PART I

WHAT IS RESEARCH DESIGN?

1

THE CONTEXT OF DESIGN

Before examining types of research designs it is important to be clear about the role and purpose of research design. We need to understand what research design is and what it is not. We need to know where design fits into the whole research process from framing a question to finally analysing and reporting data. This is the purpose of this chapter.

Description and explanation

Social researchers ask two fundamental types of research questions:

1 *What* is going on (descriptive research)?
2 *Why* is it going on (explanatory research)?

Descriptive research

Although some people dismiss descriptive research as 'mere description', good description is fundamental to the research enterprise and it has added immeasurably to our knowledge of the shape and nature of our society. Descriptive research encompasses much government sponsored research including the population census, the collection of a wide range of social indicators and economic information such as household expenditure patterns, time use studies, employment and crime statistics and the like.

Descriptions can be concrete or abstract. A relatively concrete description might describe the ethnic mix of a community, the changing age profile of a population or the gender mix of a workplace. Alternatively

the description might ask more abstract questions such as 'Is the level of social inequality increasing or declining?', 'How secular is society?' or 'How much poverty is there in this community?'

Accurate descriptions of the level of unemployment or poverty have historically played a key role in social policy reforms (Marsh, 1982). By demonstrating the existence of social problems, competent description can challenge accepted assumptions about the way things are and can provoke action.

Good description provokes the 'why' questions of explanatory research. If we detect greater social polarization over the last 20 years (i.e. the rich are getting richer and the poor are getting poorer) we are forced to ask 'Why is this happening?' But before asking 'why?' we must be sure about the fact and dimensions of the phenomenon of increasing polarization. It is all very well to develop elaborate theories as to why society might be more polarized now than in the recent past, but if the basic premise is wrong (i.e. society is not becoming more polarized) then attempts to explain a non-existent phenomenon are silly.

Of course description can degenerate to mindless fact gathering or what C.W. Mills (1959) called 'abstracted empiricism'. There are plenty of examples of unfocused surveys and case studies that report trivial information and fail to provoke any 'why' questions or provide any basis for generalization. However, this is a function of inconsequential descriptions rather than an indictment of descriptive research itself.

Explanatory research

Explanatory research focuses on *why* questions. For example, it is one thing to describe the crime rate in a country, to examine trends over time or to compare the rates in different countries. It is quite a different thing to develop explanations about why the crime rate is as high as it is, why some types of crime are increasing or why the rate is higher in some countries than in others.

The way in which researchers develop research designs is fundamentally affected by whether the research question is descriptive or explanatory. It affects what information is collected. For example, if we want to explain why some people are more likely to be apprehended and convicted of crimes we need to have hunches about why this is so. We may have many possibly incompatible hunches and will need to collect information that enables us to see which hunches work best empirically.

Answering the 'why' questions involves developing *causal* explanations. Causal explanations argue that phenomenon Y (e.g. income level) is affected by factor X (e.g. gender). Some causal explanations will be simple while others will be more complex. For example, we might argue that there is a *direct* effect of gender on income (i.e. simple gender discrimination) (Figure 1.1a). We might argue for a causal chain, such as that gender affects choice of field of training which in turn affects

