# Social Research and Data Analysis: Demystifying Basic Concepts

## Introduction

This book is about the *analysis* of certain kinds of *data*, that is, only *quantitative* data. We need to begin by discussing the three concepts that make up the main title of this book. The core concept is ‘data’. On the surface, it appears to be a simple and unproblematic idea. However, lurking behind it are complex and controversial philosophical and methodological issues that need to be considered. This concept is qualified by the adjective ‘quantitative’, thus indicating that only one of the two main types of data in the social sciences will be discussed. Just what constitutes ‘quantitative’ data will be clarified. The purpose of the book is to discuss methods of ‘analysis’ used in the social sciences, methods by which research questions can be answered. The variety of methods that are available for basic analysis will be reviewed.

This chapter deals with three fundamental questions:

- What is the purpose of social research?
- What are data?
- What is data analysis?

The chapter begins with a discussion of the role of research objectives, research questions and hypotheses in achieving the purpose of research. This is followed by a consideration of the relationship between social reality and the data we collect, and of the types and forms of these data. Included is a discussion of ‘concepts’ and ‘variables’, the ways in which concepts can be measured, and the four levels of measurement. The chapter concludes with a review of the four main types of data analysis that are covered in subsequent chapters.[1] Let us start with the first question.

## What is the Purpose of Social Research?

The aim of all scientific disciplines is to advance knowledge in their field, to provide new or better understanding of certain phenomena, to solve intellectual puzzles and/or to solve practical problems. Therefore, the critical issues for any discipline are the following:

- What constitutes scientific knowledge?
- How does scientific knowledge differ from other forms of knowledge?
- How do we judge the status of this knowledge? With what criteria?
- How do we produce new knowledge or improve existing knowledge?

In order to solve both intellectual and practical puzzles, researchers have to answer questions about *what* is going on, *why* it is happening and, perhaps, *how* it could be different. Therefore, to solve puzzles it is necessary to pose and answer questions.

## The Research Problem

A social research project needs to address a *research problem*. In order to do this, research questions have to be stated and research objectives defined; together they turn a research problem into something that can be investigated. Throughout this book the following research problem will be addressed: *the apparent lack of concern about environmental issues among many people and the unwillingness of many to act responsibly with regard to these issues*. This is a very broad problem. In order to make it researchable, it is necessary to formulate a few research questions that can be investigated. These questions will be elaborated in Chapter 2. In the meantime, to illustrate the present discussion, let us examine two of them here:

- To what extent is environmentally responsible behaviour practised?
- Why are there variations in the levels of environmentally responsible behaviour?

Each research question entails the pursuit of a particular research objective.

### *Research Objectives*

One way to approach a research problem is through a set of *research objectives*. Social research can pursue many objectives. It can explore, describe, understand, explain, predict, change, evaluate or assess aspects of social phenomena.

- To *explore* is to attempt to develop an initial rough description or, possibly, an understanding of some social phenomenon.
- To *describe* is to provide a detailed account or the precise measurement and reporting of the characteristics of some population, group or phenomenon, including establishing regularities.

- To *explain* is to establish the elements, factors or mechanisms that are responsible for producing the state of or regularities in a social phenomenon.
- To *understand* is to establish reasons for particular social action, the occurrence of an event or the course of a social episode, these reasons being derived from the ones given by social actors.
- To *predict* is to use some established understanding or explanation of a phenomenon to postulate certain outcomes under particular conditions.
- To *change* is to intervene in a social situation by manipulating some aspects of it, or by assisting the participants to do so, preferably on the basis of established understanding or explanation.
- To *evaluate* is to monitor social intervention programmes to assess whether they have achieved their desired outcomes, and to assist with problem solving and policy-making.
- To *assess social impacts* is to identify the likely social and cultural consequences of planned projects, technological change or policy actions on social structures, social processes and/or people.

The first five objectives are characteristic of *basic research*, while the last three are likely to be associated with *applied research*. Both types of social research deal with problems: basic research with theoretical problems, and applied research with social or practical problems. Basic research is concerned with advancing fundamental knowledge about the social world, in particular with description and the development and testing of theories. Applied research is concerned with practical outcomes, with trying to solve some practical problem, with helping practitioners accomplish tasks, and with the development and implementation of policy. Frequently, the results of applied research are required immediately, while basic research usually has a longer time frame.

A research project may pursue just one of these objectives or perhaps a combination of them. In the latter case, the objectives are likely to follow a sequence. For example, the four research objectives of *exploration, description*, *explanation* and *prediction* can occur as a sequence in terms of both the stages and the increasing complexity of research. *Exploration* may be necessary to provide clues about the patterns that need to be described in a particular phenomenon. *Exploration* usually precedes *description*, and *description* is necessary before *explanation* or *prediction* can be attempted. Whether all four objectives are pursued in a particular research project will depend on the nature of the research problem, the circumstances and the state of knowledge in the field.

The core of all social research is the sequence that begins with the *description* of characteristics and patterns in social phenomena and is followed by an *explanation* of why they occur. Descriptions of what is happening lead to questions or puzzles about why it is happening, and this calls for an explanation or some kind of understanding. The two research questions stated in the previous subsection illustrate these two research objectives. To be able to explain why people differ in their level of environmentally responsible behaviour, we need to first describe the range in levels of this behaviour. The first question is concerned with description and the second with explanation.

## *Research Questions*

To pursue such objectives, social researchers need to pose *research questions*. Research questions define the nature and scope of a research project. They:

- focus the researcher's attention on certain puzzles or issues;
- influence the scope and depth of the research;
- point towards certain research strategies and methods of data collection and analysis;
- set expectations for outcomes.

Research questions are of three main types: 'what' questions, 'why' questions and 'how' questions:

- 'What' questions seek descriptive answers.
- 'Why' questions seek understanding or explanation.
- 'How' questions seek appropriate interventions to bring about change.

All research questions can and perhaps should be stated as one of these three types. To do so helps to make the intentions of the research clear. It is possible to formulate questions using different words, such as, 'who', 'when', 'where', 'which', 'how many' or 'how much'. While questions that begin with such words may appear to have different intentions, they are all versions of a 'what' question: 'What individuals …', 'At what time …', 'At what place …', 'In what situations …', 'In what proportion …' and 'To what extent …'. Similarly, some questions that begin with 'what' are actually 'why' questions. For example, 'What makes people behave this way?' seeks an explanation rather than description. It needs to be reworded as: 'Why do people behave this way?'.

Each research objective requires the use of a particular type of research question or, in a few cases, two types of questions. Most research objectives require 'what' questions: *exploration, description, prediction, evaluation* and *impact assessment*. It is only the objectives of *understanding* and *explanation,* and possibly *evaluation* and *impact assessment*, that require 'why' questions. 'How' questions are only used with the objective of *change* (see Table 1.1). Returning to our two research questions, the first is a 'what' question that seeks a descriptive answer, and the second is a 'why' question that asks for an explanation.

## *The Role of Hypotheses*

It is a commonly held view that research should be directed towards testing *hypotheses*. While some types of social research involve the use of hypotheses, in a great deal of it hypotheses are either unnecessary or inappropriate. Clearly stated, hypotheses can be extremely useful in helping to find answers to 'why' questions. In fact, it is difficult to answer a 'why' question without having some ideas about where to look for the answer. Hence, hypotheses provide possible answers to ‘why’ questions.

