MANAGEMENT AND BUSINESS RESEARCH

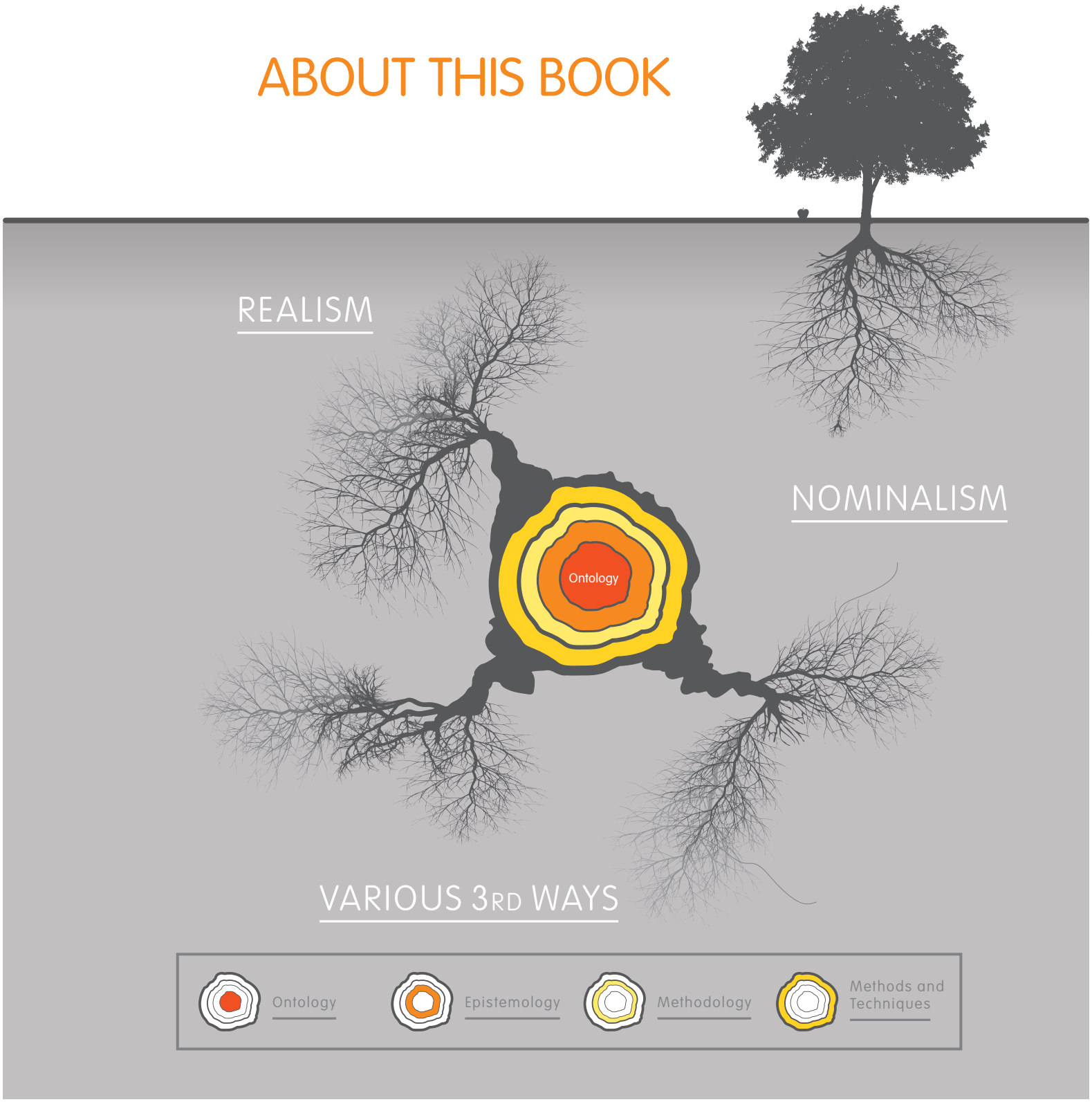
5th EDITION

Mark Easterby-Smith,
Richard Thorpe
Paul R. Jackson

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

ABOUT THIS BOOK



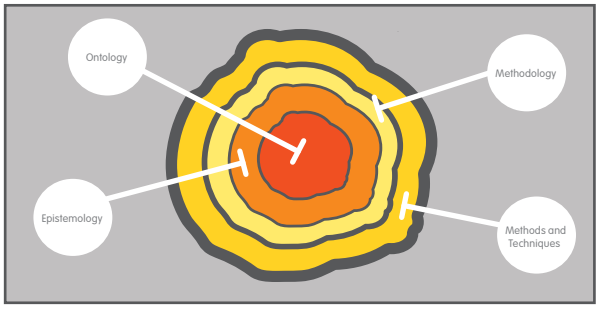
We use the metaphor of a tree to represent how the research process unfolds. The key elements of the tree are the roots, the trunk and branches, the leaves, and the fruit – and each of these parallels an aspect of conducting research.

The *roots* symbolize the research traditions within particular disciplines as well as the experiences of past researchers from particular fields. These perspectives, understandings, ideas and beliefs are drawn up (as the tree draws nutrients from the soil) to form the basis of the researcher's ideas in relation to such things as design, methods and forms of analysis.

The *trunk* transports the nutrients from the roots through the branches to the leaves and fruit; it also provides strength and shape to the tree. Here we use a simplified cross-section of the trunk to symbolize four main features of a research design. The inner ring (or heartwood) is the densest part of the trunk, and we use it to represent *ontology*, the basic assumptions made by the researcher about the nature of reality. The next ring represents *epistemology*, the assumptions about the best ways of inquiring into the nature of the world. The third ring from the centre represents *methodology*, or the way research techniques and methods are grouped together to provide a coherent picture. And the fourth ring represents the individual **methods and techniques** that are used for data collection and analysis. The four rings are named and ordered in this way, because the most visible parts of research projects are the methods and techniques used for data collection and analysis and represented by the outer ring. The three inner rings are increasingly hidden from the external observer, yet each makes a critical contribution to the strength, vitality and coherence of the research project.

Moving up and along the *branches*, the leaves and fruit form the tree's canopy. The *leaves* collect energy from sunlight, and represent the collection and analysis of data within a research project. It is the collection of research data that stimulates new ideas and enables the evaluation of existing theories. Here we distinguish between three main kinds of data based on the underlying epistemology (second ring in the trunk), according to whether they are essentially positivist, constructionist or hybrid approaches. To provide differentiation we indicate the positivist approaches in orange, constructionist approaches in green, and hybrid approaches in a combination of these colours.

In the final chapter of the book we focus on the *fruit* of the tree, which represents the way the research is written up and communicated to third parties. Here we show the coherence between the outputs of the research, and the ontology, epistemology, methodology and methods which underpin any research. In this way, the views and



values adopted by the researcher from the early stages of the framing of the research, the design of the research project and the collection of the data are coherently connected and linked.

Within the chapters that follow we have placed a number of these icons in key locations. This is not intended to be exhaustive in the sense that everything is necessarily covered by the icons; nor are we seeking to explain everything through the use of these icons. Our intention is mainly to use them in the light sense as an organizer and as a reminder of the origins of some of the ideas being discussed.

Key of symbols


Ontology


Epistemology


Methodology


Methods and Techniques

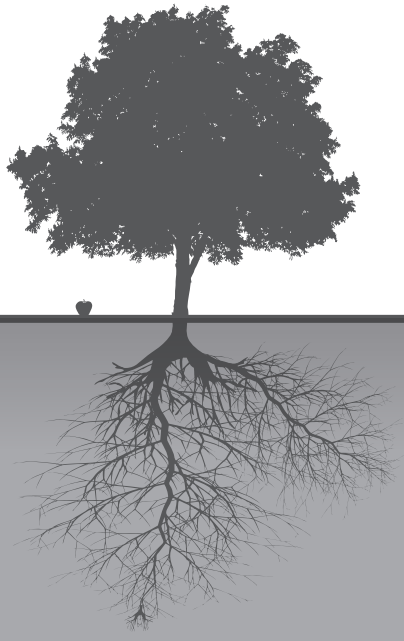

Positivist

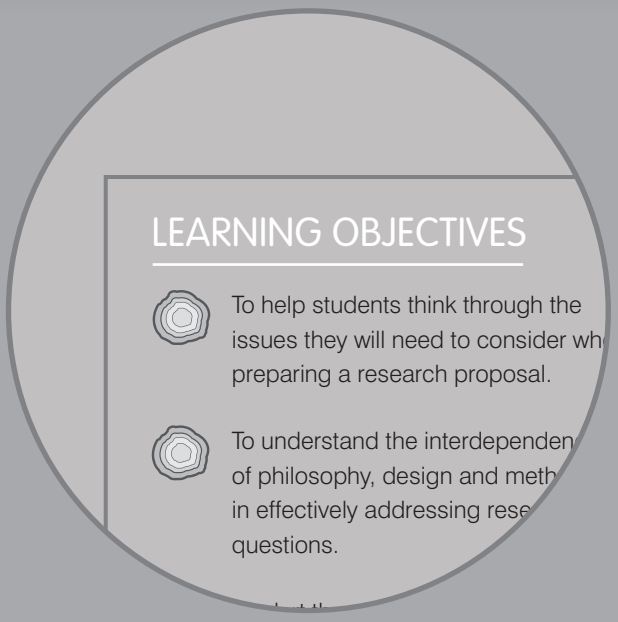

Constructionist


Hybrid approaches


Output

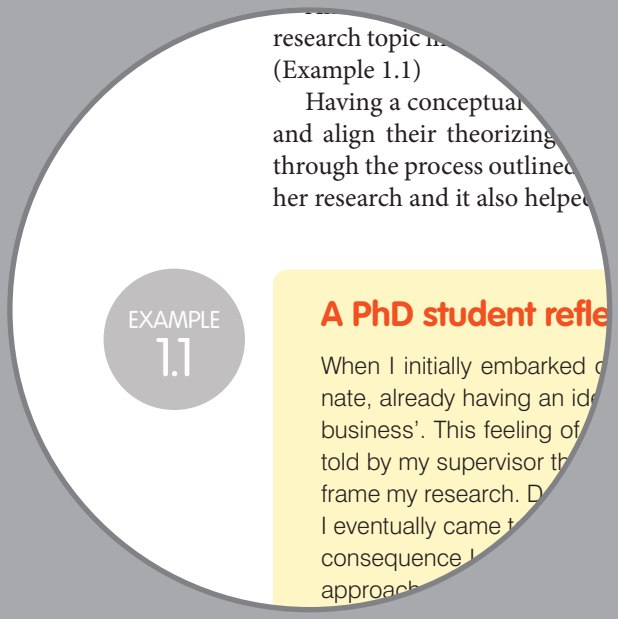
GUIDED TOUR



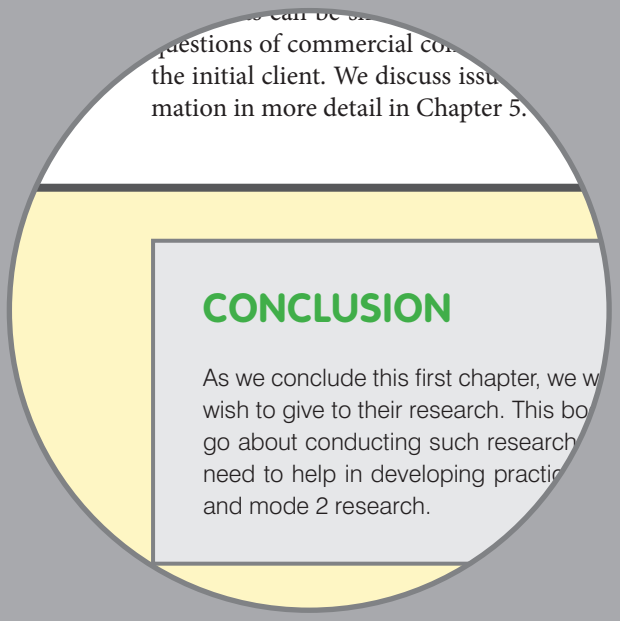


Learning objectives Each chapter starts by setting out clearly what key information you should soon understand, so you can easily track your progress.

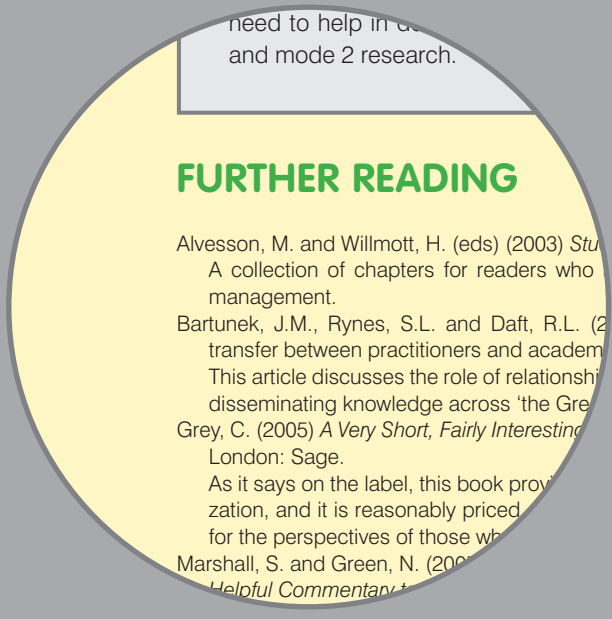
Examples Real management and business research examples are highlighted in these boxes, to show you how research happens in reality.



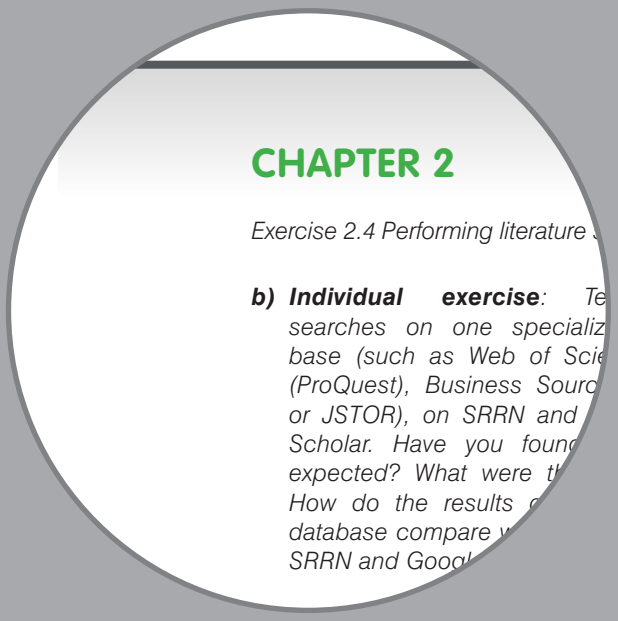
Exercises These activities are designed to get you to practise and think about the key research tasks covered in each chapter.