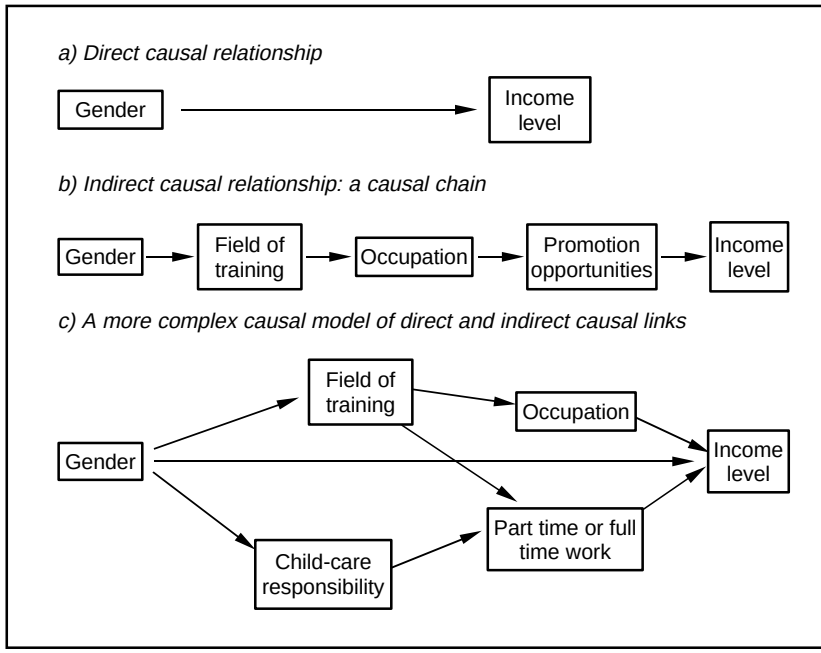

Figure 1.1 *Three types of causal relationships*

occupational options, which are linked to opportunities for promotion, which in turn affect income level (Figure 1.1b). Or we could posit a more complex model involving a number of interrelated causal chains (Figure 1.1c).

Prediction, correlation and causation

People often confuse correlation with causation. Simply because one event follows another, or two factors co-vary, does not mean that one causes the other. The link between two events may be coincidental rather than causal.

There is a correlation between the number of fire engines at a fire and the amount of damage caused by the fire (the more fire engines the more damage). Is it therefore reasonable to conclude that the number of fire engines causes the amount of damage? Clearly the number of fire engines and the amount of damage will *both* be due to some third factor – such as the seriousness of the fire.

Similarly, as the divorce rate changed over the twentieth century the crime rate increased a few years later. But this does not mean that divorce causes crime. Rather than divorce causing crime, divorce and crime rates might both be due to *other* social processes such as secularization, greater individualism or poverty.

Students at fee paying private schools typically perform better in their final year of schooling than those at government funded schools. But this need not be because private schools *produce* better performance. It may be that attending a private school and better final-year performance are *both* the outcome of some other cause (see later discussion).

Confusing causation with correlation also confuses prediction with causation and prediction with explanation. Where two events or characteristics are correlated we can predict one from the other. Knowing the type of school attended improves our capacity to predict academic achievement. But this does not mean that the school type affects academic achievement. Predicting performance on the basis of school type does not tell us *why* private school students do better. *Good prediction does not depend on causal relationships. Nor does the ability to predict accurately demonstrate anything about causality*.

Recognizing that causation is more than correlation highlights a problem. While we can observe correlation *we cannot observe cause*. We have to *infer* cause. These inferences however are 'necessarily fallible . . . [they] are only indirectly linked to observables' (Cook and Campbell, 1979: 10). Because our inferences are fallible we must minimize the chances of incorrectly saying that a relationship is causal when in fact it is not. *One of the fundamental purposes of research design in explanatory research is to avoid invalid inferences*.

Deterministic and probabilistic concepts of causation

There are two ways of thinking about causes: deterministically and probabilistically. The smoker who denies that tobacco causes cancer because he smokes heavily but has not contracted cancer illustrates deterministic causation. Probabilistic causation is illustrated by health authorities who point to the increased chances of cancer among smokers.

Deterministic causation is where variable *X* is said to cause *Y* if, and only if, *X invariably* produces *Y*. That is, when *X* is present then *Y* will 'necessarily, inevitably and infallibly' occur (Cook and Campbell, 1979: 14). This approach seeks to establish causal *laws* such as: whenever water is heated to 100 °C it always boils.

In reality laws are never this simple. They will always specify particular *conditions* under which that law operates. Indeed a great deal of scientific investigation involves specifying the conditions under which particular laws operate. Thus, we might say that *at sea level* heating *pure* water to 100 °C will always cause water to boil.

Alternatively, the law might be stated in the form of 'other things being equal' then *X* will always produce *Y*. A deterministic version of the relationship between race and income level would say that other things being equal (age, education, personality, experience etc.) then a white person will [always] earn a higher income than a black person. That is, race (*X*) causes income level (*Y*).