Table 1.1 *Research questions and objectives*

| | Research questions | | |
|---|---|---|---|
| Research objectives | What | Why | How |
| Exploration | ✓ | | |
| Description | ✓ | | |
| Explanation | | ✓ | |
| Understanding | | ✓ | |
| Prediction | ✓ | | |
| Intervention | | | ✓ |
| Evaluation | ✓ | ✓ | |
| Assess impacts | ✓ | ✓ | |

In some types of research, hypotheses are developed at the outset to give this direction; in other types of research, the hypotheses may evolve as the research proceeds. When research starts out with one or more hypotheses, they should ideally be derived from a theory of some kind, preferably expressed in the form of a set of propositions. Hypotheses that are plucked out of thin air, or are just based on hunches, usually make limited contributions to the development of knowledge because they are unlikely to connect with the existing state of knowledge.

Hypotheses are normally not required to answer ‘what’ questions. Because ‘what’ questions seek descriptions, they can be answered in a relatively straightforward way by collecting relevant data. For example, a question such as ‘What is the extent of recycling behaviour among university students?’ requires specification of what behaviour will be included under ‘recycling’ and how it will be measured. While previous research and even theory may help us decide what behaviour is relevant to this concept, there is no need to hypothesize about the extent of this behaviour in advance of the research being undertaken. The data that are collected will answer the question. On the other hand, to answer the question ‘Why are some students regular recyclers?’ it would be helpful to have a possible answer to test, that is, a hypothesis.

This *theoretical* use of hypotheses should not be confused with their *statistical* use. The latter tends to dominate books on research methods and statistics. As we shall see later, a great deal of research is conducted using samples that are drawn from much larger populations. There are many practical benefits in doing this. If such samples are drawn using statistically random procedures, and if the response rate is very high, a researcher may want to generalize the results found in a sample to the population from which the sample was drawn. Statistical hypotheses perform a role in this generalization process, in making decisions about whether the characteristics, differences or relationships found in a sample can be expected to also exist in the population. Such hypotheses are *not* derived from theory and are *not* tentative answers to research questions. Their function is purely statistical. When research is conducted on a population or a non-random sample, there is no role for statistical hypotheses. However, theoretical hypotheses are relevant in *any* research that requires ‘why’ questions to be answered.

## What are Data?

In the context of social research, the concept of *data* is generally treated as being unproblematic. It is rare to find the concept defined and even rarer to encounter any philosophical consideration of its meaning and role in research. Data are simply regarded as something we collect and analyze in order to arrive at research conclusions.

The concept is frequently equated with the notion of 'empirical evidence', that is, the products of systematic 'observations' made through the use of the human senses. Of course, in social research, observations are made mainly through the use of sight and hearing.

The concept of *observation* is used here in its philosophical sense, that is, as referring to the use of the human senses to produce 'evidence' about the 'empirical' world. This meaning needs to be distinguished from the more specific usage in social research where it refers to methods of data collection that use the sense of sight. In this latter method, 'looking' is distinguished from other major research activities such as 'listening', 'conversing', 'participating', 'experiencing', 'reading' and 'counting'. All of these activities are involved in the philosophical meaning of 'observing'.

Observations in all sciences are also made with the use of instruments, devices that extend the human senses and increase their precision. For example, a thermometer can measure temperature far more precisely and consistently than can the human sense of touch. Its construction is based on notions of hot and cold, more and less, and of an equal interval scale. In short, it has built into it many assumptions and technical ideas that are used to extend differences that can be experienced by touch. Similarly, an attitude scale, consisting of an integrated set of statements to which responses are made, provides a more precise and consistent measure than, say, listening to individuals discussing some issue.

The notion of *empirical evidence* is not as simple as it might seem. It entails complex philosophical ideas that have been vigorously contested. These disagreements centre on different claims that are made about:

- what can be observed;
- what is involved in the act of observing;
- how observations are recorded;
- what kinds of analysis can be done on them; and
- what the products of these observations mean.

There are a number of important and related issues involved in the act of observing. One concerns assumptions that are made about what it is that we observe. A second issue has to do with the act of observing, with the connection between what impinges on the human senses and what it is that produces those impressions. A third issue is concerned with the role of the observer in the process of observing. Can reality be observed directly or can we only observe its 'surface' features? Is it reality that we observe, or do we simply process some mental construction of it? Does what we observe represent what actually exists, or, in the process of observing, do we have to interpret the physical sensations in order to make them meaningful?

Can we observe objectively, that is, without contaminating the impressions received by our senses, or does every act of observing also involve a process of interpretation? These are the kinds of complex issues that lie behind the generation of data. Consciously or unconsciously, every social researcher takes a stand on these issues. The position adopted is likely to be that of the particular research tradition or paradigm within which the researcher has been socialized and/or has chosen to work.

The issue of 'objectivity' is viewed differently in these research traditions. In some traditions it is regarded as an ideal towards which research should strive. It is assumed that a conscientious and well-trained researcher can achieve a satisfactory level of objectivity. The 'problem' of objectivity is dealt with by establishing rules for observing, for collecting data. In other traditions, 'objectivity' is regarded as not only being unattainable but also as being meaningless. In these traditions, the emphasis is on producing 'authentic' accounts of the social reality described by social actors rather than accurate representations of some external reality.

Collecting any kind of data involves processes of interpretation. We have to 'recognize' what we see, we have to 'know' what it is an example of, and we may have to 'relate' it to or 'compare' it with other examples. These activities require the use of concepts, both lay and technical, and whenever we use concepts we need to use meanings and definitions. For example, if we identify a particular interaction episode as involving conflict, the observer needs to have a definition of conflict and to be able to recognize when a sequence of behaviour fits with the definition. Incidents of conflict do not come with labels attached; the observer (with technical concepts) or, perhaps, the participants (with lay concepts) must do the labelling. Defining concepts and labelling social activities are interpretative processes that occur against the background of the observer's assumptions and prior knowledge and experiences. Data collected about, say, the frequency of conflict between parents and children will have been 'manufactured' by a particular researcher. While a researcher may follow rules, criteria and procedures that are regarded by her research community as being appropriate, such rules etc. are simply agreements about how research should be done and cannot guarantee 'pure' uncontaminated data. What they can achieve is comparable data between times, places and researchers.

## *Data and Social Reality*

All major research traditions regard data as providing information about some kind of social phenomenon, and an individual datum as relating to some aspect of that phenomenon. Just what the relationship is between the data and the phenomenon depends to a large extent on the assumptions that are made about the nature of social reality, that is, the *ontological assumptions*. In turn, the procedures that are considered to be appropriate for generating data about that phenomenon depend on the assumptions that are made about how that social reality can be known, that is, the *epistemological assumptions*.

One major research tradition assumes that social reality is external to the people involved: that it is the context in which their activities occur; and that it has the capacity to constrain their actions. Knowledge of this reality can be obtained by establishing a bridge to it by the use of concepts and their measurement. Concepts identify aspects of the reality and instruments are designed to collect data relevant to the concepts. In this way, data are supposed to represent aspects of, or what is going on in, some part of reality. Only those aspects that can be measured are regarded as relevant to research. This tradition is associated with *positivism* and *critical rationalism*, and its data-gathering procedures are mainly quantitative.

