

measurement to be used and on the kind of analysis to be undertaken. If some of the variables are nominal, and three of these are to be used in a three-way cross-tabulation, the size of the group would need to be about ten times the number of cells in this table. For example, if country of birth (coded into three categories) is to be cross-tabulated with political party preference (three categories), and if the first variable were to be controlled by year of migration (coded into three time periods), then a table of twenty-seven cells would be produced requiring a sub-sample of 270. If the smallest ethnic community makes up 10 per cent of the population, then a total sample of 2,700 would be required. Of course, one way to reduce the total sample size would be to use a stratified sample by ethnic community, and different sampling ratios in each stratum, to make all sub-samples of 270. If there were five ethnic communities, the total sample would be exactly half (1,350), thus producing a considerable reduction in the cost of the study. However, if the variables to be analysed are interval or ratio, much smaller numbers can be used. One rule of thumb is to have a minimum of fifty in each subgroup, but, clearly, many things should be considered in making this decision.

Here are some important relationships between sample size, error and accuracy.

- As sample size increases, sampling error decreases and sample reliability increases.
- As population homogeneity decreases, sample error increases and sample reliability decreases.
- As sample error increases, sample reliability decreases, and vice versa.

'Sample size must take into account the degree of diversity in the population on key variables, the level of sampling error that is tolerated and the reliability required of the sample' (de Vaus 2002: 81). To reiterate, a small increase in the size of small samples can lead to a substantial increase in accuracy, but this is not the case for larger samples.

It will be clear from this discussion that a decision on sample size is rather complex. The best a researcher can do is to be aware of the effects of accuracy requirements, population characteristics, levels of measurement and the types of analysis to be used. The latter consideration reinforces the need to include in any research design decisions about how the data will be analysed. It is easy to think that this can be put off until later, but it cannot. Failure to make the decision is likely to lead to samples that are the wrong size, to data that cannot be sensibly analysed and, hence, to research questions that cannot be answered properly. In some studies, it is not possible to know in advance how the population is distributed on the characteristics being studied. Even rough estimates may be impossible to make. In this case, the researcher must be conservative and use a sample that will cope with the worst possible situation, which means making it larger!

Having said all this, one other major consideration enters into the equation. It is the practical issue of resources. The ideal sample needed to answer a set of research questions may be beyond the scope of the available resources. *Sample size decisions are always a compromise between the ideal and the practical, between the size needed to meet technical requirements and the size that can be achieved*

*with the available resources.* In the end, the researcher must be able to defend the decision as being appropriate for answering the research questions, given the particular conditions. If resources require that the sample size be reduced beyond minimum practical limits, then the design of the study would need to be radically changed, or the project postponed until sufficient resources are available.

It is always important to discover what conventions are used for your kind of research in your discipline or sub-discipline, in your university or research organization, and what the consumers of the research, including thesis examiners, find acceptable. These conventions do not always fit well with the technical requirements, but, in the end, may be politically more important.

### *Non-probability sample size*

As it is not possible to estimate population parameters from the data acquired using a non-probability sample, the discussion in the previous section on confidence levels and acceptable errors in estimates is not relevant. If, however, quantitative analysis is to be undertaken, then sample size *will* be influenced by the requirements of the type of analysis to be undertaken.

When a research project involves the use of time-intensive, in-depth methods, particularly when directed towards theory development, the issue of sample size takes on a very different complexion. As we saw in the case of *theoretical sampling*, sampling decisions evolve along with the theory. It is not possible to determine in advance what the size should be. However, time and resource limitations will inevitably put some restrictions on it. In this kind of research, it may be more useful to think of selecting cases for intensive study, rather than getting distracted by sampling concerns that are irrelevant. It is to the discussion of case studies that we now turn.

## 8.7 Case Studies

Throughout the history of social research, *case studies* have been regarded in a variety of ways:

- as a type of research design;
- as involving the use of particular kinds of research methods, usually qualitative; and
- as being a method for selecting the source of data.

