

3

CAUSATION AND THE LOGIC OF RESEARCH DESIGN

Establishing causal relationships is at the heart of explanatory research design. However, it is not a simple matter to establish that one event causes another (Blalock, 1964; Hage and Meeker, 1988). The main reason why it is difficult to establish causal relationships is because we cannot actually *observe* one phenomenon producing change in another. Even though one event might always follow another we do not know that this is because one event causes the other. Causal relationships must therefore be inferred rather than observed. *The purpose of research design in explanatory research is to improve the quality of our causal inferences.*

Inferring causal relationships

Criteria for inferring cause

In Chapter 1 I distinguished between probabilistic and deterministic concepts of causation. Probabilistic approaches to causation are those that argue that a given factor increases (or decreases) the probability of a particular outcome. For example we may argue that there is a causal relationship between gender and working part time – that gender affects the probability of working part time.

In order to infer that a probabilistic causal relationship exists between two variables, two basic criteria must be met. First, there must be co-variation of causal and outcome variables (e.g. between gender and being a full time or part time worker); and second, the assertion that one variable affects the other must make sense.

CO-VARIATION

If two factors are causally related they must *at least* be correlated: they must co-vary. If X causes Y then people who differ from one another on X should tend to differ from one another on Y. For example, if we were to argue that working in the private sector rather than the public sector makes people more achievement oriented at work we would, *at the very*

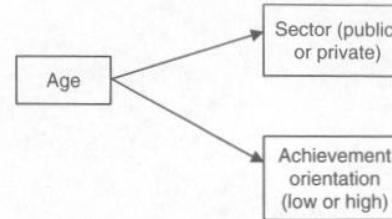


Figure 3.1 *A spurious relationship*

least, need to demonstrate that private sector workers had a higher achievement orientation than public sector workers. If two factors did not co-vary – that is public and private sector workers had identical levels of work achievement – then we would be hard pressed to argue that they are causally related.¹

However, while co-variation is a precondition it is not enough for us to assert that the variables are causally related. Correlations can also reflect non-causal relationships. When two variables or events are correlated but not causally related the relationship between the two variables is said to be spurious (see Chapter 1). For example, the correlation between sector of employment and employment orientation might be due entirely to a third variable such as age (Figure 3.1). Younger people might be more likely than older people to work in the private sector and younger people might also have a higher achievement orientation than older people. These two patterns will mean that sector and achievement orientation are correlated (more young people in the private sector, so therefore the private sector is associated with achievement orientation). However the link between sector and achievement orientation is coincidental rather than causal.

IT MUST MAKE SENSE

Any assertion that co-variation reflects a causal relationship must be plausible. It must make sense at three levels.

Time order If two variables are correlated the cause must come *before* the effect. Causal reasoning has no time for the assertion that a future event can have a present effect (teleological explanation). Our causal proposition must be such that the causal variable occurs before the presumed effect. The time gap between cause and effect can be minutes or may be years (e.g. the effect of education on income can take many years to show itself).

Even though two variables might be causally related it can sometimes be difficult to work out which variable comes first and therefore to

establish which variable is the cause and which is the effect. For example, does sector of employment affect achievement orientation or is it the other way around? Even where we assert that one variable comes first the causal relationship may be two-way. That is, sector of employment may affect achievement orientation which in turn influences future decisions about the sector of employment in which one works. Causal relationships can be reciprocal (two-way) rather than one-way.

Dependent variable must be capable of change If we say that a correlation between two variables is because one is causing the other, we must make sure that the dependent variable (the effect) is *capable* of being changed. If it cannot be changed then a causal account of the relationship makes no sense. For example, any causal relationship between sex and income could only be in the direction of sex affecting income. The opposite proposition (income→sex) makes no sense.

Theoretical plausibility The causal assertion must make sense. We should be able to tell a story of how X affects Y if we wish to infer a causal relationship between X and Y . Even if we cannot empirically demonstrate *how* X affects Y we need to provide a plausible account of the connection (plausible in terms of other research, current theory etc.). For example, to support the assertion that sector of employment affects achievement orientation we might argue that the private sector fosters the development of an achievement orientation by strategies such as paying performance bonuses, developing a culture of higher expectations, providing better resources and creating less job security. When backed up by this type of reasoning, any correlation between employment sector and achievement orientation can be plausibly interpreted in causal terms.