Stated like this the notion of deterministic causation in the social sciences sounds odd. It is hard to conceive of a characteristic or event that will invariably result in a given outcome even if a fairly tight set of conditions is specified. The *complexity* of human social behaviour and the *subjective, meaningful and voluntaristic* components of human behaviour mean that it will never be possible to arrive at causal statements of the type 'If *X*, and *A* and *B*, then *Y* will always follow.'

Most causal thinking in the social sciences is *probabilistic* rather than *deterministic* (Suppes, 1970). That is, we work at the level that a given factor increases (or decreases) the probability of a particular outcome, for example: being female increases the probability of working part time; race affects the probability of having a high status job.

We can improve probabilistic explanations by specifying conditions under which *X* is less likely and more likely to affect *Y*. But we will never achieve complete or deterministic explanations. Human behaviour is both *willed* and *caused*: there is a double-sided character to human social behaviour. People *construct* their social world and there are creative aspects to human action but this freedom and agency will always be constrained by the structures within which people live. Because behaviour is not simply determined we cannot achieve deterministic explanations. However, because behaviour is constrained we can achieve probabilistic explanations. We can say that a given factor will increase the likelihood of a given outcome but there will never be certainty about outcomes.

Despite the probabilistic nature of causal statements in the social sciences, much popular, ideological and political discourse translates these into deterministic statements. Findings about the causal effects of class, gender or ethnicity, for example, are often read as if these factors invariably and completely produce particular outcomes. One could be forgiven for thinking that social science has demonstrated that gender completely and invariably determines position in society, roles in families, values and ways of relating to other people.

Theory testing and theory construction

Attempts to answer the 'why' questions in social science are theories. These theories vary in their *complexity* (how many variables and links), *abstraction* and *scope*. To understand the role of theory in empirical research it is useful to distinguish between two different styles of research: theory testing and theory building (Figure 1.2).

Theory building

Theory building is a process in which research begins with observations and uses *inductive* reasoning to derive a theory from these observations.

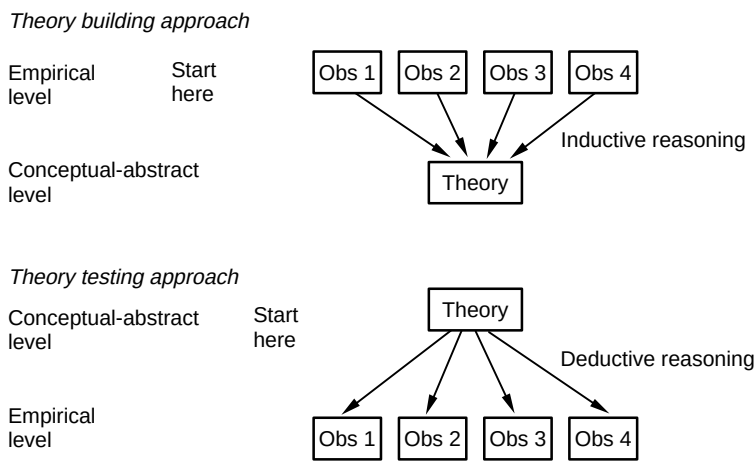

Figure 1.2 *Theory building and theory testing approaches to research*

These theories attempt to make sense of observations. Because the theory is produced *after* observations are made it is often called *post factum* theory (Merton, 1968) or *ex post facto* theorizing.

This form of theory building entails asking whether the observation is *a particular case of a more general factor*, or how the observation *fits into a pattern or a story*. For example, Durkheim observed that the suicide rate was higher among Protestants than Catholics. But is religious affiliation a particular case of something more general? Of what more general phenomenon might it be an indicator? Are there other observations that shed light on this? He also observed that men were more suicidal than women, urban dwellers more than rural dwellers and the socially mobile more than the socially stable. He argued that the common factor behind all these observations was that those groups who were most suicidal were also less well socially integrated and experienced greater ambiguity about how to behave and what is right and wrong. He theorized that one of the explanations for suicidal behaviour was a sense of normlessness – a disconnectedness of individuals from their social world. Of course, there may have been other ways of accounting for these observations but at least Durkheim's explanation was consistent with the facts.