In the first view, case studies are included alongside surveys, experiments and ethnography/field research. As we saw in chapter 3, this way of classifying research designs is inappropriate. The second view goes back many decades to debates about the relative merits of survey research, with its statistical techniques, compared to participant observation and field research. 'Case study' was the collective term used for the latter. While the third view is probably the least common, it is the one that we want to emphasize here. 'Case study is not a methodological choice but a choice of what is to be studied' (Stake 2005: 443).

The idea of a case study has some relationship to the notions of both clinical studies in medicine and psychology, and case histories in the helping professions, such as social work. However, there are some important differences that should become evident as this discussion proceeds.

The case study has a long history. It has been used extensively in social anthropology, and it is now used in political science (e.g. policy and public administration research), sociology (e.g. community studies), management (e.g. organizational studies) and planning (e.g. research on cities, regions and neighbourhoods). Some writers have suggested that case studies are suitable for single-person research on a limited budget, and that the study of one case provides a manageable opportunity for a researcher to study one aspect of a problem in some depth within a limited time-scale (see, for example, Blaxter et al. 2002; Bell 2005). It is implied that they are appropriate for student research, particularly for postgraduate theses, and that most researchers are capable of doing a case study.

Case studies have been used for various purposes: in exploratory, descriptive and explanatory research (Yin 2003a: 3), to generate theory (Mitchell 1983; Eckstein 1992; Hammersley et al. 2000) and to initiate change (Gummesson 1991). These uses will depend on the research questions asked, and the extent to which the researcher has control over the events being studied (Yin 2003a).

### Background

In the 1920s in the United States, and before survey research gained dominance, the case study was the acceptable way of doing social research. This occurred mainly in the Chicago tradition of sociology. With the rise of survey research, and the view that it was more 'scientific', there was a period of debates about the relative merits of 'statistical methods' and 'case studies' (e.g. Burgess 1927).

Following the Second World War, debates continued between advocates of 'surveys' (and sometimes 'experiments') and 'participant observation'; essentially between supporters of quantitative and qualitative methods. However, by the 1950s, the discussion of case studies had all but disappeared from textbooks on social research methods, although considerable attention was still given to specific techniques, such as participant observation. This decline in interest was no doubt due to the increasing availability of computers to speed up quantitative analysis (Mitchell 1983).

Thirty years later, as a result of a revival of interest in qualitative methods, particularly in educational, nursing and evaluation research, and in much British sociology, the discussion of case studies re-emerged (e.g. Mitchell 1983; Yin 1984; Platt 1988). They were now identified with methods that had been commonly used in anthropological research, with techniques of data collection such as participant observation, the use of informants, unstructured interviewing, and the study of personal documents and records. The general tenor of the discussion of such qualitative methods was that they were inferior to quantitative methods; that they were only really useful in exploratory stages of research, and

that researchers were unfortunate if they had no alternative but to use them as major methods (Goode and Hatt 1952; Platt 1988). Any research using qualitative methods and data was assumed to be a case study. However, *the case study is not one or a number of specific techniques.*

### Definitions

The literature abounds with various definitions of 'case study'. It is still common to use the concept as 'an umbrella term for a family of research methods having in common the decision to focus on inquiry around an instance' (Adelman et al. 1977). While some methods may be used frequently, such as observation and interviewing, any method is regarded as being legitimate. Even a survey may be used in a case study. What, then, is distinctive about case studies?

Goode and Hatt (1952) focused their definition on the notion of a social unit and the manner in which it is studied. A *social* unit is a 'real' individual, social event or group of people; and the individual, group or event are treated as a whole. What this means in practice is that the case study attempts 'to keep together, as a unit, those characteristics which are relevant to the scientific problem being investigated' (Goode and Hatt 1952: 333). This is in contrast to survey research that deals with individuals, but only as a collection of traits or variables.