Conclusion Review the contents of each chapter in the Conclusion, an easy-to-read review and discussion of the most important information you've read.

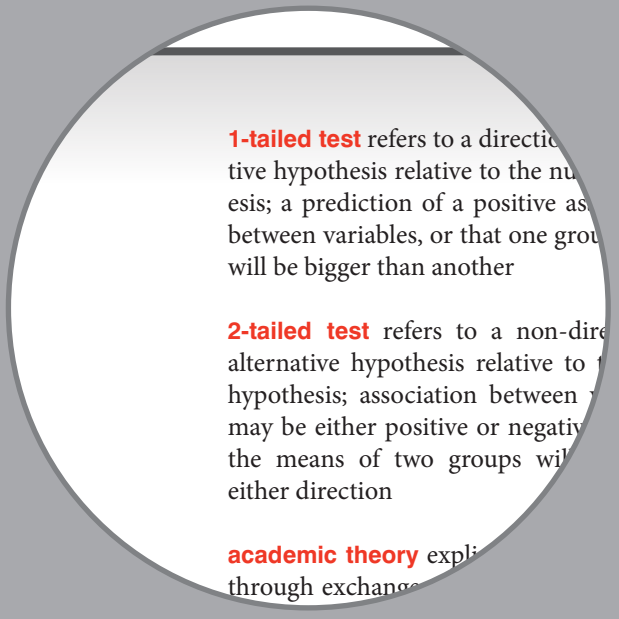


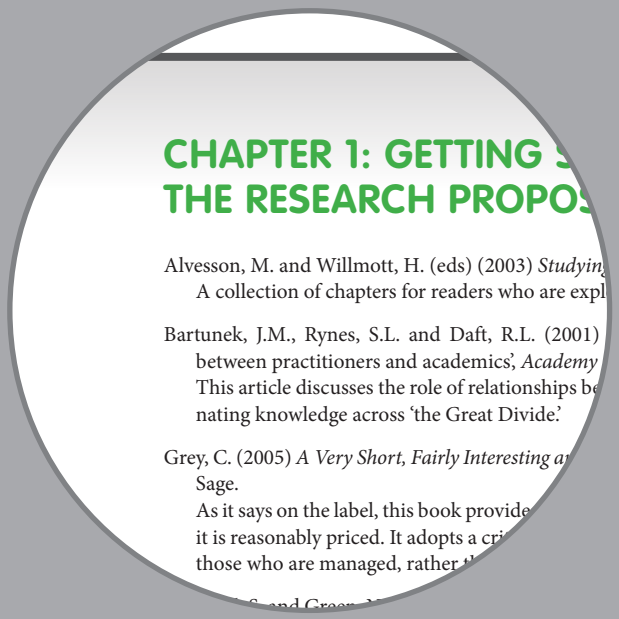
Further reading Each chapter ends by highlighting books and articles that provide more in-depth information on the methods presented in each of the chapters and insights into how these methods have been used successfully in existing studies.



Answers to exercises Turn to the back of the book to get responses to those exercises that have specific answers.

Glossary Key terms appear in bold in the text, and their definitions are presented in the Glossary near the back of the book.





CHAPTER 1: GETTING S
THE RESEARCH PROPOS

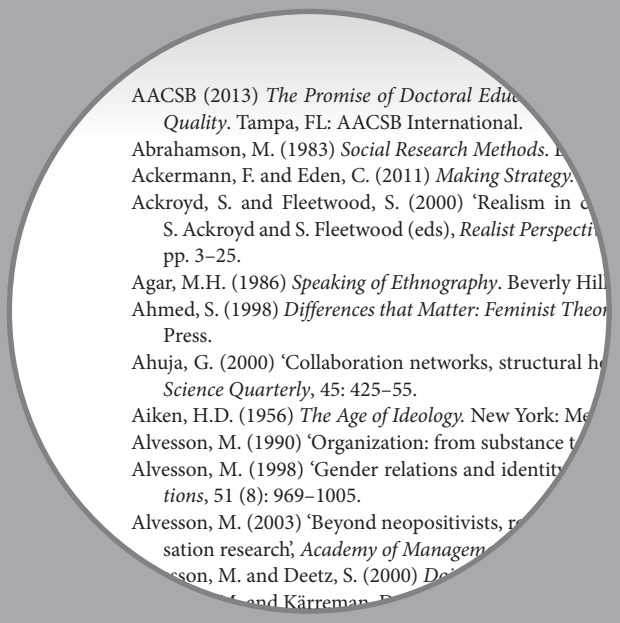
Alvesson, M. and Willmott, H. (eds) (2003) *Studyin*
A collection of chapters for readers who are expl

Bartunek, J.M., Rynes, S.L. and Daft, R.L. (2001)
between practitioners and academics', *Academy*
This article discusses the role of relationships be
nating knowledge across 'the Great Divide.'

Grey, C. (2005) *A Very Short, Fairly Interesting a*
Sage.
As it says on the label, this book provide
it is reasonably priced. It adopts a cri
those who are managed, rather t

Further Readings For ease of reference we have gathered the further reading from all chapters into one section.

Bibliography A wealth of reading material is covered here, as all the sources that have collectively informed the writing of this book are gathered in one place.



AACSB (2013) *The Promise of Doctoral Educ*
Quality. Tampa, FL: AACSB International.
Abrahamson, M. (1983) *Social Research Methods.*
Ackermann, F. and Eden, C. (2011) *Making Strategy.*
Ackroyd, S. and Fleetwood, S. (2000) 'Realism in c
S. Ackroyd and S. Fleetwood (eds), *Realist Perspecti*
pp. 3–25.
Agar, M.H. (1986) *Speaking of Ethnography*. Beverly Hil
Ahmed, S. (1998) *Differences that Matter: Feminist Theor*
Press.
Ahuja, G. (2000) 'Collaboration networks, structural h
Science Quarterly, 45: 425–55.
Aiken, H.D. (1956) *The Age of Ideology.* New York: Me
Alvesson, M. (1990) 'Organization: from substance t
Alvesson, M. (1998) 'Gender relations and identity
tions, 51 (8): 969–1005.
Alvesson, M. (2003) 'Beyond neopositivists, r
sation research', *Academy of Managem*
sson, M. and Deetz, S. (2000) *Do*
and Kärreman, D

SAGE edge™

edge.sagepub.com/easterbysmith

For the cyber-hungry among you, or for those who are more inclined to read in pixels than in print, this textbook comes with SAGE edge (**edge.sagepub.com/easterbysmith**), a range of free tools and resources that will enhance the teaching and learning experience.

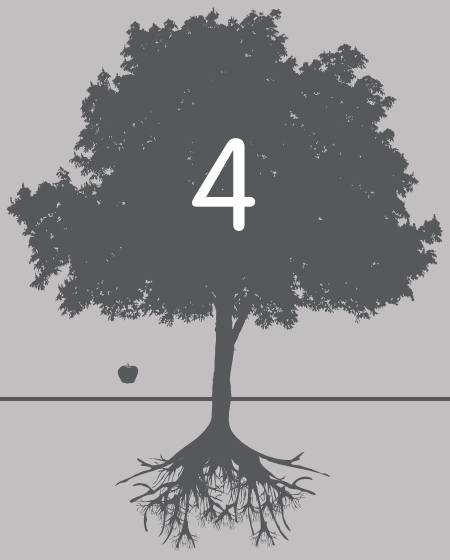
SAGE edge™ **for Students** provides a personalised approach to help you accomplish your coursework goals in an easy-to-use learning environment and features:

- **Learning objectives** for each chapter that reinforce the most important material
- Practice **quizzes** that allow you to assess your understanding of course material, with 3 short answer questions per chapter and 10 multiple choice questions per chapter
- Exclusive access to **SAGE journal articles and chapters** which have been carefully selected to support and expand on key concepts and to encourage critical thinking
- Links to **websites and videos** that offer a new perspective and an alternative learning style
- A complete **dataset** so you can practice data analysis with helpful **notes** from the authors
- **Action plan** for a more personalised learning experience
- Mobile friendly **eFlashcards** to strengthen your understanding of key terms and concepts

SAGE edge™ **for Instructors** supports teaching by making it easy to integrate quality content and create a rich learning environment for students. SAGE edge™ instructor resources for this book include:

- **PowerPoint slides** for each chapter that can be edited for use in teaching
- **Instructor's Manual** containing chapter overviews, hints and tips, guidance on the exercises and examples in the book and additional exercises (some incorporating the use of video).

DESIGNING MANAGEMENT AND BUSINESS RESEARCH



LEARNING OBJECTIVES

- To appreciate how research philosophies impact research designs.
- To understand what is regarded as good theory within each tradition.
- To be able to critique research designs.
- To be able to develop and justify research designs.

THE ESSENCE OF RESEARCH DESIGN

Research designs are about organizing research activity, including the collection of data, in ways that are most likely to achieve the research aims. Let us start with an example.

In 1985, the US businessman Kenneth Lay (see Figure 4.1, for a police mugshot of Lay in 2004) founded Enron after merging Houston Natural Gas and InterNorth. The company expanded very rapidly, first as an energy company and then through diversification, so that by mid-2000 its stock was valued at over $60 billion. But on 2 December 2001, it filed for bankruptcy following the failure of a rescue bid from a rival company. It quickly emerged that the company had systematically developed accounting practices that had inflated revenue while keeping debts and liabilities off the books. Not only had these practices been fostered by Lay and his colleagues at the top, but also the global consultancy firm Arthur Anderson, which had regularly audited Enron, had failed to report any problems.

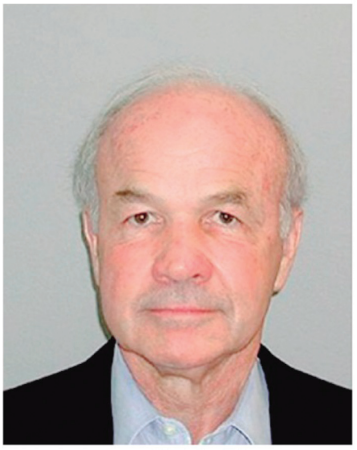


FIGURE 4.1
Businessman Kenneth Lay

This raised major implications about the efficacy of accepted practices for auditing corporate accounts. Imagine that you wish to conduct empirical research into the changes in corporate accounting practices following this scandal. In the previous chapter we outlined an ontological dimension containing the positions of realism, internal realism, relativism and nominalism. Following each of these positions in order, your research design might involve: (1) conducting a review of new legislation and accountancy practices published over the period 2002–2005; (2) sending out a postal questionnaire to 200 members of the Chartered Institute of Management Accountants; (3) arranging to interview one accountant from each of 20 different organizations including companies and consultancies; or (4) getting a job for a year in the accounting department of a US energy company.

Each of these brief statements includes at least three decisions about research designs. In (1) there is a decision to focus on two categories of written documents published over a specific period of time; in (2) the decision is to design a questionnaire, which will be mailed to a specific number of people who belong to one professional association; in (3) the aim is to gather views from a medium sample of people likely to have different perspectives and experiences; and in (4) the decision is to invest personal time in observing accountancy practices in a US company within a specific industry. Each of these decisions specifies courses of action in preference to other options. For example, the focus on published sources precludes internal corporate documents; the decision to mail the questionnaire precludes face-to-face interviews; the decision to interview only one person from each company precludes multiple perspectives from any company; and the decision to work in one company precludes obtaining direct data from other companies.

This is the essence of research design: it is about making choices about what will be observed, and how. But each of these designs is incomplete, and there are many other choices to be made, and features to be specified. A research design is a statement written, often before any data is collected, which explains and justifies what data is to be gathered, how and where from. It also needs to explain how the data will be analysed and how this will provide answers to the central questions of the research.

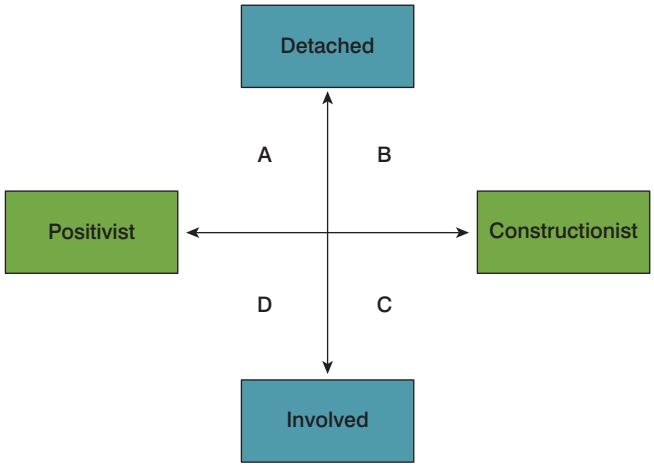
We can also analyse the Enron example against the Detached Engaged dimension identified at the end of Chapter 3. In the first case, since the focus is on documents in the public domain, the researcher remains quite *detached* from the source of the investigation. In each of the subsequent examples, he or she is likely to become increasingly involved with the people who influence the evolution and implementation of accounting practices. This defines a second major dimension of choice for the researcher, which is the degree of *engagement* with the object of the research. It is similar to the pure/applied research distinction, where the former tries to maintain the independence and objectivity of the researcher, and the latter encourages greater engagement and interchange between the researcher and the research setting.

In this chapter we explain: what a research design is; the main choices that need to be made; how research designs vary according to the underlying philosophical position; and how the quality of research designs can be judged. In the later chapters of the book we will be looking in detail at techniques and methods for gathering and analysing qualitative and quantitative data, though questions of research design need to be resolved before gathering (much) data.

We produce the basic dimensions of the matrix in Figure 4.2.

Remember that the poles in the matrix represent extreme positions, and because it is relatively rare for research designs to be purely one or another we will now focus on the space between them, essentially the quadrants, which we have labelled A, B, C and D. We also need to emphasize that the horizontal dimension is a continuum between strong positivism on the left-hand side, and strong constructionism on the right-hand side. In the next two sections we give examples of typical methodologies that fit into each of these quadrants, although we have grouped them as primarily positivist or constructionist methodologies. There are also some methods and designs that bridge across quadrants, or combine elements of constructionism and positivism. We refer to these as mixed methods, and we will cover these in the third section.