Types of causal patterns

DIRECT AND INDIRECT CAUSAL RELATIONSHIPS

Causal relationships can be either *direct* or *indirect*. A direct relationship is one where we assert that the cause affects the outcome directly rather than via other variables. An indirect causal relationship is one where the cause has its effect by operating via its influence on another variable that, in turn, produces the effect. The variable through which the two variables are related is called the *intervening* variable: it comes in time and in a causal sequence between the initial cause and the effect. For example, we might argue that the way the private sector produces higher achievement orientation is by making employees fear for their jobs (the intervening variable) (Figure 3.2).

Indirect causal relationships may be simple (as in Figure 3.2) or consist of an extended causal chain or a number of different causal paths (Figure 3.3).

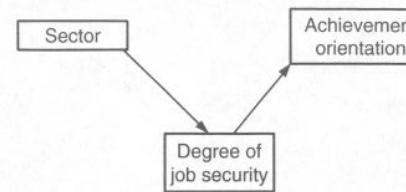


Figure 3.2 An indirect causal relationship

a) Long causal chain

$$X \rightarrow Z \rightarrow W \rightarrow V \rightarrow P \rightarrow Y$$

b) Multiple indirect paths

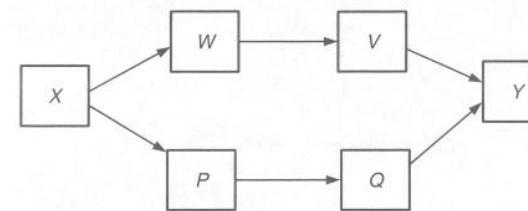


Figure 3.3 More complex indirect causal relationships

TYPES OF RELATIONSHIPS IN A THREE-VARIABLE MODEL

Any relationship between two variables will consist of two components – a causal component and a non-causal (spurious) component. The causal component can consist of a direct component, an indirect component or both.

It follows then that any relationship between two variables can be interpreted as:

- a direct causal relationship
- an indirect causal relationship
- a spurious relationship
- any combination of these.

Figure 3.4 illustrates the possibilities where we have three variables which, for the purpose of the example, I will call X , Y and Z . The relationship between X and Y could be any of the following:

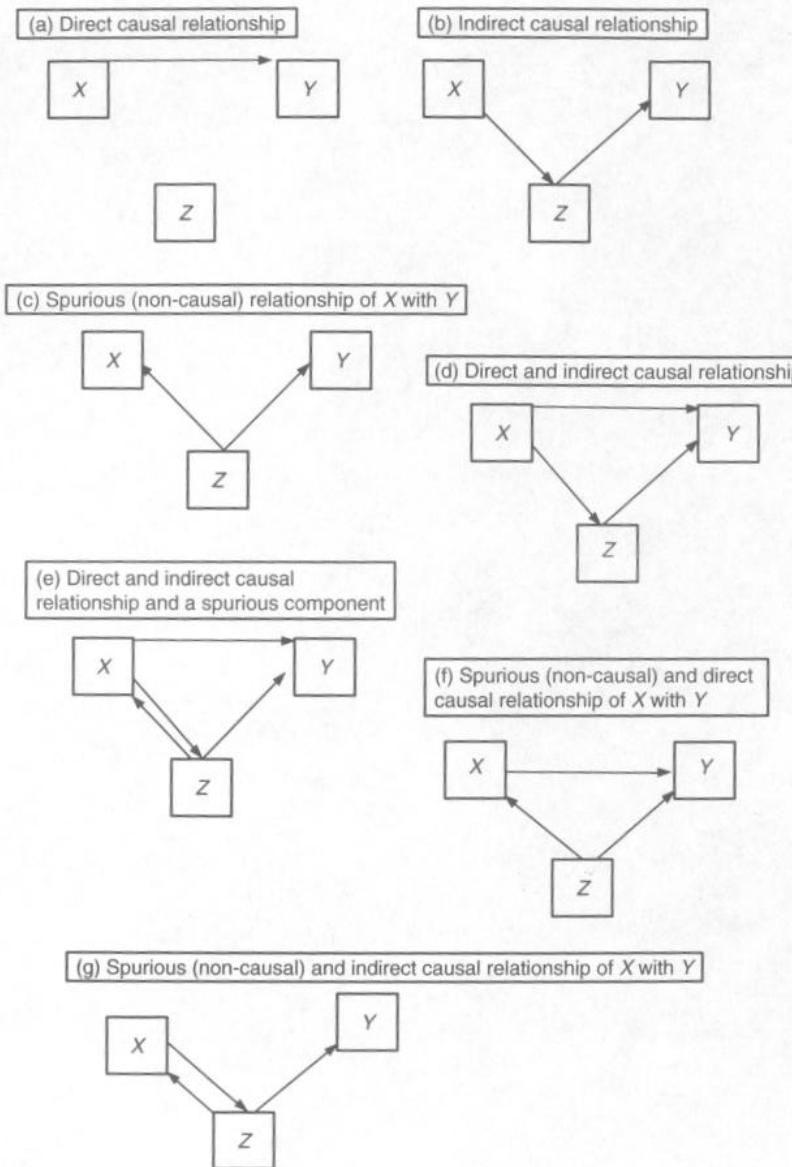


Figure 3.4 *Types of relationships between three variables*

- (a) *Direct causal* Y follows X in time, Y is capable of being changed, and it is plausible that X could produce changes in Y. In the absence of finding any other variable that is responsible for this relationship we may continue to argue that the observed correlation is direct and is causal (Figure 3.4a).
- (b) *Indirect causal* Y follows X in time, Y is capable of being changed, and it is plausible that X could produce changes in Y. However, in this case we are spelling out the mechanism by which X affects Y. We may think of Z as a single mechanism or a whole lot of intervening variables (Figure 3.4b).
- (c) *Spurious* X and Y are not causally related to one another. Even though Y might follow X in time and be capable of being changed, both X and Y are joint effects of some third variable Z. X and Y covary purely because Z has a simultaneous effect on both X and Y (Figure 3.4c).
- (d) *Both direct and indirect* The effect of X may be partly via its effect on an intervening variable and partly direct (Figure 3.4d).
- (e) *Direct and indirect causal and spurious* The relationship between X and Y could consist of three components: a direct causal part ($X \rightarrow Y$), an indirect causal part ($X \rightarrow Z \rightarrow Y$) and a spurious part ($X \leftarrow Z \rightarrow Y$) (Figure 3.4e).
- (f) *Direct causal relationship combined with a spurious component* (Figure 3.4f).
- (g) *Indirect causal relationship combined with a spurious component* (Figure 3.4g).

This set of possibilities results from situations in which we have only three variables. The more variables we take into account, the more complex matters become.

When we collect and analyse data it can be helpful to draw diagrams to spell out the ways in which we propose variables are interrelated. We need to specify:

- 1 whether relationships are presumed to be causal or spurious
- 2 whether causal relationships are expected to be direct or indirect
- 3 the mechanisms (intervening variables) underlying any indirect causal relationships.

Resolving these matters allows us to articulate our research question and the most plausible line of explanation.

Providing a frame of reference

The logic of making comparisons is fundamental to testing causal models. Consequently a central element in designing an explanatory

Proposition Divorce leads to emotional problems in young children.
[divorce → emotional problems of children]

Observation We find that many children with divorced parents have emotional problems.

Proposition Private schools produce high academic performance among their students.
[private schools → high performance]

Observation There are many instances of students from private schools achieving highly.

Proposition Youth unemployment is responsible for youth suicide.
[unemployment → high suicide rate]

Observation 50% of young people who commit suicide are unemployed.

Figure 3.5 *Propositions and observations without a comparative frame of reference*

research study is providing a comparative frame of reference. Different research designs go about the task of providing comparisons in different ways.

Comparing groups

By making comparisons we provide a frame of reference within which to try to make sense of particular findings. The importance of this can be seen in the illustrations in Figure 3.5. In these examples the observation in each case hardly provides convincing support of the proposition. One reason for this is that there is no frame of reference within which to make sense of the observations. There are no groups with which to compare and contrast the observations. It is only by making *comparisons* that our observations take on much meaning and we are able to eliminate alternative explanations.