Goode and Hatt avoided identifying the case study with a particular technique for collecting data. Rather, they regarded it as 'a *mode of organizing data* in terms of some chosen unit, such as the individual life history, the history of a group, or some delimited social process' (1952: 339).

While being sympathetic with this emphasis on the unitary character of a case, Mitchell (1983) was critical of it on two counts. First, it gives no place for extrapolation from case studies, leaving this activity to studies using statistical analysis. Second, the stress on the whole appears to ignore the context in which the case is located. In other words, Mitchell objected to case studies being limited to social descriptions in social isolation.

The issue of what constitutes the unit of analysis in a case study has been dealt with in some detail by Yin (2003a). He has suggested that the key to resolving the problem is the way research questions are stated. For example, a research question might ask: 'What internal and external changes are related to changes in religious practices and orientation to the world adopted by new religious movements?' In this example, the emphasis is on new religious movements, one or more of which could be treated as case studies. If, however, the research question was 'What kinds of religious movements arise in times of rapid social change?', then one or more rapidly changing societies could be selected as case studies.

Yin defined a *case study* as an empirical inquiry that:

- investigates a contemporary phenomenon within its real-life context, especially when
- boundaries between phenomenon and context are not clearly evident.

In addition, a case study inquiry:

- copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result
- relies on multiple sources of evidence, with data needing to converge in a triangulation fashion, and as another result
- benefits from the prior development of theoretical propositions to guide that data collection and analysis (Yin 2003a: 13–14).

Yin argued that there is no necessary connection between case studies and qualitative research: 'case studies can include, and even be limited to, quantitative evidence' (Yin 2003a: 14).

Hammersley (1992) provided a much more limited definition of the case study. Like Yin, he contrasted case studies with experiments and surveys and confined the comparison to the manner in which each one selects its units for study. The case study is viewed as just one *method of selection* and is not associated with certain kinds of methods of data collection, or a particular logic of inquiry. What distinguishes case studies from both experiments and surveys is that they use a comparatively small number of units in naturally occurring settings, and these units are investigated in considerable depth (Hammersley 1992: 185; Hammersley et al. 2000: 3).

### Types and Uses

A number of writers have provided classifications of the uses of case studies. For Gluckman (1961), the case study has had three main uses: as *apt illustrations*; as *social situations*; and as *extended case studies*. The *apt illustration* is a description of an event in which some general principle is in operation; the analysis of a *social situation* involves a more complex collection of connected events that occur in a limited time span, and which demonstrate the operation of general principles of social organization; and the *extended case study* follows the same social actors as they move through a series of linked situations and events.

While Gluckman's classification was based mainly on differing degrees of complexity, and to a lesser extent on duration of time, Eckstein's (1975, 1992) five-way classification focused on the different uses of case studies in theory development: *configurative-ideographic* studies, *disciplined-comparative* studies, *heuristic* case studies, *plausibility probes* and *crucial-case* studies.

Yin (2003a: 39–45) included both single and multiple cases within case study research. He regarded single-case studies as being of three types: *critical*, *extreme* and *revelatory*. In all three uses, the single-case study will be a complete study. However, Yin divided them into two types: *holistic* and *embedded*. The *holistic* case study has only one unit of analysis while the *embedded* case study may have a sub-unit or a number of sub-units. In contrast to single-case studies, Yin argued that the main use of multiple-case designs is analogous to conducting a series of experiments. He claimed that a well-developed theory can be tested by carefully selecting a series of cases in the same way as theories are tested experimentally.

The use of multiple-case studies has both advantages and disadvantages compared to the use of single-case studies. According to Yin, all three types of single-case studies – *critical*, *extreme* or *revelatory* – serve useful purposes on their own. While the use of a number of cases may add greater weight to a study, and make its findings more convincing, their use is only appropriate when replication rather than sampling logic is used.