FIGURE 4.2
Epistemology and research style



POSITIVIST RESEARCH DESIGNS: QUADRANTS A AND D

As we noted in Chapter 3, positivist methods usually incorporate the assumption that there are true answers, and the job of the researcher is either to start with a hypothesis about the nature of the world, and then seek data to confirm or disconfirm it, or the researcher poses several hypotheses and seeks data that will allow selection of the correct one. The ideal methodologies for doing this are experimental and quasi-experimental methods, and we will describe the key principles of each below. We then look at survey methodologies, which are generally associated with positivism in the sense that they are looking for patterns and causal relations which are not directly accessible.

EXPERIMENTAL METHODS

Classic experimental method involves **random assignment** of study participants to either an experimental or a control group. Conditions for the experimental group are then manipulated by the experimenter/researcher in order to assess their effect in comparison with members of the control group who are receiving no unusual conditions.

Possibly the most famous experimental studies in the field of management were the Hawthorne experiments conducted by Elton Mayo at the General Electric Hawthorne Plant in Illinois between 1927 and 1932. One experiment involved the relocation of six women (the experimenters selected the first two, who were each asked to select two more) into a room separate from the rest of the employees who assembled telephone relays. Their working conditions were modified systematically in order to establish whether there was any link between physical conditions and productivity. An observer was located in this room, making notes of what was happening and also keeping them informed about progress of the experiment and listening to their ideas and complaints. Over a period of many months changes were made, including shortening the working day, introducing increasing amounts of breaks into the day, and eventually providing a hot meal in the middle of the morning shift. With each change, productivity increased, which would suggest a correlation between productivity and the easing of working conditions. However, at the end of the experiment they returned conditions to the situation at the outset, expecting productivity to decrease to the initial level – but it increased once more. This observation led to the development of **human relations theory**, which stressed that positive relationships between employees and their supervisors were more significant than the physical circumstances of their work as predictors of productivity.

Questions for discussion about the Hawthorne experiment

EXERCISE 4.1

1. What is the primary question/hypothesis of the researchers?
2. What are the key features of the research design?
3. In what ways does the Hawthorne experiment diverge from classical experimental design?

There are three important implications from the Hawthorne experiment. First, the study showed that the most significant findings emerged because the experiment went *wrong*,

in the sense that the expected results were not obtained. Second, the design was very systematic, including the return to the original condition. These features reflect the ideas of Austin about what leads to scientific breakthroughs as discussed in Chapter 1. The third implication is that it has raised awareness of the **experimenter effect** whereby human behaviour can be affected, and potentially 'distorted', by the presence of an observer. Most people behave differently if they think they are being watched.

The main *advantages* of experimental research designs are that they encourage clarity about what is to be investigated, and should eliminate many alternative explanations because the random assignment ensures that the experimental and control groups are identical in all respects, except for the focal variable. It is also easier for another researcher to replicate the study, and hence any claims arising from the research can be subjected to public scrutiny. The *disadvantages* are practical and ethical. With medical research there is always the danger that volunteers will be harmed by drug tests; hence stringent ethical guidelines have been developed, which are now filtering into social science and management research (see Chapter 5). Also, when dealing with people and business situations it is rarely possible to conduct true experiments with randomization. For example, if a company wants to assess the value of an elite highflier career development scheme, it cannot simply assign managers at random to the scheme because the managers themselves will be aware of what is happening, and there will also be strong ethical, performance-related and employment law objections to this arbitrary assignment. For this reason, quasi-experimental methods have been developed in order to circumvent the problem of random assignment.

Quasi-experimental designs

A key feature of **quasi-experimental design** is the use of multiple measures over time in order to reduce the effects of control and experimental groups not being fully matched (Shadish et al., 2002). Individuals are not allocated randomly to the treatment group and the control group, but rather allocation takes place on some other criterion, usually by using intact groups. As a result, the **validity** of inferences from this type of design depends critically on how equivalent the two groups actually are. Since equivalence cannot be guaranteed in this type of design, some purists insist that they be called **non-experimental designs**; although in practice many forms of quasi-experimental design can allow relatively strong inferences in settings where true experiments would be impossible to achieve.

Quasi-experimental methods share some of the advantages of full experimental methods such as clarity, transparency and repeatability. However, as we have indicated, they have problems accommodating the politics and agency of human beings in work settings. And there are also other subtle problems with pre-test and post-test designs because changes over time may be a consequence of measurement itself (a **testing effect**). The first measurement may get respondents to reflect on their initial answers, and this can lead to them answering differently the next time – not because of the intervention itself, but because they have been measured before. Thus, the process of measurement itself becomes a kind of intervention, but one that cannot easily be directly assessed. Returning to Figure 4.2, the aspiration of quasi-experimental methods is to conduct research without affecting the people under study, which would locate it in quadrant A. But as we have seen, the necessary level of detachment is often difficult to sustain. Figure 4.3 shows four common forms of research design, in increasing order of sophistication.

The cross-sectional comparison

The first form of design (Figure 4.3a) is the commonest in business and management research, but also the weakest. It involves selecting a group of people who have experienced something that you are interested in (attended a training course, graduated from

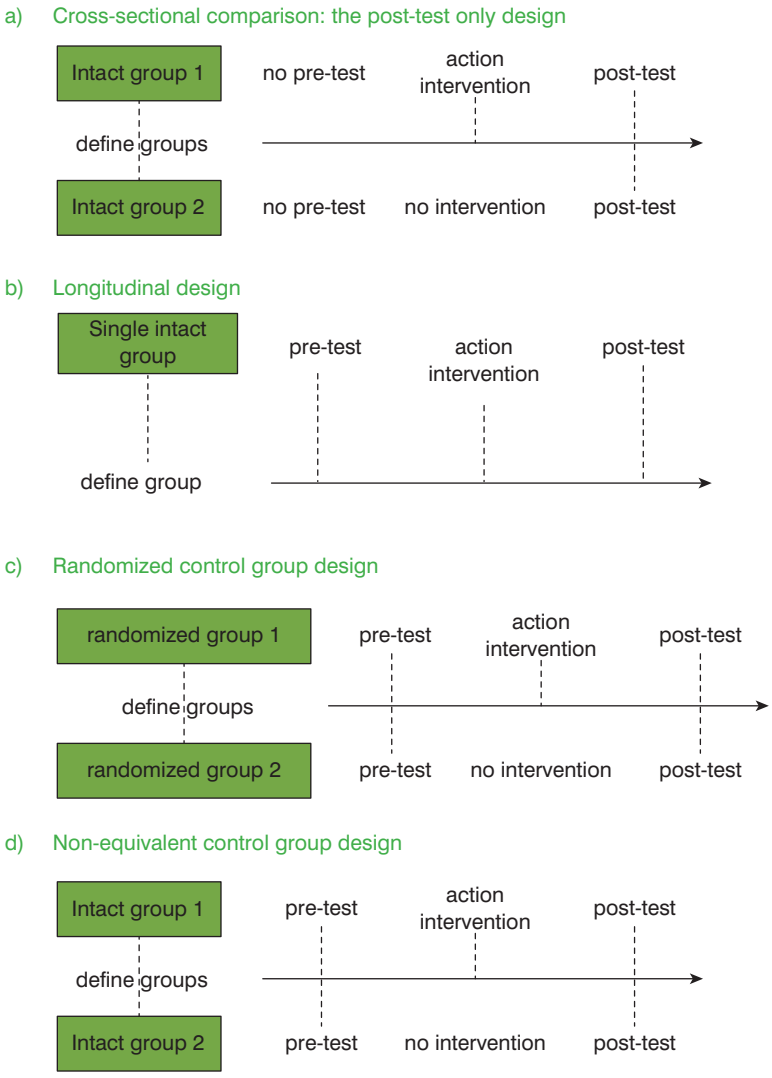


FIGURE 4.3 Common forms of quasi-experimental design

a prestigious university, worked with a charismatic mentor etc.), and then comparing that group with others who have not had the same experience. If we find a difference between the two groups on some variable that is theoretically interesting, it is tempting to jump to conclusions about causality. Although a great deal of the empirical literature relies on cross-sectional comparisons of this kind, it is obvious that no firm conclusion can be reached about cause-and-effect relationships unless groups being compared differ only on those variables (such that no other factor could be responsible for the observed difference in group means). Since this can never be guaranteed outside of the research laboratory, cross-sectional studies have only limited potential for advancing knowledge in a convincing way.

The pre-test/post-test design

It is a commonplace principle that when something changes, you have to measure before and after the change. Figure 4.3b shows a simple form of design with a single group measured twice, with an intervention of some kind taking place between the two measurements. The intervention might be the introduction of group exercises in a classroom, or a new form of recording product defects on an assembly line. In another example, the effects of a

leadership course on a group of managers might be evaluated by measuring the managers' attitudes and behaviour before and after the course.

What interpretation could be placed on a change in means from pre- to post-test? The most obvious question to ask is whether the same change would have been observed if the intervention had never taken place. Change could reflect the influence of some other event that took place between pre- and post-test (this is called a *history* effect); or *maturation* in respondents (due to them growing older, wiser, stronger or more disillusioned) in ways that have nothing to do with the intervention. Both history and maturation effects are those that would have happened anyway, with or without the researcher's presence.

More subtly, changes over time may be a consequence of measurement itself (a *testing* effect), arising from effects of the first measurement in leading to respondents reflecting on their answers for the first time. This may then lead to them answering differently the second time, not because of the intervention itself, but because they have been measured before. Thus, the process of measurement itself becomes a kind of intervention, but one which cannot easily be directly assessed. If there is a danger that the properties of a measuring instrument change if it has been answered before, it may make sense to change the instrument from pre- to post-test. However, this brings new problems unless it can be guaranteed that the instrument, though different, measures the same thing in the same way (the *instrumentation* effect).

The randomized control group design

The classical way to deal with history and maturation as threats to the internal validity of a design is by using a control group of individuals who are the same as the treatment group in every way except that they do not receive the intervention (Figure 4.3c). The way in which this is achieved is by randomization in allocation to groups. The consequence of randomization is that the whole of the prior history of individuals is detached from the intervention itself: individuals either receive the intervention or not based on a criterion (tossing a coin, or the equivalent) which is quite separate from any characteristics on which they differ. Any change in the treatment group which is not paralleled by an equivalent change in the control group is thus attributed to the intervention, on the basis that the control group shows what the treatment group would have been like but for the intervention. History effects and maturation effects will show in changes in the control group, and the effect of the intervention can be seen by comparing changes in the two groups.

The non-equivalent control group design

A weaker form of research design, which captures some of the strength of the randomized control group design, is the non-equivalent control group design (Figure 4.3d). Although the formal specification of this design looks similar to design (c), the crucial difference is that individuals are not allocated randomly to the treatment group and the control group, but rather allocation takes place on some other criterion, usually by working with intact groups. As a result, the validity of inferences from this type of design depends critically on how equivalent the two groups actually are. The non-equivalent control group design is the simplest of a family of research designs, which are termed 'quasi-experimental designs'. Since equivalence cannot be guaranteed in this type of design, some statisticians insist that they be called non-experimental designs; though the reality is that many forms of quasi-experimental design can allow relatively strong inference in settings where true experiments would be impossible to achieve. (It is also the case that even where a true experiment is possible, the phenomenon being tested is not of interest because its generalizability to the real world is questionable.)

There are substantial problems when using this form of design in real organizations. For example, the design assumes that 'nothing' happens to the control group during the period that the treatment (for example, course attendance) is being given to the experimental group.

This is a naive assumption, as Easterby-Smith and Ashton (1975) found, when attempting to evaluate a project-based management development programme held at Durham University Business School. While the 'chosen few' were away on the course, several members of the control group seized the opportunity to improve relationships with their bosses and strengthen their political standing in the company, thus harming the career prospects of a number of managers who had attended the course.

EXERCISE 4.2

Designing a call centre

Call centres are important for many companies for lots of reasons. First, they are often the primary point of contact between the company and its customers, and so the call centre staff *are* the company as far as customers are concerned. The company's reputation is in the hands of those people. Second, call centres have to be accessible 24 hours a day for global companies operating in many time zones and in many languages. Finally, the turnover of staff in call centres is very high, and can be over 90 per cent per year.

People who run call centres therefore have to pay attention to recruiting the best people and treating them well so that they stay with the company. There is some truth in the saying that a happy worker is a productive worker (not always, but often); and it is not a coincidence that First Direct (the online banking arm of HSBC) has year after year won awards for the quality of its customer service and also for how it treats its employees.

The task

You have been contacted by a company seeking to set up its own call centre, and your task in this exercise is to design a study to identify which factors are important for the company to consider in the design of its call centre. Consultants have identified the following general factors as being important:

- density of workstations
- layout of the space
- ergonomics of the workstation (the chair, the desk, lighting etc.)
- colour of the walls
- the view – is there a view out of the windows?

Discussion questions

What primary questions/hypotheses would you investigate?

Based on the list of factors identified by the consultants, how would you go about designing a study for the most effective call centre?

Validity of positivist experimental designs

Experimental methods are particularly concerned to ensure that results provide accurate reflections of reality. They distinguish between internal and external validity, with the former relating to *systematic* factors of bias and the latter being concerned with how far the conclusion can be generalized across other types of person, settings and times.

The aim of experimental designs is to maximize **internal validity**, and this requires the elimination of plausible alternative explanations for any differences observed between

groups. That is why full experiments require random assignment to control and experimental groups, and efforts are then made to ensure that the subsequent experiences of the two groups are identical in all respects, except for the focal variable. But there are many threats to internal validity, including history (experiences of the two groups diverge in some unexpected way), maturation (group members get older or other life changes take place) or mortality. The latter can be a problem in medical research where people literally die before the experiment is completed, and in organizational studies people may vanish from the research because they move jobs, leave the company or lose interest. Threats to internal validity are systematic rather than random, and they tend to focus on factors that cloud the interpretation of differences between groups in change over time.