The proposition that *divorce leads to emotional problems in young children* would encourage us to expect to find emotional problems among children whose parents are divorced. But finding such evidence will not get us very far. To go further down the path of explanation we must make comparisons. Are emotional problems greater than, the same as or less than those among children from intact families? Do the observations among children of divorced parents simply reflect that which we would find among *any* group of children?

Proposition Divorce leads to emotional problems in young children.

Restated proposition Children whose parents have divorced are more likely to display emotional problems *than are children whose parents have not divorced*.
[parental marital status → level of emotional problems]

Proposition Private schools produce high academic performance among their students.

Restated proposition Private schools will have higher levels of academic performance among their students *than will government schools*.
[school type → level of academic performance]

Proposition Youth unemployment is responsible for youth suicide.

Restated proposition Unemployed youth are more likely to commit suicide *than employed youth*.
[employment status → suicide rate]

Figure 3.6 *Using propositions with an explicit comparative frame of reference*

It will be recalled that when we say that two variables are related it means that *variation* or difference on one variable is linked to differences on the other variable. In this instance we have two variables: (1) parents' marital status (married versus divorced) and (2) child's emotional adjustment (low versus high). If our research design included only divorced parents we could say nothing about the impact of divorce on children. Even if 100 per cent of the children from divorced parents had a low level of emotional adjustment we could draw no conclusions about the impact of parental divorce on children's emotional adjustment. We would need to compare the adjustment of the children with divorced parents with the adjustment of children from intact families. The *difference* in adjustment levels in the two groups will provide strong evidence regarding the alleged effect of divorce on children's emotional adjustment.

The propositions in Figure 3.5 only mention one particular group (children of divorced parents; private school students; suicidal youth). It can be helpful to restate the propositions to highlight the implicit comparisons. Figure 3.6 restates these propositions using explicit comparisons.

MULTIPLE COMPARISON GROUPS

The independent variables above (parents' marital status, school type, youth employment status) have all been treated as two-category variables (dichotomies). However, we are not restricted to comparisons of

just two groups. Multiple comparison groups are possible and will arise from at least two circumstances.

Independent variables with more than two categories Where independent variables have more than two categories we can compare more groups and provide stronger and subtler tests of theories. For example, in examining the effect of divorce on children we could simply classify marriages into divorced or intact, or we could classify them as single (never married), intact, separated, widowed, divorced (repartnered) and divorced (not repartnered). If we took the latter approach we could make multiple comparisons. In doing so we can get a clearer picture of what is going on. If we find that it is the children of divorced parents who have repartnered who have the most problems we achieve a more focused understanding of the impact of parental marital status on children. We have learned that it is not divorce *per se* but the arrangements that follow divorce that are critical. Had we simply compared divorced and intact marriages we might never have identified this.

Combined effects of different independent variables It is possible that it is only when people have a particular *combination* of characteristics that an effect is produced. For example, it may be the *joint* effects of gender and age rather than each independently that is linked to suicide. We may hypothesize that when a person is both male and young, the likelihood of suicide is at its highest. If, for the purpose of this example, we think of age as a dichotomous variable (15–35 = young; 35+ = old), then we can think in terms of four groups based on the combination of these two independent variables. These are illustrated in Figure 3.7.

Comparing time points

A finding that children from divorced families are less well adjusted emotionally than those from intact families *presumes* that divorce is responsible for this difference. The problem with a simple comparison between groups (divorced and intact) is that it does not tell us whether divorce has actually produced any *change* in the emotional adjustment of children. Any conclusions would be much more convincing if we could track the emotional adjustment of children both before and after divorce to see if there was any change in the emotional adjustment of children following divorce.

By tracking children over time we could see if there was a change in emotional adjustment of the children and evaluate whether this was attributable to their parents' divorce.

