

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 simply 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.<sup>2</sup> 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



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 Religion 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) 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 IQ score Celsius scale
Ratio	As for interval-level measurement	Categories may be discrete or continuous but with an absolute zero.	Age Income No. of children

stage, and the transformations that may be made, including data reduction, will determine the types of analysis that can be used.

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.