External validity is about generalizability of results beyond the focal study. In the physical sciences we assume that Newton's laws of motion will have equal validity whether applied in New York, Bogota, Xi'an or on Mars. If they do not apply in the same way everywhere, then there should be a clear way of understanding how they vary in different circumstances. Thus Albert Einstein predicted through his theory of relativity that bodies do not follow Newton's laws when they are moving at relative velocities near to the speed of light, and that light does not travel in straight lines when subject to strong gravitational forces. The latter prediction was confirmed by observations of the total eclipse of the Sun in 1919 (see Example 4.1).

EXAMPLE 4.1

Relativity and the 1919 total eclipse of the Sun

Probably the most important eclipse in the history of science occurred on 29 May 1919. Just six months after the end of the First World War, British astronomers used it to test a new idea that came from Germany in 1915. Expeditions of astronomers photographed the eclipse in difficult tropical conditions in Brazil and on the African island of Principe. At the time, the Sun was in front of a useful cluster of stars, the Hyades. The astronomers compared the relative positions in the sky near the Sun with the positions of the same stars as previously photographed in the night sky.

The proposition was that gravity affected light, space and time itself, and as a result the Sun would deflect starlight passing by it. Changes in the apparent location of stars in the sky, seen close to the Sun during a total eclipse, could confirm the idea. It was predicted that for stars almost in line with the Sun, the shift in apparent position would be slightly less than two seconds of arc, or a few ten-thousandths of a degree. The 1919 measurements confirmed that the Sun bent the light rays by roughly the right extent – less than predicted in Principe, more than predicted in Brazil.

Source: www.esa.int/esaSC/SEM7I9R1VED_index_0.html.

However, in order to demonstrate external validity, management research designs need to demonstrate a number of features. First, they need to demonstrate that the results observed are not just a product of the selection of individuals or organizations. Sometimes the people who volunteer to take part in research are open minded and keen to help; sometimes they will put themselves forward because they have strong opinions or 'axes to grind'. Managers will often 'volunteer' employees to take part in research because they believe they will show the organization in a positive light, and will not offer individuals who are likely to be critical; or they will allow access to the organization because they hope research will add legitimacy to a new innovation or policy that they wish to promote. These issues of access and sampling are discussed further in later chapters (Chapters 5 and 8), but the key point to remember is that selection should avoid sources of bias as far as possible.

Other threats to external validity come from the setting and history. In the first case, the results of a piece of research in the health service may be difficult to generalize to an automobile manufacturer. Similarly, research conducted in large organizations may not apply

to small organizations; and there is also increasing awareness that research conducted in one national setting may not apply to other national settings. With regard to history, it is important to note that patterns and relationships observed in one era may not apply in another era. For example, at the present time, with the emergence of new economies such as China and India as global economic forces, the theories about the behaviour of financial markets that were developed during the era of US dominance are now having to be rewritten. Similarly, a study on reward systems conducted in one country where the supply of skilled labour is plentiful might not be relevant in another country where there is a marked shortage of skilled labour.

SURVEY RESEARCH

The dominant epistemology underlying survey research methods is positivism. As explained in the previous chapter, this assumes that there are regular, verifiable patterns in human and organizational behaviour, although they are often difficult to detect and extremely difficult to explain due to the number of factors and variables that might produce the observed result. Consequently, survey research tends to use cross-sectional designs with large samples, which enable multiple factors to be measured simultaneously and hence potential underlying relationships to be examined.

There are three main types of survey that seek to take a detached viewpoint: factual, inferential and exploratory studies. **Factual surveys** can also be used in an engaged way, where they are established as survey feedback processes. The three main types generally fit with quadrant A of Figure 4.2, whereas engaged factual surveys fit into quadrant D. We will briefly describe and illustrate each type here. More detailed information on the technical design of surveys can be found in Chapter 8.

Factual surveys are most often associated with opinion polls and market research, and involve collecting and collating relatively 'factual' data from different groups of people. Thus, in order to assess market share or loyalty we might be seeking to identify what percentage of the population of Manchester entered either an Aldi or a Tesco supermarket at least once in the previous week. This is reasonably factual data, which could be gathered by a postal questionnaire or structured interviews; however, it could be affected by people's ability to recall what they did in the previous week, and possibly by **social desirability** factors where they claim loyalty to one supermarket over another in order to project a particular image of themselves to the researcher.

A common use of factual surveys within companies is **survey feedback**. This involves distributing a questionnaire to all employees asking for their views on the management of the organization, the quality of communications, their job satisfaction, feelings of security and so on. Data is then aggregated by department, level or job category and results are discussed in public. This puts pressure on management to change systems, procedures and their own behaviour in order to get 'better' results next year.

Inferential surveys predominate in academic management research, particularly in the fields of strategy, marketing and organizational psychology. They are aimed at establishing relationships between variables and concepts, whether there are prior assumptions and hypotheses regarding the nature of these relationships. Inferential surveys generally assume an internal realist ontology, although epistemologically they involve a weaker form of positivism than experiments. The usual starting point for inferential surveys is to isolate the factors that appear to be involved, and to decide what appears to be causing what. This means that researchers have to identify the main **dependent variables** and **predictor variables**:[1] it is the latter that are assumed to be causing the former.

[1]The term **independent variable** is often used instead of 'predictor variable'. We prefer the latter term because in practice, even independent variables tend to be related to each other, and therefore use of the former term is misleading.

In Figure 4.4, we are suggesting (hypothesizing) that the predictor variables of salary, leadership and responsibility have an impact on the dependent variable, motivation at work. In order to test this hypothesis it would be necessary to define ways of measuring each of these variables, generally through a small number of items in a questionnaire, and this would need to be completed by a sample of employees in one or more places of work. Naturally, this requires that the measures of the four variables are accurate, and that the sample is appropriate in terms of size and constitution in order to test the hypothesis; we discuss how to do this in more detail in Chapter 8. Moreover, the four factors identified in Figure 4.4 could be examined in more detail. For example, one might be interested in the interactions between some of the variables, such as whether some forms of leadership result in greater responsibility being distributed around the workforce; or one might be interested in whether some of the arrows might work in other directions, so that a highly motivated workforce would lead to particular styles and strategies of leadership. These are some of the relationships that can be analysed, particularly with more complex models, through the use of structural equation modelling (see Chapter 10). Studies of this kind are often known as **cross-sectional surveys** because they involve selecting different organizations, or units, in different contexts, and investigating how other factors, measured at the same time, vary across these units.

Similarly, Lyles and Salk (1996) were interested in the conditions that led to greater transfer of knowledge from foreign parent companies into international joint ventures. So they selected a sample of 201 joint ventures that were regarded as small or medium-sized across four manufacturing industries in Hungary. Through comparing indicators of performance across the whole sample, they were able to conclude first that there was a strong link between knowledge transfer and performance, and second that this transfer was most likely to take place when the foreign and domestic parents had equal (50/50) equity stakes in the new venture. The sample size of 201 was sufficient for them to demonstrate that the results were statistically significant, but one of the key problems for researchers using cross-sectional designs is to know how large the sample needs to be.

Geert Hofstede's (1980/1984) study of national cultural differences provides an example of an exploratory survey. He attempted to develop a universal set of principles against which any culture can be measured, in the hope that this would provide a basis for predicting the behaviour of individuals and organizations in almost any country. However, he did not start with an explicit set of hypotheses; rather, he had a large number (about 216,000) of questionnaires completed by employees of IBM with regard to their views and values, and he was looking for patterns in the data. His four dimensions – power distance, individualism/collectivism, uncertainty avoidance and masculinity – emerged from his data, and by demonstrating that they fitted reasonably well with prior research into this topic, he was able to substantiate his claim to the importance of these four dimensions.

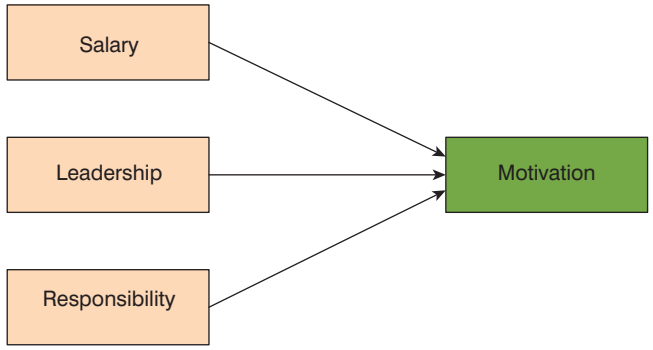


FIGURE 4.4
Possible predictors of motivation at work

EXAMPLE 4.2

Which variable is which?

Chen's (2008) doctoral thesis examined how workers in virtual teams doing different tasks used instant messaging to support their work. Her respondents were all Chinese-language speakers in China or Taiwan. Some of them worked with people in the same location (building or city), while others worked with colleagues in other countries (such as Canada, USA, Germany).

- Instant messaging is not a variable because everyone in the study used it; but other communication technologies (such as email, videoconferencing and face-to-face meetings) are variables because some people use them while others do not. Similarly, dual-language use (Chinese and English) is not a variable because the whole sample spoke both languages.
- The focus of the study was on what led people to switch communication media or languages, and this therefore was the *dependent* variable. The *predictor* variables were task characteristics, indices of relationship quality (how well people knew each other), and whether they worked in the same or different locations.

Principles in judging the quality of a sample design

When decisions are made that can have significant consequences for people it is important that those decisions are based on evidence. The trustworthiness of the evidence base for decisions depends on many factors. We begin with the difference between a **population** and a **sample** drawn from that population. The term 'population' refers to the whole set of entities that decisions relate to, while the term 'sample' refers to a subset of those entities from which evidence is gathered. The **inference** task then is to use evidence from a sample to draw conclusions about the population. Samples are very commonly used, both for research and for policy-making.

Companies estimate morale among employees from quantitative surveys of samples of employees (either by deliberately picking names from the payroll, or by letting employees choose whether or not to respond). Here, the population is the whole set of employees within the organization, and management want to gain a picture of their staff by summarizing the results from a sample of those who respond to the survey. For example, companies such as Gallup and Towers Watson are leaders in carrying out such opinion surveys, and one of the benefits that they claim is that they can compare the profile of a company (and sub-units within it) with that of other companies where the same survey questions have been used. This gives a method of benchmarking, which senior management often find very informative and powerful in shaping their strategy.

Many companies also use small-scale surveys repeated frequently (perhaps once a month) as a kind of barometer of how employees or customers are feeling. Companies may set up a panel to represent the entire body of employees, and their responses are used to monitor the state of the organization at that time. This can be very valuable during the course of a period of major change. During the UK general election in 2010, opinion polls appeared every day monitoring the state of the electorate and their response to the offerings of the political parties.

The UK government uses sample surveys of companies on their annual pay negotiations in order to get an estimate of wage inflation in the economy. In most countries, estimates of activity in the labour market (whether people are unemployed, self-employed, economically inactive etc.) are derived from samples and not from complete counts.

Researchers who are interested in the rhythms of working life often use diary methods to study what employees are doing during their working day and how they feel about those activities. This is often done by using a bleeper (or a mobile phone) set to go off at specific

times, and individuals are asked to answer a set of specific questions. Here again, sampling is involved since it is impossible to collect information continuously (unlike, for example, a heart rate monitor which someone might wear). The researcher will generally want to consider a 'typical' working day, or perhaps compare responses obtained while carrying out different kinds of work activities. In this example, the working day could be considered as the population, and the bleeps could be considered as samples of activities during the day.

Sometimes, a research project involves collecting evidence from every member of an organization (or every member of a project group), but more often this is not the case and the researcher needs to decide on a sampling strategy. A sample might be a proportion of employees in an organization, a selection of companies operating in a specific market, a number of specific periods of time for assessing the quality of customer responses in a call centre, or a selection of transactions for audit purposes. For each of these examples, the researcher has to make a decision about what the sampling unit is (the person, the company, the transaction etc.), how many sampling units to take, and on what basis sampling is to be undertaken.

Generally speaking, the purpose of collecting data from a sample is to enable the researcher to make statements about a larger group that the sample is drawn from. Many people say that they only want to talk about their sample data, but we almost never believe them. Even if there is no formal generalization beyond the sample, both the writer and the reader are interested in what this study's findings tell us that would be useful when thinking about other settings. This places a responsibility on researchers, whether we like it or not, to say where our findings might be relevant.

The claims that can legitimately be made from sample data depend absolutely on the relationship between the sample and the population (the larger group that the researcher wants to make claims about). Get the sampling wrong, and accuracy in calculating the results is of no consequence. The opposite is also true: get the sampling right, and even rough-and-ready calculations or 'eyeballing' of the data can be immensely valuable. There are two basic principles that underlie decisions about sampling design: **representativeness** and **precision**.

Representativeness in sampling

The accuracy of conclusions drawn from a sample depends on whether it has the same characteristics as the population from which it is drawn. If the sample is systematically different in some way, then the sample is biased. A simple definition of bias in sampling is that it occurs when some members of the population have a higher chance of being included in the sample than others. There are two steps involved in defining a sampling design: first, to draw up a **sampling frame**, a list of all who are eligible to be included in the study; second, to achieve a valid response from all those included in the sampling frame. **Bias** can be introduced into a sampling strategy in many ways, through choices made in the design of the study itself and also through features of the process of collecting research data:

1. *Exclude groups of people*: leave out home-workers, casual staff, new recruits, part-timers. If your study involves selecting people to interview, it is tempting to prefer people who are articulate and have 'interesting' stories to tell, or people who have been cooperative in previous studies within the organization.
2. *Distribution method*: send out a questionnaire survey using an out-of-date list of mailing addresses to exclude those who have recently moved; distribute an invitation to interviews or a focus group by email to exclude those without a PC on their desks.
3. *Language used*: use English to exclude those who don't speak English, and introduce biasing factors for those who do because of differences in how well people can use the language.

One of the key ways of judging the representativeness of a sample is to compare the characteristics of the sample to those of the population; this kind of information is commonly reported in published papers that are based on quantitative surveys but less so in interview

studies. Even if the sampling frame accurately represents the population, non-response is a major source of problems in getting outsiders to believe the results. In itself though, non-response is not a problem, as long as those who do take part in the study have similar characteristics to those who do not. Of course, there is usually a big problem in assessing whether this is true, because (obviously) you do not have data from those who did not take part. It is sometimes possible to get some idea about potential bias due to systematic non-response by comparing those who respond quickly to a questionnaire survey or an invitation to attend an interview with slow responders on demographic variables and also on the central variables in the research. The idea is that the slow responders will be more similar to non-responders than are those who reply quickly. If the slow responders have similar characteristics to rapid responders, then the researcher can have greater confidence that non-responders would also have been similar, and this helps to build credibility in a study.