Theory testing

In contrast, a theory testing approach *begins* with a theory and uses theory to guide which observations to make: it moves from the general to the particular. The observations should provide a test of the worth of the theory. Using *deductive* reasoning to derive a set of propositions from the theory does this. We need to develop these propositions so that

		Parents divorced?	
		No	Yes
Parental conflict	Low	(a)	(b)
	High	(c)	(d)

Figure 1.3 *The relationship between divorce and parental conflict*

if the theory is true then certain things should follow in the real world. We then assess whether these predictions are correct. If they are correct the theory is supported. If they do not hold up then the theory needs to be either rejected or modified.

For example, we may wish to test the theory that it is not divorce itself that affects the wellbeing of children but the level of conflict between parents. To test this idea we can make predictions about the wellbeing of children under different family conditions. For the simple theory that it is parental conflict rather than divorce that affects a child's wellbeing there are four basic 'conditions' (see Figure 1.3). For each 'condition' the theory would make different predictions about the level of children's wellbeing that we can examine.

If the theory that it is parental conflict rather than parental divorce is correct the following propositions should be supported:

- *Proposition 1: children in situations (a) and (b) would be equally well off* That is, where parental conflict is low, children with divorced parents will do just as well as those whose parents are married.
- *Proposition 2: children in situations (c) and (d) should be equally poorly off* That is, children in conflictual couple families will do just as badly as children in post-divorce families where parents sustain high conflict.
- *Proposition 3: children in situation (c) will do worse than those in situation (a)* That is, those with married parents in high conflict will do worse than those who have married parents who are not in conflict.
- *Proposition 4: children in situation (d) will do worse than those in situation (b)* That is, those with divorced parents in high conflict will do worse than those who have divorced parents who are not in conflict.
- *Proposition 5: children in situation (b) will do better than those in situation (c)* That is, children with divorced parents who are not in conflict will do better than those with married parents who are in conflict.
- *Proposition 6: children in situation (a) will do better than those in situation (d)* That is, children with married parents who are not in conflict will do better than those with divorced parents who are in conflict.

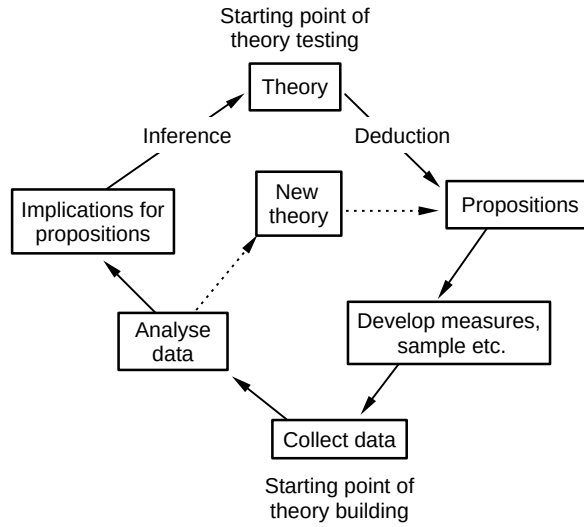

Figure 1.4 *The logic of the research process*

No single proposition would provide a compelling test of the original theory. Indeed, taken on its own proposition 3, for example, would reveal nothing about the impact of divorce. However, taken as a package, the *set* of propositions provides a stronger test of the theory than any single proposition.

Although theory testing and theory building are often presented as alternative modes of research they should be part of one ongoing process (Figure 1.4). Typically, theory building will produce a plausible account or explanation of a set of observations. However, such explanations are frequently just one of a number of possible explanations that fit the data. While plausible they are not necessarily compelling. They require systematic testing where data are collected to specifically evaluate how well the explanation holds when subjected to a range of crucial tests.

What is research design?

How is the term 'research design' to be used in this book? An analogy might help. When constructing a building there is no point ordering materials or setting critical dates for completion of project stages until we know what sort of building is being constructed. The first decision is whether we need a high rise office building, a factory for manufacturing machinery, a school, a residential home or an apartment block. Until this is done we cannot sketch a plan, obtain permits, work out a work schedule or order materials.