A second research tradition adopts different ontological assumptions. In this case, reality is assumed to consist of layers or domains. The 'surface' or empirical layer can be observed in much the same way as the tradition just described. However, reality also has an 'underlying' layer that cannot usually be observed directly. This is the 'real' layer consisting of the structures and mechanisms that produce the regularities that can be observed on the surface. Knowledge of this 'real' layer can only be gained by constructing imaginary models of how these structures and mechanisms might operate. Then, knowing what kinds of things are worth looking for, painstaking research will hopefully produce evidence for their existence, and perhaps will eventually expose them to the surface layer. This position is known as *scientific realism*, and it uses a variety of quantitative and qualitative data-gathering procedures.

A third major research tradition adopts yet another set of ontological assumptions. Social reality is regarded as a social construction that is produced and reproduced by social actors in the course of their everyday lives. It consists of intersubjectively shared, socially constructed meaning and knowledge. This social reality does not exist as an independent, objective world that stands apart from social actors' experience of it. Rather, it is the product of the processes by which social actors together negotiate the meanings of actions and situations. It consists of mutual knowledge – meanings, cultural symbols and social institutions. Social reality is the symbolic world of meanings and interpretations. It is not some 'thing' that may be interpreted in different ways; it *is* those interpretations. However, because these meanings are intersubjective, that is, they are shared, they both facilitate and constrain social activity. With these ontological assumptions, knowledge of social reality can only be achieved by collecting social actors' accounts of *their* reality, and then redescribing these accounts in social scientific language. This position is known as *interpretivism* or *social constructionism*, and its data-gathering procedures are mainly qualitative.

This book is concerned with the first of these traditions.

## *Types of Data*

An important issue in social research is the extent to which a researcher is removed from the phenomenon under investigation. Any 'observer' is, by definition, already one step removed from any social phenomenon by dint of the fact of viewing it from the 'outside'. This means that the processes involved in 'observing' require degrees of interpretation and manipulation. Even data generated first-hand by a researcher have already been subjected to some processing. As we have seen, there is no such thing as 'pure' data. However, not all data are first-hand. A researcher may use data that have been collected by someone else, either in a raw form or analyzed in some way. Hence, social research can be conducted that is more than one step removed from the phenomenon.

This notion of distance from the phenomenon can be categorized into three main types: primary, secondary and tertiary. *Primary data* are generated by a researcher who is responsible for the design of the study and the collection, analysis and reporting of the data. These 'new' data are used to answer specific research questions. The researcher can describe why and how they were collected. *Secondary data* are the raw data that have already been collected by someone else, either for some general information purpose, such as a government census or another official purpose, or for a specific research project. In both cases, the purpose in collecting such data may be different from that of the secondary user, particularly in the case of a previous research project. *Tertiary data* have been analyzed by either the researcher who generated them or an analyst of secondary data. In this case the raw data may not be available, only the results of this analysis.

While primary data can come from many sources, they are characterized by the fact that they are the result of direct contact between the researcher and the source, and that they have been generated by the application of particular methods by the researcher. The researcher, therefore, has control of the production and analysis, and is in a position to judge their quality. This judgement is much more difficult with secondary and tertiary data.

Secondary data can come from the same kind of sources as primary data; the researcher is just another step removed from it. The use of secondary data is often referred to as secondary analysis. It is now common for data sets to be archived and made available for analysis by other researchers. Such data sets constitute the purest form of secondary data. Most substantial surveys have potential for further analysis because they can be interrogated with different research questions.

> Secondary information consists of sources of data and other information collected by others and archived in some form. These sources include government reports, industry studies, archived data sets, and syndicated information services as well as traditional books and journals found in libraries. Secondary information offers relatively quick and inexpensive answers to many questions and is almost always the point of departure for primary research. (Stewart and Kamis, 1984: 1)

While there are obvious advantages in using secondary data, such as savings in time and cost, there are also disadvantages. The most fundamental drawback stems from the fact that this previous research was inevitably done with different aims and research questions. It may also have been based on assumptions, and even prejudices, which are not readily discernible, or which are inconsistent with those a researcher wishes to pursue. Secondly, there is the possibility that not all the areas of interest to the current researcher may have been included. Thirdly, the data may be coded in an inconvenient form. Fourthly, it may be difficult to judge the quality of secondary data; a great deal has to be taken on faith. A fifth disadvantage for some research stems from the fact that the data may be old. There is always a time lag between collection and reporting of results, and even longer before researchers are prepared to archive their data sets. Even some census data may not be published until at least two years after they were collected. However, this time lag may not be a problem in historical, comparative or theoretical studies.

With tertiary data, the researcher is even further removed from the social world and the original primary data. Published reports of research and officially collected 'statistics' invariably include tables of data that have summarized, categorized or have involved the manipulation of raw data. Strictly speaking, most government censuses report data of these kinds, and access to the original data set may not be possible. When government agencies or other bodies do their own analysis on a census, they produce genuine tertiary data. Because control of the steps involved in moving from the original primary data to tertiary data is out of the hands of the researcher, such data must be treated with caution.

Some sources of tertiary data will be more reliable than others. Analysts can adopt an orientation towards the original data, and they can be selective in what is reported. In addition, there is always the possibility of academic fraud. The further a researcher is removed from the original primary data, the greater the risk of unintentional or deliberate distortion.

The purpose of this classification is to sensitize the researcher to the nature of the data being used and its limitations. This discussion brings us back to the key issue: what are data? In particular, it highlights the problem of the gap between the researcher and the social phenomenon that is being investigated.

There is an interesting relationship between types of data and ontological assumptions. Such assumptions about the nature of the reality being investigated will not only have a bearing on what constitutes data but also determine how far a researcher is seen to be removed from that reality. This can be illustrated with reference to the operation of stock markets. All major stock markets in the world produce a numerical indicator that is used to follow movements in that particular market. For example, the New York stock exchange uses the Dow Jones index, the London exchange uses the FTSE 100, and the Tokyo exchange the Nikkei. The share prices of a selection of stocks are integrated into a summary number. This number or indicator is used to measure the behaviour of ‘the market’. Trends can be calculated and, perhaps, models and theories developed about cycles or stages in these trends.

But what kind of data are these indices? The answer to this question depends on what view of reality is adopted. The notion of ‘the market’ is an abstract idea that can refer to an entity that exists independently of the people who buy and sell shares. Analysts frequently attribute the market with human or animal qualities: it has ‘sentiments’, it ‘looks for directions’, it acts like a bull or a bear. Hence, ‘the market’ can be regarded as constituting an independent reality. From these assumptions, the market indicator might be regarded as primary data; it measures the behaviour of ‘the market’. The share prices are the raw data.

Another (albeit much less common) set of assumptions would be to regard the worldviews and behaviour of the people who buy and sell shares as constituting the basic social phenomenon. The decisions and actions of these people generate the fluctuating prices of shares. The stockbrokers through whom these people conduct their share transactions are equivalent to researchers who then feed the outcomes of the decisions of these people into a particular market’s database from which the price of any shares, at any time, can be determined and trends plotted. Other researchers then take these average prices and do some further analysis to produce a share price index. Further researchers can then use the changes in the index to trace movements in ‘the market’. Therefore, the price that individual investors pay for their parcel of shares is equivalent to primary data, the closing or average price of the shares in any particular company represents secondary data, and the share price index represents tertiary data.