Precision in sampling: sampling proportion and sample size

Precision is about how credible a sample is. For example, opinion polls conducted before the 2014 general election in India used samples that are very much smaller than the population of registered voters. How confident is it possible to be about predicting election outcomes from opinion poll samples? Does the precision of the estimate depend on how big the sample is? It seems plausible that it should. Does it also depend on what proportion of the population the pollsters talk to? If the number of electors is 10,000 then a sample of 1,000 (i.e. 10 per cent) might be OK. However, many people would be less happy with a sample of 1,000 if the number of electors was one million instead (0.1 per cent). Intuitively, this proportion seems too small. However, the first intuition is correct, but the second is not.

Nguyen (2005) provides a clear and graphic example of cooking chicken soup to show why the size of a sample matters but how big a proportion the sample is of the population (the **sampling proportion**) does not. Consider three scenarios: cooking at home for four people using a small pot; cooking for a dinner party with 12 guests using a medium-size pot; and cooking a banquet for 200 wedding guests using a huge pot. Regardless of the number of guests, the only way to tell if there is enough salt in the soup is to taste it. The way to find this out is: first, stir the soup so it is well mixed and is the same all the way through, and second, use a tablespoon to draw off some soup. A tablespoon will do because there is no point taking more than that: taste it all and there is no soup left for the guests; taste more spoonfuls and each will taste just the same, so nothing is learned. It is not necessary to use a large ladle to sample from the large pot simply because the pot is bigger, or a tiny spoon to sample from the small pot because the pot is smaller: the same sized tablespoon is enough to judge the adequacy of the seasoning regardless of how big the batch of soup is, as long as the pot is stirred first.

The soup in the pot is the *population*; the spoonful to taste is the *sample*. The size of the spoon is the **sample size**, and that is what matters. The cook needs to taste enough soup to be able to make a judgement about the pot as a whole. Now apply these principles to the task of making judgements about attitudes in a society towards an issue of concern. Consider the question of whether organizations should aim to maximize their profit or should consider their social responsibilities. The precision of the answer to this question has nothing to do with the size of the population but rather depends on the size of the sample. Small samples will always be less precise than large samples.

Combining precision and representativeness to achieve a credible sample

We have looked at the two design principles of bias and precision, and clearly both are important in achieving a credible sampling design for a quantitative research project. Low bias means that conclusions from a specific sample can reasonably be applied to a larger population, and high precision means that the margin of error in the claims that are made will be low – the researcher can expect to be precisely right (see Table 4.1). However, high precision is no way of saving a study where the sample is biased (the claims the researcher makes are precisely wrong). Giving very precise answers to the wrong question will not

endear a researcher to his or her supervisor, just as it does not help in getting high marks in an examination! Most projects carried out by students (at whatever level) are a compromise in some way, simply because resources are limited. As a result, there will always be trade-offs when it comes to decisions about design. Is it better to have a large sample (giving higher precision) if the cost of achieving it is to introduce bias into the sample achieved? Put more simply, which is better: to be imprecisely right or to be precisely wrong? In our opinion, the answer is straightforward. Imprecisely right is better: it is preferable to have a sample that properly represents the population even if the precision is lower because of a small sample.

TABLE 4.1 Principles in designing a sample

		Bias	
		High	**Low**
Precision	High	Precisely wrong	Precisely right
	Low	Imprecisely wrong	Imprecisely right

Probability sampling designs

This section describes forms of sampling design where the probability of each entity being part of the sample is known. Some sampling methods have the same probability for every entity in the sample, while others have the same probability within segments of the design but differing probabilities across segments.

Simple random sampling

With **simple random sampling**, every sample entity (company, employee, customer etc.) has an equal chance of being part of the sample. In the past, this was done using printed random number tables. Now computers are used for this, and it is easy to draw up a list of random numbers as a basis for selecting a sample.

Stratified random sampling

One drawback of simple random sampling is that it can mean that small but important parts of a population are missed altogether or sampled so little that the researcher cannot make confident statements about them. For instance, customer surveys of a healthcare facility would be badly served by a simple random sample. Most users of a healthcare facility have relatively minor ailments and perhaps visit only once or twice in a year. There will, however, be a small number of patients with major health problems whose treatment is perhaps extensive. It is quite reasonable to expect that a sample should be informative about the chronically ill minority as well as the occasionally ill majority. The way to achieve this is to divide the population up into homogeneous groups called *strata*, and then take a simple random sample within each stratum. **Proportional stratified random sampling** has the same sampling proportion within all strata; but this has the disadvantage that rare groups within the population would be badly represented. The way to deal with this problem is to take a larger proportion of sample units in small strata, and a smaller proportion in the larger strata. This is called **non-proportional stratified random sampling**.

Systematic random sampling

Systematic random sampling relies on there being a list in some form or other of the units in the population that the researcher is interested in. This might be a customer database,

or a list of employees of a company or students registered in a university. Suppose that a researcher wants to achieve a sample of 500 students in order to assess their satisfaction with the virtual learning environment (VLE) system that a university has just introduced. If there are 20,000 students, then 500 represents a sample of 2.5 per cent, corresponding to selecting 1 in 40 students from the population. This proportion could be achieved by choosing a number at random between 1 and 40. If that number were 27, then the researcher would go down the list taking every 27th student in order to derive a sample list of 500 names. What this process relies on is that the population list is essentially organized randomly, so that picking in this systematic way does not introduce bias. There could be a problem if the list is ordered alphabetically by individuals' last name, since all those students with the same name will be listed together and individuals with the same name will have less chance of being selected than if the list were randomly ordered.

Cluster sampling

Any method that involves random sampling will lead to practical problems where the population units are spread very widely, such that the cost of approaching them is very high. **Cluster sampling** deals with this by first dividing up the population into what are called 'clusters', and then by sampling all the units within the selected clusters. A study of company success in emerging economies might first identify all the countries where the company operates, select randomly a number of those countries to study in detail, and then approach all the relevant contacts in those countries. This would allow the study to use local research staff who are familiar with the language and culture of each country.

Multi-stage sampling

Multi-stage sampling combines together the methods described above in order to achieve higher operational and technical efficiency. For example, **stratified random sampling** divides the population into strata and then samples from within all of them. Instead, a study might use a sampling approach at each level, and this is very common in large-scale social research. Suppose there was a national change in the organization of schools to give greater management autonomy to head teachers, and researchers wanted to know whether this change had any effect on the performance of students in schools. It would be very inefficient to select students at random, even if the research team actually had a national database of all school students. It makes more sense to divide the country up into regions, select some regions for detailed attention, identify all the schools in the targeted regions, and then select a sample of schools. Having defined a sample of schools within selected regions, the same process could be used to sample classes within the selected schools, or perhaps take a sample of students from all of the classes in a selected school. In this example, the criterion of randomness applies at each of several stages in the design of the study; hence the name 'multi-stage sampling'. The aim is to balance the need for representativeness of the sample with the highest possible cost effectiveness.

Why are probability sampling designs valuable?

It is only with probability sampling that it is possible to be precise about the relationship between a sample and the population from which the sample is drawn. Knowing what this relationship is allows the researcher to make a firm judgement about the relationship between characteristics of a sample and characteristics of the population from which the sample was drawn. All forms of **probability sampling design** have this feature in common: it is always possible to state the probability of each individual respondent being selected for inclusion in the research study. The statistical theory behind the inference process (often called 'significance testing' – we discuss this in Chapter 10) relies fundamentally on sampling based on probabilities. There is always uncertainty about any claim made from data, but using probability sampling designs can allow the researcher to quantify that uncertainty. Thus, sample data can tell us whether the risk associated with using a mobile phone while driving is higher, but only a properly constructed sample design can allow us to decide how worried we should be about that extra risk.

Non-probability sampling designs

Non-probability sampling designs all share the same characteristic: that it is not possible to state the probability of any member of the population being sampled. As a result, it is harder for the researcher to be confident that claims made about the sample can apply to the larger group that the sample is taken from.

Convenience sampling

Convenience sampling involves selecting sample units on the basis of how easily accessible they are, hence the term 'convenience sampling'. A student who uses a list of friends on Facebook for their dissertation is taking a convenience sample. Such a sample may well reflect the individual's own social network, but is clearly not representative of students as a whole or of the population of the UK. This may not matter: it depends what claims the researcher wants to make. Facebook may be the medium of choice for Generation Yers and their social networks, but is much less common for Baby Boomers. Convenience samples are very common in research because they are – well – convenient! It is impossible to guarantee that any sample achieved in this way represents a specific population that may be of interest. However, they can still have a value. It rather depends on what the purpose is for collecting data. For a very long time, people thought that all swans were white because no one had ever seen one of any other colour. It only takes an Australian researcher with a convenience sample of one black swan to prove the old generalization to be wrong.

Quota sampling

Quota sampling divides the relevant population up into categories (perhaps male/female, or country of origin for students) and then selection continues until a sample of a specific size is achieved within each category. The aim is to make sure that each of the categories is represented according to the quota proportions. For example, in doctoral research on whether the Internet empowers consumers, quota sampling enabled the researcher to ensure that she had users of a variety of ages, while a convenience sample would be more likely to result in a preponderance of people similar to the researcher and her friends.

Purposive sampling

In **purposive sampling**, the researcher has a clear idea of what sample units are needed according to the purposes of the study, and then approaches potential sample members to check whether they meet eligibility criteria. Those that do are used, while those that do not are rejected. The guiding principle for sampling might be theory (theoretical sampling), and the basis of sampling could therefore change as a study is carried out if analysis of initial data is used to guide theory development.

Snowball sampling

Snowball sampling starts with someone who meets the criteria for inclusion in a study who is then asked to name others who would also be eligible. This method works well for samples where individuals are very rare and it is hard to identify who belongs to the population. Dissertation students often do this by starting out with people they or their supervisor know personally, and then ask those people to pass them on to others who would also be suitable. It works well too for individuals, groups or companies that are part of networks whose membership is confidential (LinkedIn is a useful resource here, and its owners actively promote the professional network as a useful way of finding interesting people to get in touch with).

Why are non-probability sampling designs valuable?

The sampling approaches described in this section are answers to a variety of practical problems that researchers have encountered in carrying out their work. But how do these

sampling methods stack up against the key quality criteria of bias and precision that we started this section with? *Precision* is most straightforward, since its main focus is the size of the sample achieved. A convenience sample can meet the first requirement of a big-enough sample most easily; quota sampling and purposive sampling both aim to ensure that every sector in a sampling design is filled, while snowball sampling addresses the problem of ensuring an adequate sample of hard-to-find people.

However, the principle of *bias* is where non-probability sampling methods can most easily fall down, especially for convenience sampling. Many management researchers have been seduced by the lure of large samples for quantitative studies (perhaps feeling themselves under pressure from journal editors to report large samples) and achieved them simply through collecting respondents by any means (MBA classes, lists of Facebook friends etc.). There is no guarantee that the findings reported are credible, since the credibility of findings relies in large measure on the character of the sample. It is not surprising then, that reviews of research often highlight contradictions in findings between different studies, given that researchers often take little care in defining their sampling design. Qualitative researchers are not immune to issues of bias, and can be seduced by the richness of data from small samples. Rich data can be powerful persuaders, but every research study involves a sample, and it is vitally important to the credibility of results that the researcher considers how the sample used sits within the larger group from which the sample is drawn.

Validity in survey research

Since *survey designs* are informed by internal realist ontology, the issues of validity are reasonably similar to those of strong positivist studies. Thus, there is a major concern about whether the instruments and questionnaire items used to measure variables are sufficiently accurate and stable. Most of this is done through pre-testing instruments before the actual research is carried out, and hence measures of **reliability** are important because they assesses whether an instrument will produce the same score for each occasion that it is used. There is also the question of external validity: whether the patterns observed from the sample data will also hold true in other contexts and settings. And again, the technicalities of assessing reliability and validity with survey data will be discussed further in Chapter 9.

EXERCISE 4.3

Which supermarket gives the best deal?

This exercise examines the methods used by different UK supermarkets to answer the question 'Who gives you the best deal?'.

Which?

Which? uses a basket of 100 popular branded food and grocery products sold by four online supermarkets, and it calculates the average price of each individual product for each month. It then compares the prices for each supermarket to answer the question 'Who gives you the best deal?'.

MySupermarket.co.uk

MySupermarket is an online comparison site which shows the cost of items at Tesco, Asda, Sainsbury's, Waitrose and Ocado. It has recently included Aldi in its list of supermarkets, as a result of the increase in popularity of the low-cost supermarket

(Continued)

(Continued)

in the UK. Since Aldi's range consists mostly of own-label products, MySupermarket compares them to the equivalent own-label products at other retailers.

Waitrose price match promise

Waitrose selects 1,000 everyday branded products from their food and drink grocery ranges, and checks the prices against the online price at www.tesco.com for a post-code served by large Tesco stores (excluding Tesco Metro or Tesco Express). The standard individual unit selling prices are matched for identically branded products; but this does not include multi-buy promotions.

Tesco price promise

Tesco check the price of a shopper's basket of items against the prices at its three main supermarket competitors. If the comparable grocery shopping would have been cheaper there, then they give a voucher for the difference. There must be at least ten items in the basket in order to qualify for the comparison.

Sainsbury's brand match

Sainsbury's check the price of an identical basket of branded products (same size, flavour, colour etc.) at Asda. If the identical basket bought on the same day would have been cheaper at either of those supermarkets, then they give the shopper a coupon at the till for the difference.

There are several different sampling issues involved when trying to decide who gives the best deal, for example:

- *Sampling of supermarkets* – where do you buy your products? Each supermarket compares itself against selected others in order to persuade the consumer that they offer the best deal. What are the implications of the different alternatives?
- *Sampling of items* – what products do you buy? Each of the comparisons takes a different approach to deciding what to compare. Which option (if any) do you find most convincing as giving an answer to the question 'Who gives you the best deal?'.