MULTIPLE PRE-TESTS AND POST-TESTS

In the above example the measurement of emotional adjustment at the beginning of the study is often referred to as the 'pre-test' and the

	Male	Female
Young	Group 1: young males	Group 2: young females
Old	Group 3: old males	Group 4: old females

Figure 3.7 *Multiple comparison groups created by combining two independent variables*

remeasurement at the end is called the 'post-test'. Between the pre-test and the post-test certain critical events occur – in this case parental divorce (or non-divorce as the case may be). These events serve as the independent variables in the study – i.e. the presumed causal variable behind the observed change. In some designs these intervening events are called the intervention or the treatment (see Chapter 4).

We can collect data about more than two time points rather than being restricted to simple 'before' and 'after' data collections. We might have repeated measurements over an extended period to track 'ups' and 'downs' and to track *trends* before and after any critical event. Multiple 'pre-tests' and multiple post-tests can help distinguish between short term and long term trends. They also help identify the effect of the 'intervention' or independent variable. For example, a simple measure of emotional adjustment before and after divorce might show a decline in adjustment. But what if there was strong evidence of the decline having commenced well before the divorce? Only multiple measures before divorce would show that a trend had already begun. Similarly, a simple post-test might show that adjustment is lower after divorce but multiple post-tests might show poor adjustment immediately after divorce but a steady improvement over time (see Chapter 6).

Making meaningful comparisons

Ideally the groups we are comparing should be the same in all relevant respects except in regard to the independent variable. For example, if we want to test the idea that non-government schools produce students that achieve better academically than government schools we would need to be confident that both types of schools contained comparable students. We would need this so that we could be confident that the only relevant difference between the two sets of students is in the type of school attended. If the students differed in additional ways, how would we

know which of the differences was responsible for any differences in academic achievement?

Comparisons of children from divorced and intact families are complicated by the fact that typically children of divorced parents are older than children from intact families. Any differences between the level of emotional adjustment of the two sets of children could be due to age differences rather than the marital status of their parents.

The same problem can arise when looking at the same group or category of people over time. Ideally with comparisons over time the only differences between the pre-test and the post-test should be the event we are proposing as the cause of any change. However, many events can occur which can account for the change and these may confuse our comparisons and our attribution of what lies behind the change in the outcome variable (see Chapter 4 for further discussion).

Since comparisons are central to good research design we must ensure that they are meaningful. The more we can remove unintended and unknown differences between groups, the more we reduce the risk of mistaking spurious relationships for causal relationships – that is, the more we eliminate alternative plausible explanations.

There are four main strategies for maximizing the comparability of groups.

MATCHING

When recruiting the groups we can deliberately match them on relevant characteristics. Thus when comparing students attending government schools on the one hand and fee paying private schools on the other hand we should ensure that the two sets of students are similar in terms of intelligence, aspirations, parental resources, education values, family history, gender and age. By comparing like with like we should be able to isolate the effect of the type of school on academic achievement.

The problem in comparing like with like is to establish the identity of all the variables on which we need to match the groups. We can match for the characteristics that we know might contaminate our results but there may be other factors that we have not thought of.

Ex post facto matching Ideally groups should be matched at the beginning of the study before critical events (e.g. divorce, attending the school). Another approach referred to by Spector (1981: 48) as a 'patch-work procedure' involves creating matched groups from a whole pool of study participants *after* all the data have been collected. For example, we might have a large number of students from government and fee paying private schools. From this pool we could extract a group of government school students and a group of fee paying private school students who are comparable in terms of intelligence, aspirations, parental resources, values about the importance of education, family history, gender, and

age. We could then compare these groups and see if there were differences in their level of academic achievement.

There are many problems with this approach. The obvious one is that we must have the relevant information on which to match. Since this will always be limited, the groups are likely to remain unmatched on important but unknown factors. Another problem is that many cases from both groups simply will not have matches from the other group and will have to be discarded. If we match on more than a small number of factors we will often end up with only a very small number of people in the study since the number of cases from both groups that match can be quite small.²

RANDOMIZATION

A simpler and more effective way of making groups comparable is to randomly allocate people to different groups at the beginning of a study. By randomly assigning individuals to each group any differences between groups should be random rather than systematic. So long as groups are large enough, random assignment should automatically produce groups with comparable profiles on both known and unknown factors. From a statistical perspective, random assignment of people to groups will make the groups identical for all intents and purposes and provides what Davis calls 'The all purpose spurious insurance of randomization' (1985: 35). We control for an infinite number of plausible rival hypotheses *without specifying what any of them are* (Campbell, 1989).