This example illustrates two things. First, it shows that how data are viewed depends on the ontological assumptions about the social phenomenon being investigated. Second, it shows that what is regarded as reality determines what types of data are used. Reality can be either a reified abstraction, such as ‘the market’, or it can be the interpretations and activities of particular social actors, such as investors. Movements in a share price index can mean different things depending on the assumptions that are adopted. It can be a direct, primary measure of a particular reality, or it can be an indirect, tertiary measure of a different kind of reality. Hence, knowing what data refer to, and how they should be interpreted, depends on what is assumed as being the reality under investigation, and the type of data that are being used.

## *Forms of Data*

Social science data are produced in two main forms, in *numbers* or in *words*. This distinction is usually referred to as either *quantitative* or *qualitative* data. There seems to be a common belief among many researchers, and consumers of their products, that numerical data are needed in scientific research to ensure objective and accurate results. Somehow, data in words tend to be regarded as being not only less precise but also less reliable. These views still persist in many circles, even although non-numerical data are now more widely accepted. As we shall see shortly, the distinction between words and numbers, between qualitative and quantitative data, is not a simple one.

It can be argued that all primary data start out as words. Some data are recorded in words, they remain in words throughout the analysis, and the findings are reported in words. The original words will be transformed and manipulated into other words, and these processes may be repeated more than once. The level of the language will change, moving from lay language to technical language. Nevertheless, throughout the research, the medium is always words.

In other research, the initial communication will be transformed into numbers immediately, or prior to the analysis. The former involves the use of pre-coded response categories, and the latter the post-coding of answers or information provided in words, as in the case of open-ended questions in a questionnaire. Numbers are attached to both sets of categories and the subsequent analysis will be numerical. The findings of the research will be presented in numerical summaries and tables. However, words will have to be introduced to interpret and elaborate the numerical findings. Hence, in quantitative studies, data normally begin in words, are transformed into numbers, are subjected to different levels of statistical manipulation, and are reported in both numbers and words; from words to numbers and back to words. The interesting point here is whose words were used in the first place and what process was used to generate them. In the case where responses are made into a predetermined set of categories, the questions and the categories will be in the researcher’s words; the respondent only has to interpret both. However, this is a big ‘only’. As Foddy (1993) and Pawson (1995, 1996) have pointed out, this is a complex process that requires much more attention and understanding than it has normally been given.

Sophisticated numerical transformations can occur as part of the analysis stage. For example, responses to a set of attitude statements, in categories ranging from ‘strongly agree’ to ‘strongly disagree’, can be numbered, say, from 1 to 5. The direction of the numbering will depend on whether a statement expresses positive or negative attitudes on the topic being investigated, and on whether positive attitudes are to be given high or low scores. Subject to an appropriate test, these scores can be combined to produce a total score. Such scores are well removed from the respondent's original reading of the words in the statements and the recording of a response in a category with a label in words.

So far, this discussion of the use of words and numbers has been confined to the collection of primary data. However, these kinds of manipulations may have already occurred in secondary data, and will certainly have occurred in tertiary data.

The controversial issue in all of this is the effect that any form of manipulation has on the relationship of the data to the reality it is supposed to measure. If all observation involves interpretation, then some kind of manipulation is involved from the very beginning. Even if a conversation is recorded unobtrusively, any attempt to understand what went on requires the researcher to make interpretations and to use concepts. How much manipulation occurs is a matter of choice.

A more important issue is the effect of transforming words into numbers. Researchers who prefer to remain qualitative through all stages of a research project may argue that it is bad enough to take lay language and manipulate it into technical language without translating either of them into the language of mathematics. A common fear about such translations is that they end up distorting the social world out of all recognition, with the result that research reports based on them become either meaningless or, possibly, dangerous if acted on.

The reason for this extended discussion of issues involved in transforming words into numbers is to highlight the inherent problems associated with interpreting quantitative data and, hence, its analysis. Because of the steps involved in transforming some kind of social reality into the language of mathematics, and the potential for losing the plot along the way, the interpretation of the results produced by quantitative analysis must be done with full awareness of the limitations involved.

## *Concepts and Variables*

It is conventional practice to regard quantitative data as consisting of *variables*. These variables normally start out as concepts, coming from either research questions or hypotheses. First, it is necessary to define the concept in terms of the meaning it is to have in a particular research project. For example, age might be defined as 'years since birth', and education as 'the highest level of formal qualification obtained'. Unless there is some good reason to do otherwise, it is good practice to employ a definition already in use in that particular field of research. In this way, results from different studies can be easily compared.

The second step is to *operationalize* the concept to show how data related to it will be generated. This requires the specification of the procedures that will be used to classify or measure the phenomenon being investigated. For example, in order to measure a person's age, it is necessary either to ask them or to obtain the information from some kind of record, such as a birth certificate. Similarly, with education, you can either ask the person what their highest qualification is, or you can refer to appropriate documents or records. The way a concept is defined and measured has important consequences for the kinds of data analysis that can be undertaken.

The idea behind a variable is that it can have different values, that characteristics of objects, events or people can be measured along some continuum that forms a uniform numerical scale. This is the nature of metric measurement. For example, age (in years) and attitudes towards some object (in scores) are variables. However, other kinds of characteristics, such as religion, do not share this property. They are measured in terms of a set of different categories. Something can be identified as being in a particular category (e.g. female), but there is no variation within the category, only differences between categories (e.g. males and females). As there is no variability within such categories, the results of such measurement are not strictly variables. They could be called variates, but this concept also has another meaning in statistics. Therefore, I shall follow the established convention of referring to all kinds of quantitative measurement as variables. It is to the different kinds or levels of measurement that we now turn.

## *Levels of Measurement*

In quantitative research, aspects of social reality are transformed into numbers in different ways. *Measurement* is achieved either by the assignment of objects, events or people to discrete categories, or by the identification of their characteristics on a numerical scale, according to arbitrary rules. The former is referred to here as *categorical* measurement and the latter as *metric* measurement. Within these *levels of measurement* are two further levels: nominal and ordinal, and interval and ratio, respectively.

### *Categorical Measurement*

Everyday life would be impossible without the use of numbers. However, using numbers does not mean that we need to use complex arithmetic or mathematics. Frequently, numbers are simply used to identify objects, events or people. Equipment and other objects are given serial numbers or licence numbers so that they can be uniquely identified. Days of the month and the years of a millennium are numbered in sequence. The steps involved in assembling an object are numbered. People who make purchases in a shop can be given numbers to ensure they are served in order. In none of these examples are the numbers manipulated; they are simple used as a form of identification, and, in some cases, to establish an order or sequence. The alphabet could just as easily be used, and sometimes is, except that it is much more restricted than our usual number system as the latter has no absolute limit. This elementary way of using numbers in real life and in the social sciences is known as *categorical measurement*.

As has already been implied, categorical measurement can be of two types. One involves assigning numbers to categories that identify different types of objects, event or people; in the other, numbers are used to establish a sequence of objects, events or people. Categories can either identify differences or they can be ordered along some dimension or continuum. The former is referred to as nominal-level measurement, and the latter as ordinal-level measurement.