CONSTRUCTIONIST RESEARCH DESIGNS (QUADRANTS B AND C)

Constructionist research designs are linked to the relativist and nominalist ontologies. They start from the assumptions that verifiable observations are potentially subject to very different interpretations, and that the job of the researcher should be to illuminate different truths and to establish how various claims for truth and reality become constructed in everyday life. Hence it is not surprising that there is a wide range of methodologies which fit within the constructionist paradigm. Here we cover some of the main methodologies – action research and cooperative inquiry, archival research, ethnography and narrative methods – which are primarily based on constructionist designs. In the subsequent section we will look at methods that often bridge the epistemological divide, notably case method, grounded theory and so-called mixed methods.

Action research and co-operative inquiry

One of the key assumptions of positivism, and of natural scientific methods, is that the researcher should be objective, maintaining complete independence from the object of study. In the social sciences, where claims of the researcher's independence are harder to sustain, many people have tried to turn this apparent problem into a virtue. This is the tradition of **action research**, which assumes that social phenomena are continually changing rather than static. With action research, the researchers are often part of this change process itself. The following two beliefs are normally associated with action research designs:

1. The best way of learning about an organization or social system is through attempting to change it, and this therefore should be an objective of the action researcher.
2. The people most likely to be affected by, or involved in implementing, these changes should as far as possible become involved in the research process itself.

Some forms of action research appear to follow the principles of positivism, for example by attempting to change the organization from the outside and then measuring the results. Kurt Lewin (1948), who originated the action research tradition, used experimental designs when investigating the efficacy of different ways of getting housewives to change their nutritional habits during the Second World War (see Example 4.3).

EXAMPLE 4.3

An early example of action research

During the Second World War, Kurt Lewin and his associates experimented with groups of American housewives to see if they could be persuaded to serve unpopular types of meat, such as beef hearts, sweetbreads and kidneys, to their families. They used two methods to try to persuade them to change their habits. In the first case, a lecturer gave an attractive talk to members of the group on the dietary and economic value of using these meats, and offered some good recipes. In the second case, the same information was provided to the housewives, but they were invited to discuss the information and at the end to indicate by a show of hands whether they intended to serve the new meats. In a follow-up study it was found that only 3 per cent of the members of the lecture groups served one of the meats, compared with 32 per cent for the discussion/decision groups.

Similar results were obtained when persuading mothers to give orange juice and cod liver oil to their infants, although in these cases the discussion/decision method was only found to be twice as effective as the lecture method.

Source: Krech et al., 1962: 229–30

Kurt Lewin's studies were, however, different from traditional experimental research because there was an emphasis on changes in behaviour, and his housewives were active participants in deciding whether or not to change. The weakness with his initial experimental design was that it did not allow him to discover *why* the changes took place. This required subsequent experiments and qualitative studies in order to build up deeper understanding of why behaviour changed under these different circumstances. Given the strong emphasis on intervention as part of the research process, and the focus on debate and discussion, these later studies fit within quadrant C in Figure 4.1.

Involvement in the research process is taken a stage further in what has come to be known as **co-operative inquiry** (Reason, 1988; Heron, 1996). This has been developed for researching human action mainly at individual and community, rather than at organizational levels.

It starts with the idea that all people have, at least latently, the ability to be self-directing, to choose how they will act and to give meaning to their own experiences. It fits with stronger versions of constructionism, and rejects traditional positivist methods where people are studied as if they were objects under the influence of external forces. Cooperative inquiry not only focuses on the experiences and explanations of the individuals concerned, it also involves them in deciding in the first place what questions and issues are worth researching. Thus the 'subjects' become partners in the research process.

A study of the development of Afro-Caribbean entrepreneurs in the West Midlands adds a further dimension. Ram and Trehan (2010) have worked for five years with a group of eight entrepreneurs. The group meets on a monthly basis and determines its own agenda, and more importantly controls its own membership. The primary goal for the entrepreneurs is to gather ideas and develop strategies from their interaction with other owners, which will enable them to grow their businesses (none of the businesses are in competition with each other). The academics are present at these meetings in the roles of process consultant, facilitator and researcher; they are also given access to company documents and conduct interviews with group members. This approach, known as **critical action learning**, is driven largely by group members and takes place within a social and political context. The academics thus become partners in the problem solving of the group, and contribute particularly through surfacing the feelings and emotions of members.

Archival research

It is not always necessary to gather new data when conducting research. An enormous amount of data already exists in the public domain as corporate and government reports, and the statistical and financial databases that can be accessed online. Our focus here is mainly on textual information and its analysis, which we call **archival research**. Given the focus on words and on existing texts, this type of research fits with quadrant B in Figure 4.1.

Favoured sources of archival data in management and business research are the annual reports from companies where statements from chief executives review achievements from the past year and outline plans and priorities for the next year. Analysis of the language used over time can demonstrate, for example, the emerging concern among large companies about sustainability issues, or increasing emphasis being placed on employee engagement. By analysing policy statements produced by central and local government in the UK between 1997 and 2008, O'Reilly and Reed (2010) were able to follow the changing rhetoric about the desirable behaviour of public sector managers over this period (see Example 4.4).

EXAMPLE 4.4

An archival study

O'Reilly and Reed (2010) analysed 29 'high-level' government documents published between 1997 and 2008, which focused on the management of the public sector. They searched these documents for strings of text containing <profession>, <manag> and <leader>. By looking at the frequency of usage they were able to show how the discourse of leadership, which emphasizes change and reform, has started to take over from the older 'New Public Management' discourse of managerialism, which emphasized efficiency and performance.

Ethnography

The key principle of **ethnography** is that the researcher should 'immerse' him or herself in a setting, and become part of the group under study in order to understand the meanings and significances that people give to their behaviour and that of others. It is

thus a strong form of constructionism. Most outsiders who are new to an organization or group will encounter things that they do not understand. These are what Agar (1986) calls 'breakdowns': events or situations where the researcher's past experience gives no help in understanding what is going on. This breakdown therefore represents something unique about that organization, and which was previously unknown to the researcher. For example, most groups have 'in-jokes', based on experiences shared only by members of the group. In order for an outsider to make sense of the breakdown provided by an in-joke it will be necessary to track back to the original experiences (Roy, 1952; Collinson, 2002). The breakdown provides a kind of window into exploring aspects of the experiences and meaning systems of groups and organizations. It will only be possible to resolve the breakdown when the researcher has understood these meaning systems.

Another important distinction is between what are known as **emic** and **etic** perspectives. These two terms were first coined by the American linguist Kenneth Pike (1954): *emic* refers to the sounds within a language which can only be distinguished by speakers of that language; and *etic* refers to features of a language that are easily identified by outsiders, but are largely inaudible to people who speak that language. For example, the four Chinese tones are *emic* because they cannot easily be distinguished by a non-Chinese speaker, yet are absolutely critical to understanding the language. On the other hand, most native English speakers are unaware that their speech is seen (from an *etic* perspective) as tight-lipped and monotonic by French and Italian speakers.

The distinction has also led to the view that better insights can be gained into management and organizations through combining insider and outsider perspectives. Thus, Bartunek and Louis (1996) advocate methods which involve research teams that combine people working inside the organization with people working from the outside. Using methods like this the ethnographer has the opportunity to challenge and extend conventional wisdom, and to generate new insights into human behaviour. However, in many organizations it can be difficult to conduct full ethnographic research because of access restrictions, although it is often possible to combine observation of meetings with interviews of participants. We provide an example of a contemporary study in the aerospace industry in Example 4.5, which shows both the limitations on access and potential that researchers have to influence their informants. In Chapter 6 we develop the idea of participant observation further. The emphasis of involvement of the researcher in the research setting, combined with the strong constructionist element, locates ethnography within quadrant C in Figure 4.2.

EXAMPLE 4.5

A study of organizational decision making

Winston Kwon and colleagues (2009) were interested in the way managers used discourse and rhetoric in meetings to influence strategic decisions. They conducted interviews and observations over a two-year period with senior managers in a large aerospace company. Regular team meetings and conferences were observed over a six-month period and all key participants were interviewed before and after the observation period. Due to other commitments of the researchers and constraints imposed by the company they were not able to be present all the time, and hence access was intermittent. They were also asked to provide non-evaluative feedback about their observations to the company, and on several occasions this feedback led managers either to change, or clarify, decisions that they thought they had already made.

Narrative methods

Another group of constructionist research designs have been given the general label of **narrative methods** (Boje, 1995, 2001; Czarniawska, 1998; Daiute and Lightfoot, 2004).

These contain both ontological and epistemological elements. The ontological view suggests that stories and myths form a central element of organizational reality, and therefore organizational research, which ignores stories, is necessarily incomplete. The epistemological position is that by collecting organizational stories, the researcher will gain insights into organizational life which could not be reached by more conventional means. This may involve participant observation, where the researcher can become part of the process of constructing and transmitting stories, or they may be collected through interviews by asking people for the stories that they have heard about particular events (see Example 4.6). In essence, the method relies on literary theory (Hatch, 1996), and hence both the position of the narrator and the role of the analyst are very important.

EXAMPLE 4.6

A narrative-based study

Humphreys and Brown (2008) investigated the way corporate social responsibility (CSR) was introduced as an important function within a financial services company. The authors, consistent with our definition above, saw stories and narratives as central to the way managers and employees make sense of what was going on in the company. But their research design also involved the collection of stories from key actors involved in the establishment of corporate social responsibility within the company. From the analysis of semi-structured interviews with 64 employees, they identified three major themes/narratives associated with CSR – idealism and altruism, economics and expedience, and ignorance and cynicism – which summarized the conflicting perspectives on CSR in that company.

One of the criticizms of narrative methods is that they do not offer much that is distinctive or additional to 'normal' qualitative research. Nevertheless, they do have a number of strengths: they provide a holistic perspective on organizational behaviour; they are particularly useful in developing social histories of identity and development; they are useful in helping to examine relationships between individuals and the wider organization; and they introduce values into the research process. Returning to the map in Figure 4.1, narrative research may be seen as more detached (quadrant B) if the collection of existing stories is emphasized, or more involved (quadrant C) if the researcher plays a role in encouraging people to invent new stories that illustrate their feelings.

Validity of constructionist designs

There is much concern about how to assure and demonstrate the quality of constructionist designs, although authors rarely use the term 'validity'. In a classic paper, Golden-Biddle and Locke (1993) identify three key criteria:

- authenticity
- plausibility
- criticality.

Authenticity involves convincing the reader that the researcher has a deep understanding of what was taking place in the organization; *plausibility* requires the research to link into some ongoing concern/interest among other researchers; and *criticality* encourages readers to question their taken-for-granted assumptions, and thus offer something genuinely novel. More recently Amis and Silk (2008), in discussing 'non-foundationalist' qualitative

research, suggest that good research within the constructionist tradition should be partisan, taking the side of the less powerful members of society and organizations, and supporting a 'moral-sacred' philosophy. Thus quality would be indicated by the presence of the audience in the text, the sharing of emotional experience, stressing political action, taking sides, moving people to reflect and act, and providing collaborative, reciprocal, trusting and friendly relations with those studied.

Another perspective is provided by Silverman (2000), who argues for a more objective stance (and hence a weaker form of social constructionism) because there are few safeguards to prevent researchers from picking evidence out of the mass of data to support their particular prejudices. In order to defend themselves against charges of 'anecdotalism' he suggests several principles, including refutability, constant comparison, comprehensive data treatment and tabulations. *Refutability* involves looking for examples that might disconfirm current beliefs; *constant comparison* follows the principles of grounded theory (see next section) in looking for new cases and settings, which will stretch the current theory; *comprehensive data treatment* involves carrying out an initial analysis of all of the data available before coming up with conclusions; and *tabulations* imply greater rigour in organizing data, and accepting that it can also be useful to add up the occurrence of phenomena sometimes.

Our own view is that the results of constructionist research should be believable, and they should be reached through methods that are transparent. Thus it is very important for the researcher to explain how he or she gained access to the particular organization, what processes led to the selection of informants, how data was created and recorded, what processes were used to summarize or collate it, how the data became transformed into tentative ideas and explanations, and how he or she felt about the research.

CASE METHOD AND GROUNDED THEORY

There are several methods that, despite having a single label, can be used in quite different ways by different proponents. This is particularly true with **case method** and **grounded theory**. Although the dominant texts about case method come from the positivist end, the method can also be designed in ways consistent with relativist and constructionist perspectives. On the other hand, grounded theory was designed as a constructionist alternative to positivist methods, yet some respected versions now contain positivist elements.[2]

Case method

Essentially the case study looks in depth at one, or a small number of, organizations, events or individuals, generally over time. There is a very extensive literature on the design, use and purposes of case studies. In the management field authors tend to coalesce around those who advocate single cases and those who advocate multiple cases. Advocates of single cases generally come from a constructionist epistemology; those who advocate multiple cases usually fit with a more positivist epistemology.

Robert Yin is the best-known exponent of case method in the social sciences (Yin, 2013). His concern is that case studies are vulnerable to a number of criticisms from more *positivist* researchers. In particular, it is suggested that they do not have the rigour of natural scientific designs; they rarely allow generalizations to be made from specific cases to the general population; and they produce huge piles of data, which allow researchers to make any interpretations they want. In response to these criticisms, he suggests that all case studies should have clear

[2]Our focus here is on single methods, which may be interpreted and practised in significantly different ways. This is distinct from 'mixed methods', which involve using combinations of different data collection methods and types for the same study (see Creswell, 2014).

designs produced before any data is collected, and these designs should cover: the main questions or propositions, the **unit of analysis**, links between data and propositions, and procedures for interpretation of data. He is anxious to demonstrate that case studies may contain the same degree of validity as more positivist studies, and therefore his exposition of the method contains both rigour and the application of careful logic about comparisons.

The contrasting position, which is informed by a *constructionist* epistemology, is much less concerned with issues of validity, and more concerned with providing a rich picture of life and behaviour in organizations or groups. Robert Stake (2006) writes about qualitative case studies, and distinguishes between *instrumental* and *expressive* studies. The former involves looking at specific cases in order to develop general principles; the latter involves investigating cases because of their unique features, which may or may not be generalizable to other contexts. An example would be Andrew Pettigrew's research into organization development within the chemical company ICI during the 1970s and early 1980s. In those days ICI was the most powerful manufacturing company in Britain, so there was naturally a lot of interest in understanding how they were managing and developing themselves. In that respect the study was expressive, but there was also an instrumental element since Pettigrew was interested in understanding the phenomenon of organization development, and ICI was regarded as one of its leading proponents. His research involved numerous interviews with key actors in the company over several years, and this provided a longitudinal element to his research, which enabled him better to understand both the contextual and historical settings of the company (Pettigrew, 1985).

