the role of the respondent isn't limited to answering questions.
- Where the atmosphere is more relaxed and less clinical the researcher is more likely to get respondents to open up about their thoughts and feelings — which may improve research validity.
- **Validity:** Qualitative methods don't have a monopoly on validity (any poorly designed research can lack validity regardless of the methods used) but when we're dealing with the complexities of human behaviour it is much more likely that research methods that explore this complexity will score highly in terms of their validity — that they will measure what they claim to measure.
- **Naturalism:** Qualitative methods allow researchers greater freedom to study people in their everyday or 'normal' settings and this gives a greater chance of either observing or revealing what people 'really believe' or how they 'really behave'. This is a bit like the difference between asking people to remember and describe something they did last Monday with having been able to follow and observe them to discover exactly what they did. As Matveev (2002) suggests, qualitative methods allow the researcher to gain a 'more realistic feel of the world that cannot be experienced in the numerical data and statistical analysis used in quantitative research'.

Weaknesses

- **Comparisons:** Qualitative research is difficult to compare across time and space (because you're not comparing 'like with like'); it also tends to be structured in ways that make the research difficult to replicate.
- **Reliability:** Qualitative research methods generally produce data with lower levels of reliability:
 - Cassell and Symon (1994) argue that, where research evolves to take account of the input of different respondents, the original research objectives may change, making it difficult for subsequent researchers to replicate.
 - Where qualitative methods produce vast amounts of data relating to a wide range of issues, the researcher, as the initial interpreter, has a pivotal role to play in choosing which data is important and which is to be discarded. Different researchers could potentially arrive at different conclusions based on the data they choose to use, which means reliability will necessarily suffer.
 - Levy (2006), however, argues reliability evidenced through the ability to replicate research is not a useful test for qualitative research methods. She suggests the concept of **trustworthiness** might be a more useful measure of the internal reliability of qualitative methods: 'it is up to the researcher to provide evidence of reliability by carefully documenting the data collection and analysis process... can we trust that the results are a 'true' reflection of our subject?'
- **Abilities:** Qualitative methods require different skills from the researcher. The qualitative researcher may want to establish a close **rapport** with their respondents

while for their quantitative counterparts this is neither necessary nor desirable (since it would lower the **objectivity** of the research). In observational research the researcher needs to be able to convincingly and consistently 'play a role' within the group they are studying — and this requires a very different set of skills from those needed to deliver a questionnaire or structured interview.

- **Generalisations:** Qualitative research generally focuses on the intensive study of relatively small groups, so opportunities to generalise research findings may be limited.



Assess the strengths and weaknesses of qualitative data as a way of understanding people's behaviour.

Quantitative and qualitative data

There are occasions when a researcher may want to combine the two types of data — for example, collecting quantitative data about educational achievement or the number of people who visit their doctor each year alongside qualitative data that seeks to explore the satisfaction levels of pupils or patients.

Alternatively, quantitative data is sometimes collected before starting qualitative research, as an **enabling study** to establish whether or not there is anything to qualitatively investigate. A researcher looking at reasons for school truancy in their locality, for example, may first carry out a quantitative analysis to discover whether or not pupils are actually absent from the classroom.

Research ethics

Ethics refer to the morality of doing something. In sociological research there are ethical issues concerning what a researcher should or should not do before, during and after their research. It is necessary to consider both legal and safety issues (for the researcher, those being researched, and any subsequent researchers). These issues are discussed below, as is the code of practice for research.

Legal considerations

- **Power:** It would be unethical to bully or blackmail people (emotionally or physically) into participating in research. In addition, especially when researching people who are relatively powerless, relationships need to be based on trust and personal integrity on the part of the researcher. For example, if the researcher promises anonymity when researching people involved in criminal or deviant activities, disclosing those identities to the authorities or the media would be unethical.
- **Illegal behaviour:** In Britain the collection, storage and retrieval of data are governed by acts and laws, such as the Data Protection Act, the Human Rights Act, copyright laws and the laws of libel. In addition, if research involves criminal or deviant activities, the researcher may have to consider the ethical question of

participation in such behaviour or their responsibilities to both the perpetrators and their possible victims.

- **Research consequences:** Data can be used in many different ways (and not always in the way the researcher intended — media reports may seriously distort the research, for example) and participants should be aware of any possible consequences of their participation. In addition, if respondents feel they have been mistreated (physically or verbally, for example) or misled, this may have legal consequences for the researcher and create problems for any subsequent research.
- **Consent:** The researcher should always (at least try to) gain the informed consent of those being researched.

Safety considerations

- **Rights and well-being:** Care should always be taken to ensure the physical and psychological safety of both the researcher and the respondents. The researcher needs to safeguard the interests, rights and general well-being of respondents. For example, the researcher should respect respondent privacy and minimise any anxiety or distress caused by the research.
- **Involvement:** Some types of research involve methods (such as covert participant observation — see below) that create high levels of involvement with respondents. Where close personal and/or intimate relationships arise, the researcher should ensure that, once the research is completed and contact diminishes, distress is not caused to potentially vulnerable people. For example, if your research involves visiting the elderly on a regular basis, it would be unethical to simply stop your visits once the research is completed.

Code of practice

The conduct of sociological research is surrounded by a range of issues that relate to what the British Sociological Association (2004) terms **professional integrity**. The behaviour of researchers is bound by a code of ethical practice that is an integral part of the professional research role. Pimble (2002) suggests there are three main ethical questions — Is it true? Is it fair? Is it wise?