Similarly, social research needs a design or a structure before data collection or analysis can commence. A research design is *not* just a work plan. A work plan details what has to be done to complete the project but the work plan will flow from the project's research design. *The function of a research design is to ensure that the evidence obtained enables us to answer the initial question as unambiguously as possible*. Obtaining relevant evidence entails specifying the type of evidence needed to answer the research question, to test a theory, to evaluate a programme or to accurately describe some phenomenon. In other words, when designing research we need to ask: given this research question (or theory), what type of evidence is needed to answer the question (or test the theory) *in a convincing way*?

Research design 'deals with a *logical* problem and not a *logistical* problem' (Yin, 1989: 29). Before a builder or architect can develop a work plan or order materials they must first establish the type of building required, its uses and the needs of the occupants. The work plan flows from this. Similarly, in social research the issues of sampling, method of data collection (e.g. questionnaire, observation, document analysis), design of questions are all subsidiary to the matter of 'What evidence do I need to collect?'

Too often researchers design questionnaires or begin interviewing far too early – before thinking through what information they require to answer their research questions. Without attending to these research design matters at the beginning, the conclusions drawn will normally be weak and unconvincing and fail to answer the research question.

Design versus method

Research design is different from the method by which data are collected. Many research methods texts confuse research designs with methods. It is not uncommon to see research design treated as a mode of data collection rather than as a logical structure of the inquiry. But there is nothing intrinsic about any research design that requires a particular method of data collection. Although cross-sectional surveys are frequently equated with questionnaires and case studies are often equated with participant observation (e.g. Whyte's *Street Corner Society*, 1943), data for any design can be collected with any data collection method (Figure 1.5). How the data are collected is irrelevant to the *logic* of the design.

Failing to distinguish between design and method leads to poor evaluation of designs. Equating cross-sectional designs with questionnaires, or case studies with participant observation, means that the designs are often evaluated against the strengths and weaknesses of the method rather than their ability to draw relatively unambiguous conclusions or to select between rival plausible hypotheses.