Stake (2005) proposed another three-way classification: *intrinsic*, *instrumental* and *collective*. *Intrinsic* case studies are used to achieve a better understanding of a particular case. *Instrumental* case studies are selected to provide insight into an issue or to obtain a better understanding of something else, perhaps to support a developing generalization or theory. *Collective* case studies are also instrumental but involve the joint study of a number of cases that represent some phenomenon, population or general condition. The aim is likely to be theory generation, which it is hoped will apply to an even wider collection of such cases.

We can see from these various attempts to categorize case studies that their use ranges from descriptive, which usually involves single cases, to explanatory, which normally requires multiple cases. Hence, the type of research question(s) being investigated, and the purposes of the research, will determine what use could be made of case studies and, hence, their role in data selection.

### Criticisms

There have been three major criticisms that have arisen from the comparison of case studies with quantitative methods (Yin 2003a: 10–11). Researchers who favour the use of probability samples, and/or who are concerned with theory testing, tend to regard the use of case studies as unsatisfactory for these purposes. This criticism is based on the mistaken belief that only numbers can be used to describe and explain social life validly and reliably. Part of this prejudice is that qualitative research, unlike quantitative research, cannot be replicated because there is too much scope for the researcher to influence the results. These issues are taken up in chapter 9.<sup>14</sup>

The second concern is that case studies are not useful for generalizing. There are two aspects to this position: that it is not possible to generalize from a single case; and that if a number of cases are used for the purpose, it is extremely difficult to establish their comparability. Each case has too many unique aspects. However, the same criticisms could be raised about a single experiment, or the study of a single population, whether or not a sample has been used. We will return to this issue shortly.

A third criticism of the case study is not so much methodological as practical. The complaint is that case studies take too long and produce unmanageable amounts of data. Again, this criticism confuses the case study with specific methods of data collection that are time-consuming, for example, participant observation in particular and ethnography in general. Yin has argued that case studies need not take a long time and that they can now be conducted in a manageable way (2003a: 11).<sup>15</sup>

## Two Issues

Both critics and advocates of the use of case studies have focused on two issues. The first concerns the possibility or even the desirability of trying to generalize findings from case study research. The second has to do with the use of case studies in theory development and/or testing. As we have seen, Mitchell, Eckstein and Yin have all raised the issue of whether it is possible to generalize or theorize from case studies. Must the findings from a case study remain just interesting description, or can they be used to generalize and to generate and test theory?

### Generalizing

While case studies cannot be used in the way that population parameters are statistically generalized from sample statistics, is there some way of generalizing from them?

A number of strategies have been proposed, all of which require a different conception of generalizability; this involves judgement rather than probability sampling techniques. However, methods of selection are still a key element. These are: selecting 'typical' cases; studying multiple cases at different sites; the use of 'natural generalization'; the use of 'analytic generalization'; and the use of 'thick' description to facilitate 'transferability' or 'reliability'.

Researchers may feel more comfortable generalizing if they work with 'typical' cases; that is, if the case being studied can be shown to be similar to other cases in terms of relevant characteristics (see, for example, Goetz and Le Compte 1984; Whyte 1984). However, it may be difficult to demonstrate this. Certainly, this was not the guiding principle when anthropologists selected small-scale societies to study. Some writers, like Mitchell (1983: 204), have argued against trying to find typical cases, but he was more concerned with theorizing than generalizing. As we shall see shortly, when generalizing is the primary concern, finding typical cases may be useful.

A second possibility is to select a number of cases across different sites (see the review by Firestone and Herriott 1984). While such a strategy can certainly strengthen the basis for generalizing, it has some obvious practical disadvantages. As it is very expensive, each site may not be studied in the depth usually associated with case study research (Schofield 1993).