Is it true?

This relates to both the research process (how it is generally conducted) and, most importantly, the relationship between research findings and their implications. At its most extreme, perhaps, unethical behaviour in this category would involve the researcher deliberately *fabricating* ('making up') data or falsifying their results.

Is it fair?

Unethical behaviour here relates to the different social relationships created during the course of a research study.

First, there is the relationship between the researcher and **other researchers**. This involves such issues as the **ownership** of a completed piece of research (who, for

example, can ethically claim to be the author?). Authorship can be significant when career advancement (in a university, for example) depends on the ability to publish original research. One kind of unfair behaviour is **plagiarism** — passing off the work of others as your own.

Second, there is the relationship between researcher and **respondents**. This relationship is usually ethically straightforward, in terms of legality and safety for example. There are 'moral grey areas' that sometimes arise during **covert** forms of research where the respondent is not aware they're being studied:

- Wallis (1977) wanted to study the Church of Scientology but the Church leaders refused to cooperate with his request for access to existing members, so he contacted ex-members instead and based his research around their opinions and experiences.
- Rosenhan (1973), who wanted to test if doctors could accurately diagnose schizophrenia, sent students displaying fake symptoms into hospitals to test his hypothesis that they could not — and the experiment discovered that doctors were unable to expose the fake patients.

- In Milgram's (1974) study of the effects of authority on people's behaviour, he investigated whether respondents were willing to inflict extreme levels of pain on innocent strangers when instructed to do so by an authority figure. Respondents were convinced they were administering electric shocks to 'learners' whenever the latter gave an incorrect answer to a question (in fact no shocks were administered and the 'victims' were under instructions to pretend they were being hurt). Such research raises important ethical issues about:

- tricking people into cooperating with research
- causing distress to respondents (some argued and protested about the instructions they were given and some broke down in the face of the pain they believed they were inflicting)
- experimenting on people who do not know they are being studied

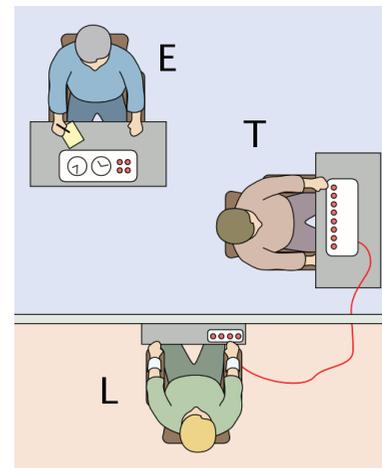


Illustration of the setup of a Milgram experiment. The experimenter (E) convinces the subject ('Teacher' T) to give what he believes are painful electric shocks to another subject, who is actually an actor ('Learner' L)

Wikipedia



Suggest and explain two reasons to justify each of the following:

- studying groups who don't want to be studied
- deceiving people in the course of research
- secretly researching people

Is it wise?

Here we are concerned with ethical questions about the relationship between 'the research agenda and the broader social and physical world, present and future':

- Can the research itself be morally justified?
- Would some other type of research have greater moral justification?

As Pimple puts it:

Will the research improve the human condition, or damage it? Will it lead to a better world, or a worse one? Or less grandly, which of the many possible lines of research would we be better off pursuing? We have finite time and money for pursuing research, and the wisdom of research programs is a valid question in research ethics.

Data analysis

Foucault (1970) argues data ‘can never speak for itself’; information has to be **analysed** and **interpreted**. Analysis involves bringing together and categorising related ideas, for example. Interpretation involves asking what the data and the overall research mean.

Analysis and interpretation take place on three levels:

- **Private or internal:** This level involves thinking about methodological concepts such as data reliability and validity (discussed earlier), to ensure data is logical and consistent.
- **Practical:** This level relates to the purpose of data collection: whether testing a hypothesis or answering a research question, the researcher must do something with the data. This might involve making **correlations** — noting how two or more things seem to occur at similar times. (Wilkinson and Pickett (2009), for example, make extensive correlations between social inequality and crime, such that the more unequal a society is, the higher its relative level of crime.) It might involve suggesting **causation** (that one thing occurs because something else has made it happen).
- **Public or external:** Whenever research is presented ‘to the world’ it represents the outcome of a process of social construction; at its most basic this means that not everything the researcher saw, heard or recorded is presented for public consumption, partly because it would make for very lengthy and unwieldy reports and partly because a great deal of collected data may be irrelevant to the overall research objectives. Decisions always have to be made about what to include (present to the world) and exclude.



Suggest and explain two reasons why data can never ‘speak for itself’.

Presentation of completed research

Glaser and Strauss (1967) suggest the presentation of research generally involves four related elements:

- **Patterns:** Is it possible to discover patterns in the data?

- **Analysis:** This involves analysing both current and related research to discover common themes and trends in the data.
- **Reflection:** This involves a questioning approach to understanding and interpreting the research. (Does it support or disprove the hypothesis? Does it answer research questions or raise further questions?)
- **Theory:** Does the research suggest ways the data can be linked to create an overall theory?



Patterns
Analysis
Reflection
Theory

Finally, once the data has been analysed and interpreted it can be presented in terms of:

- **findings**
- **conclusions** about, for example, the hypothesis (has it been disproven?)
- **limitations** — this might include discussion of various research problems that impacted on the study
- **suggestions** for further research, improvements to the research design and the like

Explain and evaluate the use of qualitative data techniques when researching the identities of young people who are seen as school failures. (52 marks)