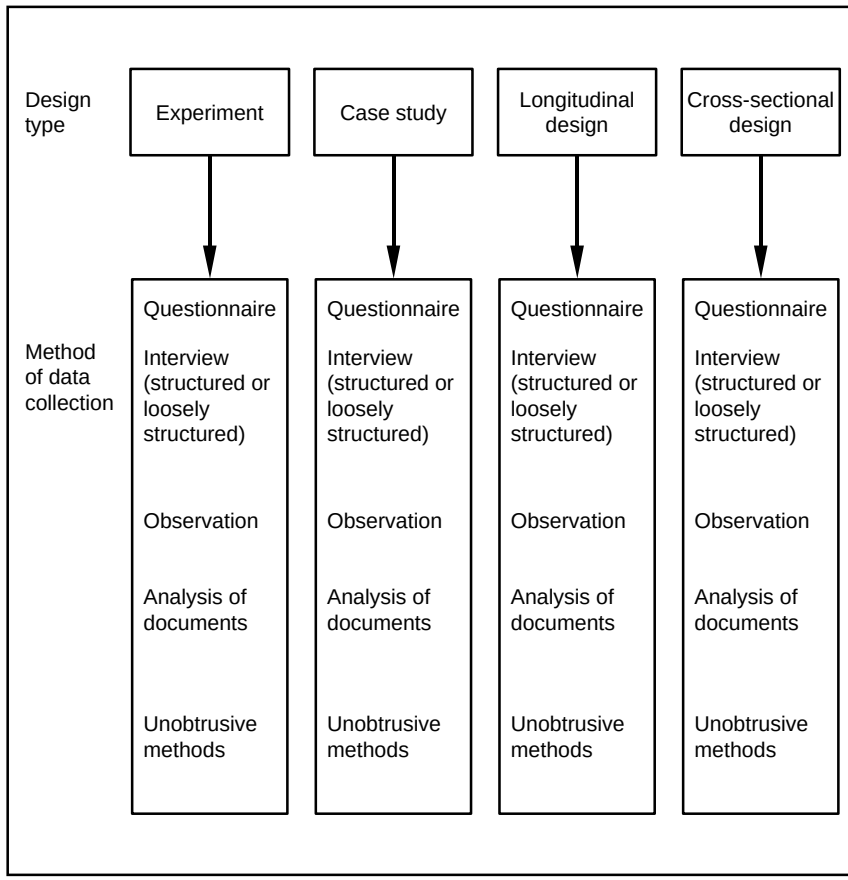

Figure 1.5 *Relationship between research design and particular data collection methods*

Quantitative and qualitative research

Similarly, designs are often equated with qualitative and quantitative research methods. Social surveys and experiments are frequently viewed as prime examples of quantitative research and are evaluated against the strengths and weaknesses of statistical, quantitative research methods and analysis. Case studies, on the other hand, are often seen as prime examples of qualitative research – which adopts an interpretive approach to data, studies 'things' within their context and considers the subjective meanings that people bring to their situation.

It is erroneous to equate a particular research design with either quantitative or qualitative methods. Yin (1993), a respected authority on case study design, has stressed the irrelevance of the quantitative/qualitative distinction for case studies. He points out that:

> a point of confusion . . . has been the unfortunate linking between the case study method and certain types of data collection – for example those focusing on qualitative methods, ethnography, or participant observation. People have thought that the case study method required them to embrace these data collection methods . . . On the contrary, the method does not imply any particular form of data collection – which can be qualitative or quantitative. (1993: 32)

Similarly, Marsh (1982) argues that quantitative surveys can provide information and explanations that are 'adequate at the level of meaning'. While recognizing that survey research has not always been good at tapping the subjective dimension of behaviour, she argues that:

> Making sense of social action . . . is . . . hard and surveys have not traditionally been very good at it. The earliest survey researchers started a tradition . . . of bringing the meaning from outside, either by making use of the researcher's stock of plausible explanations . . . or by bringing it from subsidiary in-depth interviews sprinkling quotes . . . liberally on the raw correlations derived from the survey. Survey research became much more exciting . . . when it began including meaningful dimensions in the study design. [This has been done in] two ways, firstly [by] asking the actor either for her reasons directly, or to supply information about the central values in her life around which we may assume she is orienting her life. [This] involves collecting a sufficiently complete picture of the context in which an actor finds herself that a team of outsiders may read off the meaningful dimensions. (1982: 123–4)

Adopting a sceptical approach to explanations

The need for research design stems from a sceptical approach to research and a view that scientific knowledge must always be provisional. The purpose of research design is to reduce the ambiguity of much research evidence.

We can always find some evidence consistent with almost any theory. However, we should be sceptical of the evidence, and rather than seeking evidence that is *consistent* with our theory we should seek evidence that provides a *compelling* test of the theory.

There are two related strategies for doing this: eliminating rival explanations of the evidence and deliberately seeking evidence that could *disprove* the theory.

Plausible rival hypotheses

A fundamental strategy of social research involves evaluating 'plausible rival hypotheses'. We need to examine and evaluate alternative ways of explaining a particular phenomenon. This applies regardless of whether the data are quantitative or qualitative; regardless of the particular research design (experimental, cross-sectional, longitudinal or case