The third possibility is 'natural generalizability'. Stake (1978) pointed out that natural generalization is something we routinely do in everyday life; we can recognize similarities between objects and issues and identify repetitive patterns. Case studies add to existing experience and understanding. Lincoln and Guba (1985) – using 'transferability' or 'fittingness' – and Bassey (1981) – using 'reliability' – have taken this idea further. According to Lincoln and Guba (1985: 124–5), transferability between contexts is possible if they can be judged to be similar. Fittingness is the degree of congruence between the context in which the research was conducted and the one to which the findings are to be transferred. This requires information about both contexts.

In the context of educational research, Bassey argued that 'an important cri-

terion for judging the merit of a case study is the extent to which the details are sufficient and appropriate for a teacher working in a similar situation to relate his [*sic*] decision making to that described in the case study. The reliability of a case is more important than its generalisability' (1981: 85). Hence, if researchers provide sufficiently detailed descriptions of their cases, it is possible for others to make judgements about whether the findings can be related. The burden of proof is on the user rather than the originator of the research.

Yin argued for the use of 'analytic generalization' as an alternative to 'statistical generalization' in case studies. Case studies are not sampling units and should not be selected as such. He argued that cases can be compared with a previously developed theory and a number of cases can help to support a theory against a rival. Theory can be used to make links between case studies. This kind of 'analytical generalization' can be used whether your case study involves one or several cases' (Yin 2003a: 32–3).

Gomm et al. (2000) adopted a different approach that returns to more traditional ideas on generalizing. They went back to the issue of typical cases and have suggested that researchers should collect information about the wider population to which the case study findings might be generalized. This is particularly necessary where the population is heterogeneous. Selecting a case that includes this degree of heterogeneity would be important. Another suggestion is to select a small sample of cases that cover the range of diversity in a population. Such a selection would be assisted if a survey of the population, or perhaps a probability sample drawn from it, preceded the case selection in order to give an idea of the kind of diversity that needs to be taken into account.

Hence, *generalizing is done by making judgements on the basis of knowledge of the characteristics of the case and the target population*. This process is no different from the situation where survey researchers want to generalize from a population they may have studied, with or without probability sampling, to a wider population or other populations. For example, the findings of a survey of students at one university could be generalized to students at another university only if critical characteristics of students in both universities are similar. This kind of generalization is a matter of judgement based on evidence. We shall come back to this issue in the next chapter in the context of generalizing in qualitative research.

### Theorizing

We have already noted that Eckstein (1975, 1992) proposed a five-way classification of case studies according to their role in theory development. These range from developing understanding through to theory building and theory testing. *Configurative-ideographic* studies use descriptions to provide understanding; *disciplined-comparative* studies require that each case be viewed in the context of an established or at least a provisional theory; *heuristic* case studies are deliberately used to stimulate theoretical thinking; *plausibility probes* are used in the intermediate stage between the development of a theory, whether by heuristic case studies or some other means, and the testing of that theory; and *crucial-case* studies are similar to crucial experiments when they are designed to challenge an

existing theory.

In the context of anthropological research, Mitchell championed the view that case studies are not simply descriptions of the life of the people; they have a theoretical aim of searching for general theoretical principles (1983: 192). In this way, anthropological theory has been built up from many case studies over many years.

Issues associated with the selection of case studies are different when theorizing rather than generalizing is the concern. As both Mitchell and Yin have argued, a different kind of logic is required to test a theory with case studies, such as, 'logical inference' (Mitchell); 'analytic generalization' (Yin); and 'analytic induction' (Lindesmith).

For Mitchell, *logical inference* 'is the process by which the analyst draws conclusions about the essential linkage between two or more characteristics in terms of some systematic explanatory schema – some set of theoretical propositions' (1983: 199–200). Yin (2003a) adopted a similar approach in his proposal for the use of *analytic generalization* as a way of linking cases. He also saw this as a means of theory testing, although in the form of accumulated support rather than critical testing. Platt (1988) also adopted the view that prior theory is necessary. She argued that a case study needs to be located in the context of relevant knowledge and appropriate theory. To strengthen their theoretical role, case studies need to be specifically designed rather than being chosen as a matter of convenience or by accident. Both Yin and Platt proposed the use of Inductive rather than Deductive logic.