### Nominal-level measurement

In *nominal-level measurement*, the categories must be homogeneous, mutually exclusive and exhaustive. This means that all objects, events or people allocated to a particular category must share the same characteristics, they can only be allocated to one category, and all of them can be allocated to some category in the set. The categories have no intrinsic order to them, as is the case for the categories of gender or religion. People can also be assigned numbers arbitrarily according to some criterion, such as different categories of eye colour – blue (1), brown (2), green (3), etc. However, these categories have no intrinsic order (except, of course, on the colour spectrum).

### Ordinal-level measurement

The same conditions apply in *ordinal-level measurement*, with the addition that the categories *are* ordered along some continuum. For example, people can be assigned numbers in terms of the order in which they cross the finishing line in a race, they can be assigned social class categories ('upper', 'middle' and 'lower') according to their income or occupational status, or they can be assigned to age categories ('old', 'middle-aged' and 'young') according to some criterion. A progression or a hierarchy is present in each of these examples.

However, the intervals between such ordinal categories need not be equal. For example, the response categories of 'often' (1), 'occasionally' (2) and 'never' (3) cannot be assumed to be equally spaced by researchers, because it cannot be assumed that respondents regard them this way. When the numbers in brackets are assigned to these categories, they only indicate the order in the sequence, not how much of a difference there is between these categories. They could just as easily have been identified with 'A', 'B' and 'C', and these symbols certainly do not imply any difference in magnitude.

Similarly, the commonly used Likert categories for responses to attitude statements, 'strongly agree', 'agree', 'neither agree nor disagree', disagree', and 'strongly disagree', are not necessarily evenly spaced along this level of agreement continuum, although researchers frequently assume that they are. When this assumption is introduced, an ordinal-level measure becomes an interval-level measure with discrete categories.

## Metric Measurement

There are more sophisticated ways in which numbers can be used than those just discussed. The introduction of the simple idea of equal or measurable intervals between positions on a continuum transforms categorical measurement into *metric measurement*. Instead of assigning objects, events or people to a set of categories, they are assigned a number from a particular kind of scale of numbers, with equal intervals between the positions on the scale. For example, we measure a person's height by assigning a number from a measuring scale. We measure intelligence by assigning a person a number from a scale that represents different levels of intelligence (IQ). Of course, with categorical measurement, it is necessary to have or to create a set of categories into which whatever is being measured can be assigned. However, these categories do not have any numerical relationships and, therefore, cannot have the rules of a number system applied to them.

Hence, the critical step in this transition from categorical to metric measurement is the mapping of the things being measured onto a scale. The scale has to exist, or be created, before the measurements are made, and these scales embody the properties and rules of a number system. Measuring a person's height clearly illustrates this. You have to have a measuring instrument, such as a long ruler or tape measure, before a person's height can be established. We can describe people as being 'tall', 'average' or 'short'. Such ordinal-level categories allow us to compare people's height only in very crude terms. Adding numbers to the categories, say '1', '2' and '3', neither adds precision to the measurement nor does it allow us to assume that the intervals between the categories are equal. Alternatively, we could line up a group of people, from the tallest to the shortest, and give them numbers in sequence. Each number simply indicates where a person is in the order and has nothing to do with the actual magnitude of their height. In addition, the differences in height between neighbouring people will vary and the number assigned to them will not indicate this. However, once we stand them beside a scale in, say, centimetres, we can get a measure of magnitude, and because they are all measured against the same scale we can make precise comparisons between any members of the group. Precision of measurement is only one of the considerations here. The important change is that much more sophisticated forms of analysis can now be used which, in turn, means that more sophisticated answers can be given to research questions.

All metric scales of measurement are human inventions. The way in which points on the scale are assigned numbers, the size of the intervals between those points, whether or not there are gradations between these points, and where the numbering starts, are all arbitrary. Scales differ in how the zero point is established. Some scales have an absolute or true zero, while for others there is no meaningful zero, that is, the position of zero is arbitrary.

### *Interval-level measurement*

*Interval-level measurement* is achieved when the categories or scores on a scale are the same distance apart. Whereas in ordinal-level measurement the numbers '1', '2' and '3' only indicate relative position, say in finishing a race, in interval-level measurement, the numbers are assumed to be the same distance apart – the interval between '1' and '2' is the same as the interval between '2' and '3'. As the numbers are equally spaced on the scale, each interval has the same value.

The distinguishing feature of interval-level measurement is that the zero is arbitrary. Whatever is being measured cannot have a meaningful zero value. For example, an attitude scale may have possible scores that range from 10 to 50. Such scores could have been derived from an attitude scale of ten items, using five response categories (from 'strongly agree' to 'strongly disagree') with the categories being assigned numbers from 1 to 5 in the direction appropriate to the wording (positive or negative) of the item.[2] However, these scores could just as easily have ranged from 0 to 40 (with categories assigned numbers from 0 to 4) without altering the relative interval between any two scores. In this case, a zero score is achieved by an arbitrary decision about what numbers to assign to the response categories. It makes no sense to speak of a zero attitude, only relatively more positive or negative attitudes.

### *Ratio-level measurement*

*Ratio-level measurement* is the same as interval-level measurement except that it has an absolute or true zero. For example, goals scored in football, or age in years, both have absolute or true zeros; it is possible for a team to score no goals, and a person's age is normally calculated from the time of birth – point zero.

Ratio-level measurement is not common in the social sciences and is limited to examples such as age (in years), education (in years) and income (in dollars or other currencies). This level of measurement has only a few advantages over the interval level of measurement, mainly that statements such as 'double' or 'half' can be made. For example, we can say that a person aged 60 years is twice as old as a person aged 30 years, or that an income of $20,000 is only half that of $40,000. These kinds of statements cannot be made with interval-level variables. For example, with attitude scales, such as those discussed above, it is not legitimate to say that one score (say 40) is twice as positive as another (say 20). What we can say is that one score is higher, or lower, than another by so many scale points (a score of 40 is 10 points higher than a score of 30, and the latter is 10 points higher than a score of 20) and that an interval of, say 10 points, is the same anywhere on the scale. The same applies to scales used to measure temperature. Because the commonly used temperature scales, Celsius and Fahrenheit, both have arbitrary zeros, we cannot say that a temperature of 30°C is twice as hot as 15°C, but the interval between 15°C and 30°C is the same as that between 30°C and 45°C. Similarly, not only is 30°C a different temperature than 30° Fahrenheit, but an interval of 15° is different on each scale. However, as the kelvin scale does have a true zero, the absolute minimum temperature that is possible, a temperature of 400K is twice as hot as 200K.

Compared to ratio-level measurement, it is the arbitrary zero that creates the limitations in interval-level measurement. In most social science research, this limitation is not critical; interval-level measurement is usually adequate for most sophisticated forms of analysis. However, we need to be aware of the limitations and avoid drawing illegitimate conclusions from interval-level data.