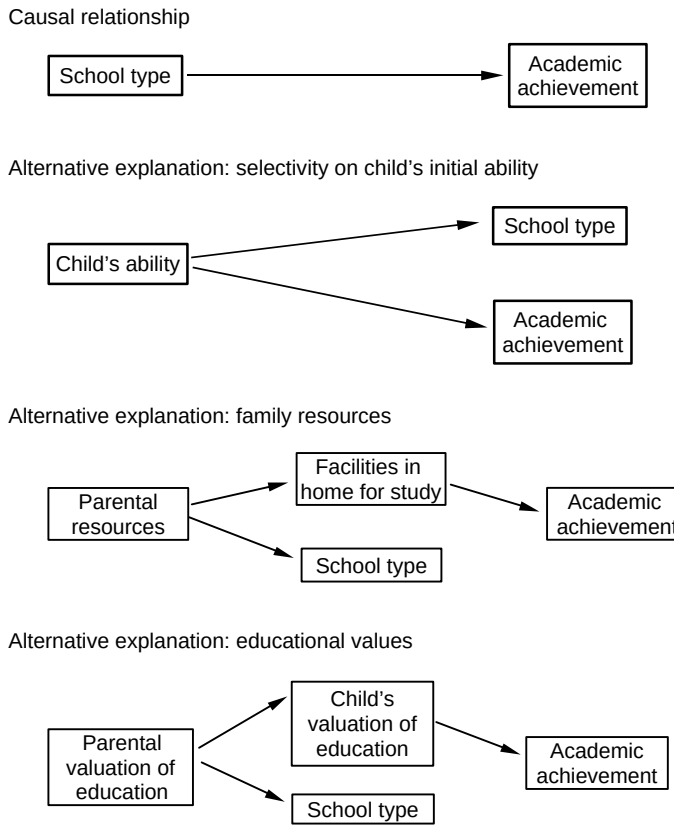

Figure 1.6 *Causal and non-causal explanations of the relationship between school type and academic achievement*

study); and regardless of the method of data collection (e.g. observation, questionnaire). Our mindset needs to anticipate alternative ways of interpreting findings and to regard any interpretation of these findings as provisional – subject to further testing.

The idea of evaluating plausible rival hypotheses can be illustrated using the example of the correlation between type of school attended and academic achievement. Many parents accept the causal proposition that attendance at fee paying private schools improves a child's academic performance (Figure 1.6). Schools themselves promote the same notion by prominently advertising their pass rates and comparing them with those of other schools or with national averages. By implication they propose a causal connection: 'Send your child to our school and they will pass (or get grades to gain entry into prestigious institutions, courses).' The data they provide are consistent with their proposition that these schools produce better results.

But these data are not compelling. There are at least three other ways of accounting for this correlation without accepting the causal link between school type and achievement (Figure 1.6). There is the *selectivity* explanation: the more able students may be sent to fee paying private schools in the first place. There is the *family resources* explanation: parents who can afford to send their children to fee paying private schools can also afford other help (e.g. books, private tutoring, quiet study space, computers). It is this help rather than the type of school that produces the better performance of private school students. Finally, there is the family *values* explanation: parents who value education most are prepared to send their children to fee paying private schools and it is this family emphasis on education, not the schools themselves, that produces the better academic performance. All these explanations are equally consistent with the observation that private school students do better than government school students. Without collecting further evidence we cannot choose between these explanations and therefore must remain open minded about which one makes most empirical sense.

There might also be methodological explanations for the finding that private school students perform better academically. These methodological issues might undermine any argument that a causal connection exists. Are the results due to questionable ways of measuring achievement? From what range and number of schools were the data obtained? On how many cases are the conclusions based? Could the pattern simply be a function of chance? These are all possible alternative explanations for the finding that private school students perform better.

Good research design will anticipate competing explanations *before* collecting data so that relevant information for evaluating the relative merits of these competing explanations is obtained. In this example of schools and academic achievement, thinking about alternative plausible hypotheses beforehand would lead us to find out about the parents' financial resources, the study resources available in the home, the parents' and child's attitudes about education and the child's academic abilities before entering the school.

The fallacy of affirming the consequent Although evidence may be consistent with an initial proposition it might be equally consistent with a range of alternative propositions. Too often people do not even think of the alternative hypotheses and simply conclude that since the evidence is consistent with their theory then the theory is true. This form of reasoning commits the logical *fallacy of affirming the consequent*. This form of reasoning has the following logical structure:

- If A is true then B should follow.
- We observe B.
- Therefore A is true.

If we apply this logic to the type of school and achievement proposition, the logical structure of the school type and achievement argument becomes clearer.

Initial proposition:

- Private schools produce better students than do government schools.

The test:

- *If A then B* If private schools produce better students (A) then their students should get better final marks than those from government funded schools (B).
- *B is true* Private school students do achieve better final marks than government school students (observe B).
- *Therefore A is true* Therefore private schools do produce better students (A is true).