It is worth noting that whereas typical cases may be useful for generalizing, Eckstein (1975), Mitchell (1983), Platt (1988) and Yin (2003a) have argued for the use of extreme, deviant or least likely cases in theory testing. If a general theoretical principle can be shown to hold in these types of cases, the degree of corroboration is stronger than in cases that might be regarded as typical.

This leads us to a form of logic known as *analytic induction*, which shares much in common with the trial and error process associated with Deductive logic. It was used by Znaniecki (1934), Lindesmith (1937, 1968) and Cressey (1950). As expounded by Lindesmith (1968), this logic begins with an initial investigation of a few cases, which leads to the formulation of a tentative causal hypothesis to account for the phenomenon. Additional cases are investigated to test this hypothesis, and this may lead to its reformulation. Further cases are examined until revisions of the hypothesis are no longer required. However, the investigation of other cases at a later time, or in other places, may very well lead to further revisions. Burgess (1984: 179) has elaborated this logic as follows.

1. Define the phenomenon to be explained.
2. Formulate a hypothetical explanation of the phenomenon.
3. Study one case to see whether the hypothesis relates to that particular case.
4. If the hypothesis does not fit the case, the phenomenon is either reformulated or redefined in order to exclude the particular case.
5. Practical certainty is achieved with a small number of cases, but negative cases disprove the explanation and require a reformulation.
6. The examination of cases, redefinition of the phenomenon and reformulation of hypotheses, is continued until a universal relationship is established.<sup>16</sup>

A major difference between *analytic induction* and Deductive logic is that deductively derived theories have a more tentative status than is implied in the last step above. It is on this point that Robinson (1951) and Turner (1953) have been critical of analytic induction.

While supporting the use of the comparative method, Hammersley et al. (2000) have been critical of Lindesmith's formulation of it. Amongst other things, they raise the issue of just how hypotheses are tested in Lindesmith's method. There is a danger that additional cases may be selected to provide confirmation rather than severe testing. They also raise some practical difficulties: the problem of selecting cases that provide good tests of a hypothesis; and the large number of cases that may be required. They have suggested that the second point can be ameliorated if researchers work cumulatively by building on each other's work.

Hammersley et al. (2000) make an important distinction between case studies that are used to describe and possibly explain features of a case or a set of cases, and those used for theory development and testing. This brings us back to the issue of generalization. Where the focus is only on the case or set of cases themselves, then generalizing is not the issue. However, in theory development and testing, generalizing is involved. The minimum aim is to develop theories that apply across many situations. *We should entertain research questions that require us to strive beyond the particulars of single cases to explanations that have some wider application.* Hammersley et al. appear to concur with this view when they suggest that pragmatic considerations, such as seeking solutions to practical problems, rather than developing coherent theoretical systems, should be given precedence.

## 8.8 Conclusion

Textbooks on social research methods usually concentrate on the tasks of obtaining and analysing data. No doubt these are central tasks, and learning how to do them, or at least enough of them to conduct the research at hand, is absolutely essential. However, a major theme of this book is that decisions about these matters can only occur after a wide range of other decisions have been made. This is what the previous chapters have been about.

This chapter has dealt with some decisions that anticipate the next chapter on methods of data collection, generation and analysis. Consideration has to be given to what will be the best *type* and *form* of data to answer the research questions; for example, will it be possible to obtain 'new' primary data, or will it be appropriate or necessary to use data from previous research? Also, will data in numerical (quantitative) form be the most appropriate, or will non-numerical (qualitative) data be relevant? Then there are decisions to be made about the *sources* of these data. Will they need to come from micro, meso or macro settings; from individuals, social groups, social collectivities or social artefacts?

Once the decision about the research site or context has been made there is