## *Discrete and Continuous Measurement*

Metric scales also differ in terms of whether the points on the scale are discrete or continuous. A *discrete* or discontinuous scale usually has units in whole numbers and the intervals between the numbers are usually equal. Arithmetical procedures, such as adding, subtracting, multiplying and dividing, are permissible. On the other hand, a *continuous* scale will have an unlimited number of possible values (e.g. fractions or decimal points) between the whole numbers. An example of the former is the number of children in a family and, of the latter, a person's height in metres, centimetres, millimetres, etc. We cannot speak of a family having 1.8 children (although the average size of families in a country might be expressed in this way), but we can speak of a person being 1.8 metres in height. When continuous scales are used, the values may also be expressed in whole numbers due to rounding to the nearest number.

## *Review*

The characteristics of the four levels of measurement are summarized in Table 1.2. They differ in their degree of precision, ranging from the least precise (nominal) to the most precise (ratio). The different characteristics, and the range of precision, mean that different mathematical procedures are appropriate at each level. It is too soon to discuss these differences here; they will emerge throughout Chapters 3–6.

However, a word of caution is appropriate. It is very easy to be seduced by the precision and sophistication of interval-level and ratio-level measurement, regardless of whether they are necessary or theoretically and philosophically appropriate. The crucial question is what is necessary in order to answer the research question under consideration. This relates to other aspects of social research, such as the choice of data sources, the method of selection from these sources and the method of data collection. The latter, of course, will have a considerable bearing on the type of analysis that can and should be used. In quantitative research, the choice of level of measurement at the data-collection stage, and the transformations that may be made, including data reduction, will determine the types of analysis that can be used.

Table 1.2 *Levels of measurement*

| Level | Description | Types of categories | Examples |
|---|---|---|---|
| Nominal | A set of categories for classifying objects, events or people, with no assumptions about order. | Categories are homogeneous, mutually exclusive and exhaustive. | Marital status<br>Religion<br>Ethnicity |
| Ordinal | As for nominal-level measurement, except the categories are ordered from highest to lowest. | Categories lie along a continuum but the distances between them cannot be assumed to be equal. | Frequency (often, sometimes, never)<br>Likert scale |
| Interval | A set of ordered and equal-interval categories on a contrived measurement scale. | Categories may be discrete or continuous with arbitrary intervals and zero point. | Attitude score<br>IQ score<br>Celsius scale |
| Ratio | As for interval-level measurement | Categories may be discrete or continuous but with an absolute zero. | Age<br>Income<br>No. of children |

Finally, it is important to note that some writers refer to categorical data as qualitative and metric data as quantitative. This is based on the idea that qualitative data lack the capacity for manipulation other than adding up the number in the categories and calculating percentages or proportions. This usage is not adopted here. Rather, 'qualitative' and 'quantitative' are used to refer to data in words and numbers, respectively. Categorical data involve the use of numbers and not words, allowing for simple numerical calculations. According to the definitions being used here, categorical data are clearly quantitative.

## *Transformations between Levels of Measurement*

It is possible to transform metric data into categorical data but, in general, not the reverse. For example, in an attitude scale, scores can be divided into a number of ranges (e.g. 10–19, 20–29, 30–39, 40–50) and labels applied to these categories (e.g. 'low', 'moderate', 'high' and 'very high'). Thus, interval-level data can be transformed into ordinal-level data. Something similar could be done with age (in years) by creating age categories that may not cover the same range, say, 20–24, 25–34, 35–54, 55+. In this case, the transformation is from ratio level to ordinal. While such transformations may be useful for understanding particular variables, and relationships between variables, measurement precision is lost in the process, and the types of analysis that can be applied are reduced in sophistication. It is important to note, however, that if a range of ages or scores is grouped into categories of equal size, for example, 20–29, 30–39, 40–49, 50–59, 60–69, etc., the categories *can* be regarded as being at the interval level; they cover equal age intervals, thus making their midpoints equal distances apart. All that has changed is the unit of measurement, in 10-year age intervals rather than 1-year intervals.

There are a few cases in which it is possible to transform lower-level measurement to a higher level. For example, it is possible to take a set of nominal categories, such as religious denomination, and introduce an order using a particular criterion. For example, religious categories could be ordered in terms of the proportion of a population that adheres to each one, or, more complexly, in terms of some theological dimension. Similarly for categories of political party preference, although in this case dominant political ideology would replace theology. In a way, such procedures are more about analysis than measurement; they add something to the level of measurement used in order to facilitate the analysis.

The reason why careful attention must be given to level of measurement in quantitative research is that the choice of level determines the methods of analysis that can be undertaken. Therefore, in designing a research project, decisions about the level of measurement to be used for each variable need to anticipate the type of analysis that will be required to answer the relevant research question(s). Of course, for certain kinds of variables, such as gender, ethnicity and religious affiliation, there are limited options. However, for other variables, such as age and income, there are definite choices. For example, if age is pre-coded in categories of unequal age ranges, then the analysis cannot go beyond the ordinal level. However, if age was recorded in actual years, then analysis can operate at the ratio level, and transformations also made to a lower level of measurement. Such a simple decision at the data-collection stage can have significant repercussions at the data-analysis stage. The significance of the level of measurement for choice of method of analysis will structure the discussion in Chapters 3–6.

## What is Data Analysis?

All social research should be directed towards answering research questions about characteristics, relationships, patterns or influences in some social phenomenon. Once appropriate data have been collected or generated, it is possible to see whether, and to what extent, the research questions can be answered. *Data analysis* is one step, and an important one, in this process. In some cases, the testing of theoretical hypotheses, that is, possible answers to 'why' research questions, is an intermediary step. In other cases, the research questions will be answered directly by an appropriate method of analysis.

The processes by which selection is made from the sources of data can also have a major impact on the choice of methods of data analysis. The major consideration in selecting data is the choice between using a population and a sample of some kind. If sampling is used, the type of data analysis that is appropriate will depend on whether probability or non-probability sampling is used. Hence, it is necessary to review briefly how and why the processes of selecting data affect the choice of methods of data analysis.

## Types of Analysis

Various methods of data analysis are used to describe the characteristics of social phenomena, and to understand, explain and predict patterns in social life or in the relationships between aspects of social phenomena. In addition, one type of analysis is concerned with estimating whether characteristics and relationships found in a sample randomly drawn from a population could also be expected to exist in the population. Hence, analysis can be divided into four types: univariate descriptive, bivariate descriptive, explanatory and inferential.

### Univariate Descriptive Analysis

*Univariate descriptive analysis* is used to represent the characteristics of some social phenomenon (e.g. student academic performance on a particular course). This can be done in a number of ways:

- by counting the frequency with which some characteristic occurs (e.g. the total marks[3] students receive on a particular course);
- by grouping scores of a certain range into categories and presenting these frequencies in pictorial or graphical form (e.g. student's total marks);
- by calculating measures of central tendency (e.g. the mean marks obtained by students on the course); and
- by graphing and/or calculating the spread of frequencies around this centre point (e.g. plotting a line graph of the frequency with which particular marks were obtained, or calculating a statistic that measures the dispersion around the mean).

There are clearly many ways in which the phenomenon of student academic performance can be described and compared. The principles of each of these methods will be elaborated later in this chapter, and they will be illustrated in later chapters.

### Bivariate Descriptive Analysis

*Bivariate descriptive analysis* is a step along the path from univariate analysis to explanatory analysis. It involves either establishing similarities or differences between the characteristics of categories of objects, events or people, or describing patterns or connections between such characteristics.