But as I have already argued, the better performance of private school students might also reflect the effect of other factors. The problem here is that any number of explanations may be correct and the evidence does not help rule out many of these. For the social scientist this level of indeterminacy is quite unsatisfactory. In effect we are only in a position to say:

- If A [or C, or D, or E, or F, or . . .] then B.
- We observe B.
- Therefore A [or C, or D, or E, or F, or . . .] is true.

Although explanation (A) is still in the running because it is consistent with the observations, we cannot say that it is the most plausible explanation. We need to test our proposition more thoroughly by evaluating the worth of the alternative propositions.

Falsification: looking for evidence to disprove the theory

As well as evaluating and eliminating alternative explanations we should rigorously evaluate our own theories. Rather than asking 'What evidence would constitute support for the theory?', ask 'What evidence would convince me that the theory is *wrong*?' It is not difficult to find evidence consistent with a theory. It is much tougher for a theory to survive the test of people trying to disprove it.

Unfortunately some theories are closed systems in which any evidence can be interpreted as support for the theory. Such theories are said to be non-falsifiable. Many religions or belief systems can become closed systems whereby all evidence can be accommodated by the theory and

nothing will change the mind of the true believer. Exchange theory (Homans, 1961; Blau, 1964) is largely non-falsifiable. It assumes that we always maximize our gains and avoid costs. But we can see almost anything as a gain. Great sacrifices to care for a disabled relative can be interpreted as a gain (satisfaction of helping) rather than a loss (income, time for self etc.). We need to frame our propositions and define our terms in such a way that they are capable of being disproven.

The provisional nature of support for theories

Even where the theory is corroborated and has survived attempts to disprove it, the theory remains provisional:

> falsificationism stresses the ambiguity of confirmation . . . corroboration gives only the comfort that the theory has been tested and survived the test, that even after the most impressive corroborations of predictions it has only achieved the status of 'not yet disconfirmed'. This . . . is far from the status of 'being true'. (Cook and Campbell, 1979: 20)

There always may be an unthought-of explanation. We cannot anticipate or evaluate every possible explanation. The more alternative explanations that have been eliminated and the more we have tried to disprove our theory, the more confidence we will have in it, but we should avoid thinking that it is *proven*.

However we can *disprove* a theory. The logic of this is:

- If theory A is true then B should follow.
- B does *not* follow.
- Therefore A is not true.

So long as B is a valid test of A the absence of B should make us reject or revise the theory. In reality, we would not reject a theory simply because a single fact or observation does not fit. Before rejecting a plausible theory we would require multiple disconfirmations using different measures, different samples and different methods of data collection and analysis.

In summary, we should adopt a sceptical approach to explanations. We should anticipate rival interpretations and collect data to enable the winnowing out of the weaker explanations and the identification of which alternative theories make most empirical sense. We also need to ask what data would challenge the explanation and collect data to evaluate the theory from this more demanding perspective.

Summary

This chapter has outlined the purpose of research design in both descriptive and explanatory research. In explanatory research the purpose is to develop and evaluate causal theories. The probabilistic nature of causation in social sciences, as opposed to deterministic causation, was discussed.

Research design is not related to any particular method of collecting data or any particular type of data. Any research design can, in principle, use any type of data collection method and can use either quantitative or qualitative data. Research design refers to the *structure* of an enquiry: it is a logical matter rather than a logistical one.

It has been argued that the central role of research design is to minimize the chance of drawing incorrect causal inferences from data. Design is a logical task undertaken to ensure that the evidence collected enables us to answer questions or to test theories as unambiguously as possible. When designing research it is essential that we identify the type of evidence required to answer the research question in a convincing way. This means that we must not simply collect evidence that is consistent with a particular theory or explanation. Research needs to be structured in such a way that the evidence also bears on alternative rival explanations and enables us to identify which of the competing explanations is most compelling empirically. It also means that we must not simply look for evidence that supports our favourite theory: we should also look for evidence that has the potential to disprove our preferred explanations.