This is the approach taken in drug trials. Individuals are randomly assigned to one of several groups and each group is then given a different drug. Since the groups should have virtually identical profiles to begin with, any differences in outcomes between the groups should be due to the different treatments administered to each group. However, this approach is often not applicable in social research because practical and ethical considerations preclude us assigning people to groups and then doing something to one group to see what effect it has (see Chapter 5). Advantages and drawbacks of this approach to social research are discussed in Chapters 4 and 5.

MATCHED BLOCK DESIGNS

While randomization minimizes the chances of initial differences between groups there is still the chance that there will be differences between the groups – especially in smaller groups. Where we want to be certain that the groups are comparable on a particular characteristic we can take an additional step to guarantee that the groups are comparable at least on specific variables. This is achieved with what is called the randomized block design – a combination of both randomization and matching.

Suppose we wanted to evaluate the effect on academic performance of three methods of delivering university courses: (1) face-to-face delivery

of lectures and tutorials, (2) correspondence using written materials and (3) electronically using the internet and e-mail. Further, suppose that we could assign individuals to one of these three modes. Random allocation would be the obvious way. But suppose we want to be absolutely certain that the students in each of the three modes are of comparable academic ability.

To achieve this we would do the following:

- 1 Obtain a measure of academic ability at the pre-test.
- 2 Rank all the students from highest to lowest ability.
- 3 Select the three students with the highest ability (because we have three learning modes). This group of three students is block 1.
- 4 Randomly allocate each of the three students to a group (one to face-to-face; one to written correspondence; and one to electronic learning).
- 5 Select the next three most able students (block 2) and repeat the random allocation to groups.
- 6 Repeat this until all students have been allocated to one of the three groups.

STATISTICAL CONTROLS

An alternative way of making groups comparable is to do so at the data analysis stage after data have been collected. It involves multivariate analysis that, in one respect, matches groups on specified variables (see Chapter 12 for a discussion of the logic of some of these techniques).

Although the procedures of multivariate analysis can be very complex, the essential logic is simple. Suppose we want to compare the emotional adjustment of children from divorced and intact families but we believe that any comparison would be confounded by the fact that, on average, children whose parents divorce are older than those from intact families. Any greater maladjustment among children whose parents have divorced could occur because they are older, as older children display greater maladjustment.

Multivariate analysis removes any effects that might be due to age by comparing like with like – by selecting, say, preschoolers and comparing those from divorced and intact families to see if, despite similarity of age, the children show different levels of adjustment. The same comparisons could be repeated among say 5–8 year olds, 9–12 year olds, 13–15 year olds and so on.

The obvious shortcomings of this approach are similar to that of matching. We can remove the influence of variables that we have thought of and on which we have data but we cannot remove the effects of unknown variables, or those for which we have no data. Since we can never know what factors we have missed there is always the danger that factors we have not thought of may be contributing to group differences.

Interventions and independent variables

Throughout this chapter I have used the terms 'intervention', 'treatment', 'independent variable' and 'groups' more or less interchangeably. Research designs vary in terms of the type of independent variable employed and in the number of independent variables built into the study.

TYPES OF INDEPENDENT VARIABLES

When thinking in causal terms ($X \rightarrow Y$) X is the independent variable. When we conduct the research and do the analysis we compare the outcomes (Y) for different groups. The groups are defined according to which category of the independent variable they belong to. In some studies the investigator introducing an *active intervention* defines the independent variable (e.g. allocate different people to one of three modes of course delivery – face-to-face, correspondence or electronic). In other situations the independent variable is defined by a *naturally occurring intervention* (e.g. a person retires or remains employed). In other situations the independent variable is determined by the relatively *fixed attributes* of participants (e.g. sex, education, race etc.) rather than by interventions of any sort.

NUMBER OF INTERVENTIONS

Research designs can also differ in terms of the number of interventions that are made. Where the independent variable involves either an active or a passive intervention we can examine the effect of single versus repeated interventions. We can see whether a single intervention has a different impact than the cumulative impact of repeated interventions. For example, we might be evaluating the impact of teachers giving students negative feedback on their work. Initial criticism of the work may boost a student's effort and performance but repeated negative feedback may lower performance as the student loses confidence. A multiple intervention design helps refine our understanding of the way in which interventions affect participants and enables us to rule out alternative ways of interpreting results.