Typically, patterns are investigated by determining the extent to which the position of objects, events or persons on one variable coincides with their position on another variable. For example, does the position of people on a measure of height coincide with their position on a measure of weight? If the tallest people are also the heaviest, and vice versa, then these two measures can be said to be associated. Sometimes this is expressed in terms of whether position on one measure is a good predictor of position on another measure, that is, whether the height of people is a good predictor of their weight.

Continuing the example about student academic performance, we can:

- compare categories in terms of averages (e.g. differences between the mean marks of female and male students, or students of different ethnic backgrounds); and
- establish the strength of the relationship between two characteristics (e.g. measuring the association between gender and honours grades, or ethnicity and grade point average).

## *Explanatory Analysis*

To go beyond describing characteristics and establishing relationships, that is, to go beyond answering 'what' questions to addressing 'why' questions, takes us into the complex and difficult territory of *explanatory analysis* and the much disputed notion of *causation*. It is a common belief that establishing an explanation involves finding the cause or causes for the patterns and sequences in social life. Explanations are supposed to tell us why certain things occur together or follow one another in time. However, there are not only many views on how this can be achieved, but also dissenting voices that claim such a task is impossible.

There are two main views on the nature of causation: the successionist and the generative (Harré, 1972; Pawson, 1989; Pawson and Tilley, 1997). The *successionist view of causation* is based on the idea that events in the world can be explained if they follow a regular sequence. In fact, according to this view, there are no such things as causes, only connections or sequences between events in the world. For example, if we apply heat to water it will turn into steam when the temperature reaches 100°C. The event of heating is followed by another event, the change from liquid to gas. The change of status is therefore explained by the event that preceded it. Hence, the sequence of these events forms a 'natural necessity'; they could not happen otherwise. The task of science is to discover these regularities and then use them for both explanation and prediction. Explanation is achieved by pointing to prior events in the sequence, and prediction is achieved by knowing what follows in the sequence. This view is associated with the view of social science known as positivism.

While the idea of events occurring in well-established sequences implies that single events have single causes, many approaches to causal explanation recognize the possibility of multiple causes, that is, that more than one event might have to precede the event to be explained in order for it to occur. These preceding events may occur in parallel (concurrently) or in a chain reaction (sequentially), or in some combination of concurrence and sequence.

The philosophical notions of causation are frequently translated into the language of two types of conditions that are responsible for the occurrence of an event, necessary and sufficient conditions. A *necessary condition* is one that needs to be present in order for an event to occur. There may be a number of necessary conditions, but even together they may not produce an event. A *sufficient condition* is one that will lead to the occurrence of an event on its own or, perhaps, in combination with one or two other conditions. However, sufficient conditions will usually need to operate in the context of some necessary conditions. Hence, necessary conditions can be regarded as the contextual factors that need to be present and sufficient conditions as those that actually produce the event in that context. Normally, the two sets of conditions will work together. The challenge for the researcher is to find the complete set of necessary and sufficient conditions in order to explain an event. Clearly, this is an impossible task. Even trying to identify as many conditions as possible has its limitations. We need reasons for selecting possible conditions, and these are supplied by a good theory.

An elaboration of the successionist view has argued that explanation is achieved by finding a well-tested theory from which the event or pattern to be explained can be deduced once the conditions under which it is known to operate have been specified. Specifying different conditions makes it possible to predict new events or patterns and, therefore, to test the theory. However, such predictions only apply to the conditions that have been specified in the theory, not to some future time in which the conditions may not be known. Such theories are made up of statements of relationships or connections between concepts or events. It is by combining such statements into a logical argument that explanation can be achieved. This tradition is associated with the view of social science known as critical rationalism.

When research does not satisfy experimental requirements, as in survey research, it is a common practice to translate causal language into relationships between two types of variables. One variable, the values of which are to be explained, is referred to as the *dependent* or *outcome variable*, while those that are involved in producing these values are referred to as *independent* or *predictor variables*. Stated differently, the values of the dependent variable are influenced or predicted by the values of the independent variable or variables. For example, academic performance (the dependent variable) might be influenced or predicted by students' ethnicity (the independent variable).[4] In this example, there is no attempt to completely explain academic performance, only to indicate a factor that might contribute to it. Hence, it is a rather low level of causal analysis.

While it might be strengthened by the inclusion of other independent variables, this kind of research cannot produce conclusive explanations. For one thing, assumptions need to be made about the direction of influence among variables and, for another, the nature of the causal relationship is frequently left rather vague. In the case of the relationship between ethnicity and academic performance, even if there is an association, and it is possible to analyze this in terms of the influence of one variable on the other, we are still left with the problem of what it is about ethnicity that is responsible for this. The connection only makes sense when it is theorized and the theory is thoroughly tested. Something else is required to link these two variables. Such a theory might include ideas about the influence of family attitudes and experiences and the quality of early formal education. This is the core argument of those who advocate the generative view of causation.

The successionist approach has been severely criticized by the advocates of the *generative view of causation* (see, for example, Bhaskar, 1979; Pawson, 1989; Pawson and Tilley, 1997). While successionists have restricted themselves to sequences of observable events, or the connections between concepts, the supporters of the generative position argue that events cannot be seen as being discrete and isolatable; they are part of a network or system of events. To isolate them artificially is to produce connections that may bear little relationship to how things actually behave. In addition, they argue that establishing connections or relationships is only the starting point. It is necessary to discover the underlying structures and mechanisms that are responsible for producing such connections. Observed patterns or regularities are explained by discovering the structures and mechanisms that generate the observed phenomenon. As these mechanisms may not always be obvious or readily observable, their existence may have to be postulated and then established. Explanations are produced by the 'causal powers' or 'tendencies' of things to behave in a particular way. 'A mechanism is not thus a single *variable* but an *account* of the constitution and behaviour of those things that are responsible for the manifest reality' (Pawson, 1989: 130).

The supporters of the generative view also argue that prediction is not possible because to be able to do so requires knowledge of all the conditions that are relevant to the operation of such structures and mechanisms. As social phenomena occur in open systems, it is not possible to know in advance what conditions will be operating and, therefore, to be able to predict what will occur. Hence, researchers must be content with trying to establish mechanisms and the conditions under which they operate, after the event. The philosophical tradition on which the generative view is based is scientific or critical realism.

One particular social science tradition totally rejects the idea of causation. This is based on the argument that the successionist view is only relevant to the natural sciences. Because the subject matter studied by the social sciences is fundamentally different from that of the natural sciences, the only appropriate approach is to try to understand social phenomena in terms of the reasons people can give for their actions rather than in terms of some notion of independent causes. It is the socially constructed nature of social reality that distinguishes it from natural or physical phenomena, and this difference requires the use of very different ways of understanding social life. This tradition is known as interpretivism or social constructionism. For a review of these traditions, see Blaikie (1993a, 2000).

## *Inferential Analysis*

*Inferential analysis* is used with data obtained from a sample to estimate the characteristics of or patterns in the population from which the sample was drawn. This kind of analysis is only appropriate when the sample is drawn using probability or random selection procedures. There are two main types of inferential analysis: that used to estimate characteristics of a population from sample data (e.g. the mean age of its members); and that used to establish whether a pattern or a relationship found to exist in a sample could be expected to also occur in the population from which it was drawn (e.g. a relationship between ethnicity and academic performance).