From a similar perspective, Nikkolaj Siggelkow (2007) provides a spirited defence of cases arguing that they are particularly valuable for demonstrating the importance of particular research questions, for inspiring new ideas and for illustrating abstract concepts. He also points out that even single cases can provide very convincing tests of theory by quoting the famous 'talking pig' example. Thus we only need to produce a single talking pig to demonstrate the error of the popular idea that pigs are incapable of intelligent speech. The logic being that we only need one example of an anomaly to destroy a dominant theory – as in the case of Einstein's refutation of Newton's theory. And although we are unlikely to identify a 'talking pig' organization, there are many examples where single cases can be uniquely interesting; for example, the company that does significantly better (or worse) than all others in the same industry, or the entrepreneur who builds a fortune from small beginnings.

EXERCISE 4.4

A longitudinal case study

A study conducted by Prieto and Easterby-Smith (2006) explored the links between dynamic capabilities and knowledge management through a case study of the evolution of a single company over several years. Because the researchers were interested in dynamic capabilities – which are by definition about continuous change – it made sense to observe processes over time so they could examine how, for example, the introduction of knowledge-sharing routines led to greater strategic flexibility. Accordingly, the researchers spent time observing management meetings, talking with participants at meetings, and interviewing other managers. They also followed information exchanges with partner organizations by conducting visits to their sites, repeating interviews with key informants, and feeding back emerging insights to senior managers to 'validate' their interpretations and to stimulate further insights.

Questions

1. How would you justify that this research was 'valid'?
2. What possibilities are there for generalizing the findings from this research?
3. Do questions about validity and generalizability make any sense in this instance?

A few points are important about constructionist studies. First, they are based on direct observation and personal contacts, generally through interviews. Second, they take place within single organizations, but then involve sampling from numbers of individuals. Third, the collection of data takes place over a period of time and may include both live observations and retrospective accounts of what has happened. Thus the unit of analysis is either the individual, or specific events such as the exchange of a piece of knowledge, or strategies employed to transfer or retain control of knowledge.

There is also an intermediate position, which has been developed particularly through the work of Kathy Eisenhardt (Eisenhardt, 1989; Eisenhardt and Graebner, 2007). This view draws inspiration from both the positivist and constructionist positions, and has been adopted widely by researchers using case methods, particularly in North America. She is eclectic in her advice about methodology: using designs that are established at the outset, but then being flexible about their adaptation; recommending data collection through using multiple methods; and conducting both within case and across case analysis.

Above all, Eisenhardt is concerned about building theory from case-based research, and this takes the form of developing hypotheses. She recommends that hypotheses can be formed, or shaped, through three main stages. The first stage involves sharpening-up the basic constructs, and this is essentially an iterative process of moving back and forth between the constructs and the data. The second stage involves verifying that emergent relationships between constructs fit with the evidence from each case. In this respect she comments that 'Each case is analogous to an experiment, and multiple cases are analogous to multiple experiments' (Eisenhardt, 1989: 542). The third stage involves comparing the emergent theory/concepts/hypotheses with the existing literature. In particular, she suggests paying attention to literature that is contradicted by the evidence, both because any evidence of having ignored contradictory findings is likely to reduce confidence in the final conclusions, and because the highlighting of contradictory conclusions is likely to make the original contribution from the research most explicit.

EXAMPLE 4.7

Comparative case study design

In a comparative study of investment decisions in Chinese and UK companies (Lu and Heard, 1995), case studies of 16 decisions in 8 companies were compared and contrasted in order to establish the cultural and institutional variations in business decision making between China and the UK. The study involved collecting both qualitative and quantitative data, including extensive site visits to companies in both China and the UK. Each UK company was matched, in terms of size and industry, with the equivalent Chinese company. This allowed for a number of comparisons, between different industries, and between China and the UK, which led to new insights. For example, in the latter case the researchers noticed that the mean time between the inception and implementation of a major investment decision (approximately £100 million) was virtually identical in both China and the UK (approximately 3.4 years). This significantly contradicted existing theory about the speed of decision-making, which suggested that decision making in China was far slower than in the UK. Of course, with the benefit of hindsight it is now possible to see how fast Chinese companies have been developing over the last two decades, but this study was one of the first to demonstrate the speed of economic development in China.

Although the variations in case study design and application are complex and sometimes blend into each other, we summarize in Table 4.2 some of the main distinctions in the use and application of case method at three points along the epistemological continuum.

TABLE 4.2 Key features of case method informed by different epistemologies

	Positivist (Yin)	Positivist and Constructionist (Eisenhardt)	Constructionist (Stake)
Design	Prior	Flexible	Emergent
Sample	Up to 30	4–10	1 or more
Analysis	Cross-case	Both	Within case
Theory	Testing	Generation	Action

Grounded theory

Grounded theory was first formulated by Glaser and Strauss (1967). They saw the key task of the researcher as being to develop theory through 'comparative method', which means looking at the same event or process in different settings or situations. For example, the researcher might be interested in the workings of performance appraisal interviews and would therefore study a number of interviews handled by different managers, in different departments or in different organizations. As a result of the studies it might be noticed that most appraisal interviews either focus on reviewing performance in relation to last year's objectives, or they focus on future goals and how the subordinate may be helped to achieve these. They might then be labelled as 'judgemental' or 'developmental' interviews, and the distinction would represent a *substantive theory* about appraisal interviews.

However, the theorizing could be taken further. For example, it might be observed that neither form of interview has much effect on individual performance, nor on the relationships between the managers and their subordinates. Then we could conclude that both forms of interview are simply organizational rituals, which have the function of demonstrating and reinforcing hierarchical power relations. This would be the beginning of a more generalized *formal theory* about power and organizational rituals. Glaser and Strauss consider both substantive and formal theory to be valuable, and they propose two main criteria for evaluating the quality of a theory. First, it should be sufficiently *analytic* to enable some generalization to take place; at the same time it should be possible for people to relate the theory to their own experiences, thus *sensitizing* their own perceptions.

It is important to note that 'I'm doing grounded theory!' should not be used as a justification for doing some vaguely qualitative research without any clear view of where it is supposed to lead. Grounded theory contains precisely articulated methods and presuppositions. The problem is, as Locke (1997) explains, that methods have evolved and developed since their initial exposition, and at the heart of this was a rather acrimonious debate between Barney Glaser and Anselm Strauss.[3] In essence, Glaser now believes that researchers should start with no pre-suppositions, and should allow ideas to 'emerge' from the data (Glaser, 1978, 1992), whereas Strauss recommends familiarizing oneself with prior research and using structured, and somewhat mechanistic, processes to make sense of the data (Strauss, 1987; Corbin and Strauss, 2015). The implication is that the researcher should be aware that there are different versions of grounded theory, and hence needs to articulate his or her own position when writing up the research. Agreed features are shown in Table 4.3, and differences between Glaser and Strauss are summarized in Table 4.4.

[3]We understand that Glaser and Strauss did meet up and resolve their differences shortly before the untimely death of Anselm Strauss (personal communication).

TABLE 4.3 Agreed features of grounded theory

Grounded Theory	
Must:	fit the substantive area; be understandable and useable by actors; be sufficiently complex to account for variation.
Key analytical operations are:	cycle of theoretical sampling; constant comparisons; evolving theory, *leading to…* theoretical saturation.

TABLE 4.4 Points of disagreement between Glaser and Strauss

	Glaser	Strauss (and Corbin)
Researcher Roles	Maintain distance and independence.	Active interrogation of data.
Theory	Emerges from data itself.	Arises from theorist/data interaction.
Ontology	World is 'out there'.	Reality and experience are constructed.
Pre-understanding	Avoid literature from immediate area.	Flexible approach. Insights from many sources.

The debate is extended further by Kathy Charmaz (2000), who characterizes the methods of both Glaser and Strauss as 'objectivist'. Her complaint is that both authors separate the researcher from the experiences of the subjects of the study. She also feels that the recommendations from Strauss and Corbin (1998) about detailed analysis of transcripts, including line-by-line analysis and 'fracturing of data', reduces the ability to represent the whole experience of individuals involved. In her view, a constructionist should recognize 'that the viewer creates the data and ensuing analysis through interaction with the viewed' (Charmaz, 2000: 523). As such she is located a little further in the constructionist direction than Strauss because she emphasizes the interaction between the researcher and the researched, rather than between the researcher and the data.

In order to make sense of these differences, we need to look both at the ontology and epistemology of the authors. Ontologically, Glaser comes across as a realist, or possibly an internal realist, whereas both Strauss and Charmaz have a more nominalist ontology because they assume that the social world is created through the interaction of actors. Epistemologically, Strauss, who was significantly influenced by Corbin (personal communication), adopts a weak positivist position, which emphasizes systematic and reductionist approaches to the analysis of data. Glaser, on the other hand, promotes a more relaxed epistemology, insisting that the data should be analysed in its entirety, and should not be reduced to discrete elements. In some respects this is similar to the constructionist perspective of Charmaz, though she goes further in emphasizing the primacy of the stories and experiences of her research subjects.

Before completing this section, it is important to note that the methods of grounded theory have been developed mainly within educational and health settings where the

researcher can have relatively easy and flexible access to data and cases. But access is far more difficult within commercial organizations, and researchers are rarely given the freedom to select their samples on theoretical grounds – hence some of the assumptions of grounded theory have to be amended further to deal with this kind of situation (Locke, 2001). Organizational researchers have to accept the interviewees assigned to them by powerful organizational members who act as gatekeepers (see the discussion on strategies for gaining research access in Chapter 5); there are also limits imposed in terms of timing, topics and the use of data. This often requires a number of compromises to be made in terms of research design, as can be seen from the reflections of Suzanne Gagnon in Exercise 4.5 about her study of identity formation among highflying managers in two different international organizations.

Grounded theory, in its different guises, is one of the most popular qualitative methods in business and management research. Suddaby (2006) gives much helpful advice for researchers seeking to use grounded theory methods and publish their work. He lists common misconceptions:

1. Grounded theory is not an excuse to ignore the literature.
2. Grounded theory is not presentation of raw data.
3. Grounded theory is not theory testing, content analysis or word counts.
4. Grounded theory is not simply routine application of formulaic technique to data.
5. Grounded theory is not perfect.
6. Grounded theory is not easy.
7. Grounded theory is not an excuse for the absence of methodology.

EXERCISE 4.5

How grounded is this? A letter from a doctoral student

Hi Mark

I started with a general area for study – the interplay of personal and organizational identities in multi-nationality, multicultural organizations (how important is organizational culture in such settings, and why? What identities do people see themselves as having in these settings, and why?).

Once having been in the sites for some time and gathered some data through interviews, I found that *identity regulation* was a term (perhaps even a central category in Strauss' and Corbin's words) that had explanatory power; I got this term from the literature, having continued to iteratively study the literature and the data, while continuing to gather data.

My 'sample' was more or less set from the beginning (all participants on two management development programmes), so in this sense I did not use theoretical sampling. However, I did add questions and change emphases in the interviews as I proceeded.

Whether I reached theoretical saturation, I am not fully sure. In a sense it was more a question of talking to everyone, and then sampling the data (with some follow-up and changes to subsequent questioning and focus, as above).

I see this as a kind of 'theory elaboration' rather than deduction per se. But there is definitely a deductive side to it. It may also be the case that I come up with my own theory (hope so), especially, perhaps, in comparing results across the two cases.

That's as far as I can go at the moment. What do you think? How grounded is this?

Suzanne

Questions

1. How grounded is this?
2. Should she be sticking more closely to grounded theory principles, and if not, why not?

MIXED METHODS

In recent years there has been growing interest in the use of research methods that draw from both positivist and constructionist epistemologies, and which combine both qualitative and quantitative methods in the same study. This has been stimulated by several influential books (Creswell, 2003; Tashakkori and Teddlie, 2010; Teddlie and Tashakkori, 2009) and by the founding of the specialist *Journal of Mixed Methods Research* in 2007.

There has also been a debate between proponents of mixed methods and those who are sceptical about their value. Those in favour argue that by using a range of different methods within the same study the researcher will increase the validity and generalizability of results and the potential theoretical contribution; the sceptics point to practical limitations such as the competence of researchers in conducting different methods, and to possible contradictions between the paradigms underlying different methods.

There are many variants of mixed methods in social research, but the key idea is that they involve combinations of qualitative and quantitative methods for data collection and analysis. We will therefore start by discussing the choices with regard to data collection, then we consider different strategies for analysis, and finally we consider some of the arguments for and against the use of mixed methods.

Designs for data collection

There are two main considerations in the design of studies that use both qualitative and quantitative methods to conduct research: **sequencing** and **dominance**. Sequencing refers to whether one method goes before the other, and if so which goes first. Dominance is a matter of whether one method uses significantly more time and resource than the other, or whether they are roughly balanced in importance. These considerations are summarized in Table 4.5.

By combining these choices we can identify three distinct designs, which incorporate both quantitative and qualitative methods. We call these master-servant, partnership and compensatory designs.

With *master-servant* designs the key point is that one method serves the needs of the other. There is usually a definite sequence in the use of methods, and naturally one method dominates the other. The most common format is the qualitative pilot study based on interviews or direct observation, which is used to develop, and maybe test, the items for the main study, which involves a questionnaire survey. Here the questionnaire survey

TABLE 4.5 Choices in designing mixed methods research

Design Features	Alternatives to Consider
Sequencing of methods	Qualitative first, or quantitative first, or both at the same time
Dominance of methods	Predominantly qualitative, or quantitative, or balanced

is dominant, and the pilot study serves no function in the final result of the work, other than helping the researchers to design a questionnaire that is likely to yield accurate and reliable data.

There are also contemporary examples of the reverse process, where a survey is used to identify a small number of 'interesting' cases for in-depth investigation, and then the survey results are largely ignored in the final results. For example, Macpherson et al. (2010) conducted a survey based on single interviews with 92 entrepreneurs. From this sample they identified three critical cases where they conducted repeated interviews over a year in order to establish how various artefacts (such as knowledge management software, benchmarking and problem-solving forums) contributed to the learning processes of the SME. In this case it is the qualitative study that dominates the published paper, with the survey merely in the background. Similarly, Detert and Edmondson (2011) investigated why employees usually fail to speak their minds to those in authority through a four-stage design. In the first stage they conducted interviews in one company to identify the implicit theories from employees about why it was unwise to speak up; the second stage used an open survey questionnaire circulated to 185 managers in different organizations to test the generalizability of these theories. Stages three and four then developed and validated quantitative survey instruments, which could be used to examine the phenomenon on a much larger scale.