Dimensions of a research design

The above discussion has identified six core elements of a research design. The particular mix of the elements in any study will yield different designs.

These six main elements in producing a research design are:

- 1 *The number of groups in the design* Designs will vary from those with no comparisons (e.g. single case design or case study) to those with many different comparison groups.
- 2 *The number of 'pre-test' measurement phases* Designs vary from those with no 'pre-test' (e.g. cross-sectional designs and some experimental designs) to those with a series of 'pre-tests' which establish pre-existing trends before an event.
- 3 *The number of 'post-test' measurement phases* All designs require at least one 'post-test' – the measurement of an outcome variable. In some designs (e.g. cross-sectional) there will be one 'post-test', while other designs can have many post-tests to help distinguish between short and long term outcomes.
- 4 *The method of allocation of cases to groups* In multiple group designs groups can be made comparable by allocating people to different groups by random allocation, 'pre-test' matching, *post hoc* matching or block matching or by using statistical controls in the analysis phase.
- 5 *The nature of the intervention* Studies that rely on existing variation (cross-sectional designs and those with 'fixed' independent variables) have no interventions. Other designs rely on interventions between a pre-test and a post-test. These 'interventions' may be either active or natural.³
- 6 *The number of interventions* Designs with an intervention can have either a single intervention or multiple interventions. Multiple interventions can be used to identify the effect of cumulative 'treatments'.

A range of research designs

The combination of the possibilities created by these six elements of a research design yields a large number of possible designs. A useful summary discussion of many of these designs is provided by Spector (1981). To impose some order on this range of possibilities it is helpful to think in terms of four broad types of design. For each design type, decisions taken by the investigator will produce variations within the type. These four types provide the structure for the rest of the book. The four broad types of design are experimental, longitudinal, cross-sectional and case study. I will examine each of these in turn.

Experimental design

The classic version of the experimental design has the following elements:

- 1 One pre-intervention (pre-test) measure on the outcome variable.
- 2 Two groups: one group that is exposed to the intervention (the experimental group) and one group that is not exposed to the intervention (the control group).

Method of allocation to groups	Pre-test	Intervention (X)	Post-test
Random (experimental group)	Measure on outcome variable (Y)	'Treatment'	Measure on outcome variable (Y)
Random (control group)	Measure on outcome variable (Y)	No 'treatment'	Measure on outcome variable (Y)

Figure 3.8 *Classic experimental design*

- 3 Random allocation to the groups before the pre-test.
- 4 One intervention (test/treatment).
- 5 One post-intervention (post-test) measure on the outcome variable.

This design is illustrated in Figure 3.8.

The analysis of any effect of the intervention focuses on changes in the experimental group before and after the intervention and a comparison with the rate of change in the control group. If the change is greater in the experimental group than in the control group the researcher will attribute this to the impact of the independent variable (the treatment).

Longitudinal design

The basic form of this design involves:

- 1 One group.
- 2 One pre-'intervention' measurement on the outcome variable.
- 3 One 'intervention' where everyone receives the 'treatment'.
- 4 One post-'intervention' measurement on the outcome variable.

In effect, this design is similar to the experimental design except that there is no control group and typically only one 'experimental' group. The design is illustrated in Figure 3.9.

The analysis in this design compares the pre-intervention measures with the post-intervention measures. Change in these scores can reflect the influence of X on Y. However, the absence of a randomized control

Method of allocation to groups	Pre-test	Intervention (X)	Post-test
Random (one group)	Measure on outcome variable (Y)	'Treatment'	Measure on outcome variable (Y)

Figure 3.9 *Simple longitudinal design*

group makes it difficult to know whether the intervention or some other factor produces any change.

Cross-sectional design

The basic elements of the cross-sectional design are as follows:

- 1 Instead of interventions the cross-sectional design relies on existing variations in the independent variable(s) in the sample.
- 2 At least one independent variable with at least two categories is present.
- 3 Data are collected at one point of time.
- 4 There is no random allocation to 'groups'.