Inferential analysis on sample characteristics receives limited attention in social research. Apart from some opinion polls (and even in these its use is rather rare), few social researchers actually calculate the likely population characteristics from their sample data. The sample data are usually just presented and assumed to be the same in the population. In contrast, a great deal of attention is given to testing whether relationships found in a sample can be expected to exist in a population. This is done by using *tests of significance*. Unfortunately, these tests are frequently applied beyond the situations is which they are appropriate. Their purpose is poorly understood, and they are also widely misused. The common forms of confusion are that:

- using such tests will tell you what is important in your data;
- chance factors that might adversely affect the collection of data can somehow be detected and allowed for by using these test;
- the tests tell you how closely two characteristics (variables) are associated; and
- the tests should be applied to all results, regardless of whether data were collected from a population, a random sample or a non-random sample.

These issues will be addressed in Chapter 6.

## Logics of Enquiry and Data Analysis

In addition to the competing views on what constitutes causation, there is another related area of dispute in the social sciences. This centres on whether the ways of answering 'why' questions that are appropriate in the natural sciences are also appropriate in the social sciences (for a review of this issue, see Blaikie, 1993a). There are two dominant schools of thought. One argues that the logic of explanation used in the natural sciences is also appropriate in the social sciences. The other argues that the peculiar nature of the subject matter studied in the social sciences limits the kind of answers the social scientist can offer to 'why' questions. All that is possible is to understand social phenomena by establishing the reasons people give for their actions. In short, there are debates about whether it is possible to establish causal *explanations* in the social sciences or whether *understanding*, based on social actors accounts, is all that is possible and necessary.

Among those who advocate the use of the logic of explanation adopted in the natural sciences, there have been disputes about what this logic should be. The earliest view – first advocated in seventeenth century by Francis Bacon (see Bacon, 1889), with important contributions during the 1840s by William Whewell and John Stuart Mill (see Whewell, 1847; Mill, 1947) was that *inductive logic* was the appropriate scientific 'method'. Accumulated data are used to produce generalizations about the patterns or connections between events or variables. In the 1930s, this view was severely criticized and an alternative proposed by Popper (1959) in the form of a *deductive logic* of explanation. In this case, a researcher starts with a theory that provides a possible explanation, and then proceeds to test the theory by deducing from it one or more hypotheses, and then matching the hypotheses against appropriate data. More recently, both of these positions have been rejected by Harré (1961, 1970, 1972; Harré and Secord, 1972) and Bhaskar (1979) who have proposed the use of *retroductive logic*. They have argued that the inductive approach simply produces descriptions that still have to be explained by locating 'real' structures and mechanisms that produce the effects that can be observed. This is done by building models or developing pictures of these structures and mechanisms such that, if they exist and act in the way postulated, they would account for the phenomenon being examined. Structures and mechanisms are not discovered by accumulating data but by looking for evidence that would confirm their existence. These authors have argued that the discovery of atoms and viruses followed this type of logic. While the idea of atoms existed long before they were observed directly (in the 1960s), scientists, acting on the assumption that they did exist and behaved as imagined, were able to create the atomic bomb.

A fourth logic of enquiry, *abductive logic*, rejects the idea of explanation and causation in favour of understanding. Such understanding comes from 'thick' descriptions and the grasping of social actors' meanings and interpretations. These different logics of enquiry have, with some modifications, been presented as alternative research strategies, that is, as alternative ways of answering research questions, particularly 'why' questions (see Blaikie, 1993a, 2000).

The different views of causation have important consequences for the way we conduct social research and undertake data analysis. Add to that the use of different research strategies, and serious implications for data analysis become evident. The methods of analysis to be discussed in this book are really only appropriate for the successionist view of causation and the inductive and deductive research strategies. Hence, there are other ways of answering 'why' research questions that require different kinds of data analysis, and these entail different views on what constitutes data. Quantitative data analysis is only one kind of data analysis.

## Summary

- Social research must start with a research problem, an intellectual puzzle or a practical problem.
- Social research is about answering three types of research questions: 'what', 'why' and 'how' questions.
- Social research pursues a range of objectives: exploration, description, explanation, understanding, prediction, intervention, evaluation and impact assessment. The objectives of explanation and understanding are expressed as 'why' questions and the objective of intervention as 'how' questions. The remaining objectives are mostly related to 'what' questions.
- Research objectives are frequently pursued in a logical sequence, the most common of which is description, explanation/understanding and intervention.

- Theoretical hypotheses provide possible answers to 'why' research questions.
- Statistical hypotheses are used to establish whether patterns found in a random sample are present in its population. This is their only role in social research.
- Data are produced by the use of the human senses, mainly sight and hearing, and through the use of instruments that extend and systematize their use. This requires agreement about rules and criteria. Such procedures do not guarantee objectivity, only comparability between times, places and researchers.
- All forms of measurement in the social sciences are socially constructed by experts, the data they produce, and the results that follow, have to be understood in terms of the assumptions and procedures adopted.
- These assumptions are both ontological and epistemological and, while they are usually taken for granted, they can be understood with reference to one of the major philosophies of social science: positivism, critical rationalism, scientific realism and interpretivism.
- There are three types of social science data: primary, secondary and tertiary. Each type has its advantages and disadvantages and varies in terms of the distance it creates between the researcher and the social reality being studied.
- Social science data can be either qualitative or quantitative, in either words or numbers. Transformations between words and numbers, or in the reverse direction, can occur at various stages in a research project.
- Quantitative data are expressed in the form of variables that are produced by operationalizing the key concepts in research questions and theoretical hypotheses.
- Concepts can be measured at four different levels. From lowest to highest, these are nominal, ordinal, interval and ratio. The first two produce categorical variables, because objects, events or people are placed into one of a set of mutually exclusive categories. The second two produce metric variables, as objects, events or people are mapped onto an established measuring scale.
- Metric variables can be either discrete or continuous. The former consist only of whole numbers, while the latter have an unlimited number of possible values between the whole numbers.
- Data can be transformed from metric to categorical. While this means some loss of information, and entails the use of less sophisticated forms of analysis, it may allow for a better understanding of the characteristics or relationships being examined.
- There are four main types of data analysis: univariate descriptive, bivariate descriptive, explanatory and inferential. The first two are concerned with characteristics and patterns in data, the third with influence between variables and the fourth with generalizing from samples to populations. Explanatory analysis is the ultimate objective in social research and is also the most complex.
- Explanation is usually associated with the idea of causation. However, this is a highly contested notion and has to be reduced to simpler ideas to be useful in social research. One way of doing this is in terms of the influence between independent (predictor) and dependent (outcome) variables.
- Different views of causation are associated with the major logics of enquiry: inductive, deductive, retroductive and abductive. These logics also constitute different research

strategies.

## Notes

[1]For a more detailed discussion of many of these issues, see Blaikie (2000).

[2]For the moment we will leave aside the debates about whether such a scoring procedure is legitimate.

[3]For simplicity, marks, say out of 100, are used in these examples rather than grades (e.g. A, B, C) or grade point averages.

[4]This example is not intended to be racist. It happened to be a very ‘hot’ political issue in Malaysia at the time of writing.