This design mirrors the post-intervention phase of the classic experimental design but without any random allocation to 'groups' being made. The data for this design are collected at one point of time and are analysed by examining the extent to which variation in the outcome variable is linked with group differences. That is, to what extent do those in different categories of the independent variable differ in relation to the outcome variable? Causal relationships are established by utilizing statistical controls rather than random allocation of people to groups.

This design is illustrated in Figure 3.10. In this case the 'intervention' is simply being in a different category of the independent variable.

Case studies

Case study designs rely less on comparing cases than on exhaustive analysis of individual cases and then on comparing cases. A distinguishing characteristic of case studies is that contextual information is collected about a case so that we have a context within which to understand causal processes.

Method of allocation to groups	Pre-test	Intervention (X)	Post-test
Non-random	None	'Treatment'	Measure on outcome variable (Y)
Non-random	None	No 'treatment'	Measure on outcome variable (Y)

Figure 3.10 *Simple cross-sectional design*

Case study designs might consist of a single case study (e.g. a community study, a study of an organization) or a series of case studies with perhaps each case testing a theory from a different angle. It is useful to think of a case study in a similar way to an experiment. We do not finally reject or accept a theory on the basis of a single experiment; we try to replicate an experiment and conduct it under a variety of conditions. Similarly, a case study project that entails a single case study is analogous to a single experiment. If similar results are found in repeated case studies, or predictable differences in results are found for particular cases in the study, then we develop greater confidence in the findings of the cases in the same way that we gain confidence in experimental results that are found in repeated experiments.

Summary

This chapter has examined ways of structuring research designs to help draw convincing causal inferences from the research. Since causes cannot be observed they must be inferred from observations. However, incorrect inferences can easily be made. The chapter has considered ways of structuring research to improve the quality of these inferences. A range of ways of interpreting correlations between variables was outlined and criteria required for inferring that a correlation reflects a causal connection were provided.

The chapter emphasized the importance of making meaningful comparisons between groups as a core element of drawing causal inferences.

One of the tasks of research design is to structure the research so that meaningful comparisons of outcomes between groups can be made.

Finally, the chapter outlined six core elements of research designs: number of groups, number of pre-tests, number of post-tests, nature of allocation to groups, type of interventions and number of interventions. Research designs vary in the way these elements are dealt with. Four main types of research design – experimental, longitudinal, cross-sectional and case study – were then briefly described. These broad categories of design provide the framework for the remainder of this book.

Notes

1 There can be exceptions to this where a suppressor variable may be operating to mask a causal relationship. See Rosenberg (1968: Chapter 4).

2 We shall see later that this problem of loss of cases can be overcome with certain forms of multivariate analysis.

3 A naturally occurring intervention is an event that takes place between pre-test and post-test without the investigator initiating any intervention. An example might be a study of families over time in which some families experience parental divorce. Divorce is the 'natural intervention'.

PART II

EXPERIMENTAL DESIGNS

4

TYPES OF EXPERIMENTAL DESIGN

This chapter builds on the elements of research design outlined in Chapter 3 to introduce a variety of types of experimental designs. It outlines the different environments in which experiments can be conducted and describes both simple and more complex variations of experimental designs. This chapter does not seek to evaluate experimental designs: that is the task of Chapter 5. The discussion also provides a useful framework within which to understand some of the strengths and weaknesses of the research designs discussed later in this book.

The classic experimental design

The classic experimental design focuses on two variables: the independent variable (the cause/intervention) and the dependent variable (outcome). The purpose of the design is to remove the influence of other variables so that the effect of the intervention can be clearly seen. Since the classic experimental design was outlined in the previous chapter there is no need to repeat it here. However, an example will help clarify its elements.

Suppose we wanted to test the proposition in Figure 4.1, namely *the more difficult it is to join a group the more desirable the group will seem*.¹ In order to test this proposition it is necessary to develop measures of desirability of membership and to define what is meant by 'difficulties of joining'. The expression could refer to things such as restricted membership, long waiting periods, high fees, harsh initiation, high time commitments. For the purpose of this exercise we will use 'severity of