Partnership designs typically involve combining more than one method, such as a questionnaire survey and interviews, where both assume similar importance in the study. For example, entrepreneurial behaviour can be investigated by interviewing a small sample of entrepreneurs about their origins, motives, strategies, successes and failures, supplemented by a questionnaire containing similar questions sent out to a larger sample. When combined, the interview data will contain greater detail, clarifications and added explanations; the questionnaire data will contain shorter answers, possibly more focused, but will be able to cover responses from a wider range of entrepreneurs who could be divided into subgroups to explore possible differences according to family history, levels of funding, types of technology and so on.

Compensatory designs combine qualitative and quantitative studies where each is used to make up for the weaknesses of the other. Typically qualitative studies are seen as weak on generalization, and quantitative studies are weak at explaining why the observed results have been obtained. Thus there is a growing trend in leading US publications such as the *Academy of Management Journal* for quantitative studies that establish statistical relationships between variables to be supplemented by quotations from substantial numbers of interviews focusing on the mechanisms and processes, which may provide explanations of the observed results.

Analysis

Another form of mixed methodology can be introduced at the analysis stage. Although qualitative and quantitative data are normally analysed within their respective traditions, there is also the possibility of cross-over designs. Thus quantitative data can be analysed in qualitative ways and qualitative data can be analysed in quantitative ways. The most common form of the latter is when frequency counts are made of the use of particular words, phrases or themes from a sample of interview transcripts. The study by O'Reilly and Reed (2010) (see Example 4.4) provides an example of qualitative archival data in the form of government policy documents being analysed quantitatively for the occurrence of particular words and expressions. Slightly less common is when quantitative data is analysed by techniques such as factor analysis, and **principal components analysis,** which look for patterns that are largely hidden. Techniques such as the **repertory grid** technique (see Chapter 7) involve starting with qualitative data, which then becomes quantified and analysed statistically, and the result is then interpreted qualitatively.

Arguments for and against mixed methods

As we have outlined above, there are many reasons why mixed methods are regarded as a good thing: they have the potential to throw new perspectives on research questions, to increase the credibility of results, to demonstrate generalizability, and to provide deeper insights that explain why things take place. But there are also plenty of reasons for being cautious about their wholesale adoption. We summarize some of these pros and cons in Table 4.6 based on the arguments of Jick (1979), Tashakkori and Teddlie (2010) and Bryman and Bell (2007).

There is, however, a more fundamental critique of the use of mixed methods, which hangs on the notion of paradigm incommensurability (Burrell and Morgan, 1979; Morgan and Smircich, 1980). The argument is that it is unwise to combine different paradigms within the same study because the different underlying assumptions mean that it will not be possible to join the two parts of the study together. At the extreme this can produce a **semi-detached design**, because like two semi-detached houses, they are physically linked together, yet there is no adjoining doorway between the two parts of the house.

The weakness of the incommensurability argument is that it assumes that paradigms are always distinct and that there can be no overlaps. Recent thinking about paradigms suggests that boundaries are more fluid than originally portrayed (Cunliffe, 2011), and hence it may be acceptable to combine paradigms up to a point. In our view the limits can be defined by the continuum presented in Figure 4.1, where it is possible to combine adjacent ontologies and epistemologies within a mixed methods study, but increasingly problematic when combining more distant positions. For example, a positivist study might demonstrate that 80 per cent of corporate performance could be predicted by three variables: size, market share and growth rate. But when combined with an ethnographic study exploring the micro-politics of constructions of corporate performance, this would not contribute in any way to identifying the remaining 20 per cent in the predictive formula. It would be more likely to undermine the credibility of the main study by arguing that the concept of 'performance' is a sham.

The use of mixed methods can often lead to contradictory results. If the ontologies are very different there will be no way of resolving the confusion. However, if they are close enough then resolution may be possible, as illustrated in Example 4.8.

TABLE 4.6 Pros and cons of mixed methods

Arguments for Mixed Methods	Arguments against Mixed Methods
● They increase confidence and credibility of results	● Replication is difficult
● They increase validity	● The research design must be relevant to the research question
● They stimulate creative and inventive methods	● They provide no help if you are asking the wrong questions
● They can uncover deviant dimensions	● They take up more resources than single method studies
● They can help synthesis and integration of theories	● Their use requires a competent overall design
● They may serve as a critical test of competing theories	● The researcher needs to be skilled in the use of both methods
● They can combine confirmatory and exploratory research at the same time	● It is not helpful if one method simply provides window dressing for the other (an extreme version of the handmaid design discussed above)
● They present greater diversity of views	
● They provide better (stronger) inferences	

EXAMPLE 4.8

Problems with mixed methods

Morgan Tanton and Mark Easterby-Smith carried out a comparative evaluation study of two executive management programmes (Courses A and B), held in two different business schools (respectively, Institutions A and B). Observations during the course, and qualitative data obtained from follow-up interviews, showed quite clearly that Course A was superior to Course B, but the quantitative data in the form of student ratings about the two courses showed clearly that Course B was preferred to Course A. Was this discrepancy caused by the methods used, or could it highlight some unusual features of the two courses being examined?

To resolve this dilemma we showed the survey results to participants and asked for their explanations. First, participants commented that they were cautious when filling in multiple choice rating forms, because they could never be sure what the data would be used for; therefore, they usually avoided unduly negative responses. Second, the course designs and institutional settings affected the criteria that participants used for evaluating the two courses. In Institution A the emphasis was on the longer-term application of what had been learnt; in Institution B the emphasis was on the immediate quality of sessions conducted within the classroom. Thus it was not surprising that the rating forms which were completed at the end of the course showed one pattern, whereas follow-up interviews conducted some months later showed another pattern. In this case it was possible to combine the two sets of data because the survey and interviews were respectively backed by internal realist and relativist perspectives, and both parts shared a common research question.

Finally, we can note that much of the interest in mixed methods comes from those on the positivist side of the spectrum, who hold at least an internal realist view of the world, on the grounds that added data and more perspectives will enable them to get closer to the intangible objects of their enquiries and cynics might say that positivists need to incorporate more constructionist methods to make up for the shallowness of their traditional methods!

Summary

We accept that in some circumstances mixed methods can be advantageous, but think it is important that a clear rationale for their use is worked out in advance, and care needs to be taken to ensure that the methods are reasonably compatible. There is always a danger in using mixed methods just because they might add to the overall credibility of the research, because the ad hoc combination of different kinds of study means that neither may be done properly. As Bryman and Bell (2003: 493) comment: 'multi-strategy research should not be considered as an approach that is universally applicable or as a panacea'.

COMMON DESIGN DILEMMAS

In this section we identify five areas that require decisions when formulating research designs, irrespective of the ontology or epistemology that informs the study. These are:

- identifying the unit of analysis
- universal theory or local knowledge
- theory or data first

- cross-sectional or longitudinal
- verification or falsification.

Identifying the unit of analysis

The unit of analysis is the entity that forms the basis of any sample. Thus, samples may be formed from one or more of the following: countries, cultures, races, industrial sectors, organizations, departments, families, groups, individuals, incidents, stories, accidents, innovations and so on. In positivist forms of research, including multiple case studies informed by an internal realist perspective, it is important to be clear about the unit of analysis in advance, because this is the basis for collating data that will subsequently be analysed. It is not essential in constructionist forms of research, but with highly unstructured data it can help to provide an initial guidance for analysis. In the above example from our research, which compared decision-making between China and the UK (Example 4.7), the unit of analysis was the company, but there was a subsidiary unit of analysis (what is sometimes referred to as an **embedded case**), which was the investment decision. Hence it is possible to have more than one unit of analysis provided the theoretical aims of the research justify this, but it is not advisable to have too many.

Universal theory or local knowledge?

One of the key principles of scientific methods and positivist knowledge is that theories and observations made in one context should be applicable to other defined contexts. As we have discussed above, being able to provide assurances of **generalizability**, or external validity, are critical features both of experimental designs and the statistical procedures that are employed to interpret realist research data. In these cases, as with the guidance of Kathy Eisenhardt on case method, the objective is to produce **universal theories**.

On the other hand, a number of scholars argue that **local knowledge** is more significant. For example, according to post-colonial theory many theories of race, economic development and culture are constructs of scholars in Western countries, which typically cast non-Western culture and institutions as being somehow inferior to their own (Said, 1978). Similarly, from feminist theory there is a strong view that many of the dominant theories of social behaviour are blind to the effects of gender and patriarchy (Ahmed, 1998). In both cases the argument is that any generalized statement about the social world is likely to contain within it assumptions that mask relations of power between those who formulate theories and those to whom they are applied. Moreover, there is a strong view that significant social theory should be understood in relation to the context whence it is derived.

Local knowledge is also important for management and organizational research. First, it is suggested that the practical knowledge used by managers is essentially contextually bound, and is learnt through engaging in practice (Cook and Brown, 1999; Rouleau, 2005). If this is the case then it follows that for research to have theoretical value it should focus on these local practices – which may well be unique to that situation. Second, some people argue that managerial behaviour is culturally relative, including both national and organizational cultures (Boyacigiller and Adler, 1991). Hence researchers should formulate their ideas separately within each cultural context, and should not try to generalize across cultures.

For example, it has been accepted for some time that models derived from Western management research are unlikely to be relevant in Asian contexts, as Nor (2000) found in his study of Malaysia, because of Malaysia's unique cultural, political and institutional circumstances. Over the last decade there has been much interest in the development of entrepreneurial capabilities in Asian countries such as China, Vietnam, Malaysia and India. It is increasingly accepted that the cultural and institutional differences between

these countries are such that local theories to explain entrepreneurial behaviour are necessary in each country (Taylor, 1999; Hobday and Rush, 2007).

Theory or data first?

The third choice is about which should come first: the theory or the data? Again this represents the split between the positivist and constructionist paradigms in relation to how the researcher should go about his or her work. The Straussian view of grounded theory assumes that pre-conceptions are inevitable. After all, it is common sense to assume that someone will not be interested in a research topic or setting without knowing something in advance about it. Hence he argues that the researcher should make him or herself aware of previous work conducted in the general field of research before starting to generate his or her own theory.

Recent developments in organizational research have led to a wide range of designs, some of which extend the range of fieldwork methods, and others that provide intermediate positions between the two extremes. In a recent research project looking at absorptive capacity within European companies, Easterby-Smith et al. (2008) became increasingly aware that the relationship between theory and data needs to be an interactive process. When researchers observe something that seems surprising or novel in a company, it is important to go back to the literature in order to see whether anybody else has remarked on it. This ongoing dialogue between existing knowledge and what the data have to say that is novel is one reason why initial literature reviews are almost always modified before the final write-up of a research project. Similarly, when a new paper gets published it may have a direct impact on the ongoing collection and interpretation of data.

Cross-sectional or longitudinal?

Cross-sectional designs, particularly those that include questionnaires and survey techniques, generally belong to positivist traditions. As we have noted earlier, they have undoubted strengths in their ability economically to describe features of large numbers of people or organizations. But a major limitation is that they find it hard to describe processes over time and to explain *why* the observed patterns are there. Thus, although Lyles and Salk (1996) were confident that balanced equity stakes led to the highest chance of knowledge transfer, their study itself could not explain what mechanisms or processes led to knowledge being transferred.

In order to understand processes of change over time it is necessary to adopt longitudinal designs. From the positivist side these include quasi-experimental methods and diary methods because repeated measurements are taken over time, but it is more often associated with constructionist research, where repeated visits are made to the same individual or companies over months or years, or when the researcher conducts an ethnographic study working continuously in the same location.

Verification or falsification

This final decision is slightly different from the four preceding ones since it is not linked to resolving the broader debate between positivist and constructionist views. However, it is very important both for researchers and for managers, as we will explain below. The distinction between **verification** and **falsification** was made by Karl Popper (1959) as a way of dealing with what has become known as Hume's 'problem of induction'. This is the philosophical problem that, however much data one obtains in support

of a scientific theory, it is not possible to reach a conclusive proof of the truth of that law. Popper's way out of this problem is to suggest that instead of looking for confirmatory evidence one should always look for evidence that will *disconfirm* one's hypothesis or existing view (as in the 'talking pig' example on page 90). This means that theories should be formulated in a way that will make them most easily exposed to possible refutation. The advantage then is that one only needs one instance of refutation to falsify a theory, whereas irrespective of the number of confirmations of the theory, it will never be conclusively proven.

The example often given to illustrate this approach takes as a start the assertion that 'all swans are white'. If one takes the verification route, the (non-Australian) researcher would start travelling around the country accumulating sightings of swans, and provided that he or she did not go near a zoo, a very high number of white sightings would eventually be obtained, and presumably no black sightings. This gives a lot of confidence to the assertion that all swans are white, but still does not conclusively prove the statement. If, on the other hand, one takes a falsification view, one would start to search for swans that are *not* white, deliberately looking for contexts and locations where one might encounter non-white swans. Thus, our intrepid researcher might head straight for a zoo, or perhaps book a flight to Western Australia where most swans happen to be black. On making this discovery, the initial hypothesis would be falsified, and it might then have to be modified to include the idea that 'all swans have either white or black feathers'. This statement has still what Popper calls high 'informative' content because it is expressed in a way that can easily be disproved, whereas a statement like 'all swans are large birds' would not be sufficiently precise to allow easy refutation.

Much of the debate about verification and falsification fits within the positivist view because ideas of 'truth' and 'proof' are associated mainly with that paradigm. But there are also important lessons that the constructionist might take from this discussion. For example, Alvesson and Deetz (2000) advise 'critical sensitivity', and Reason (1988) advocates 'critical subjectivity', which involves recognizing one's own views and experiences, but not allowing oneself to be overwhelmed and swept along by them. If the idea of falsification is to be applied more fully to constructionist research, then one should look for evidence that might confirm or contradict what one currently believes to be true. Indeed, a falsification strategy is an important element of abductive research (Dubois and Gadde, 2002) which emphasizes ways in which data can generate new theory rather than either confirm or disconfirm existing theory.

This advice not only applies to researchers but also to managers who are concerned to investigate and understand what is taking place within their own organizations. Most managers are strongly tempted to look for evidence that supports the currently held views of the world. This is not surprising if they are responsible for formulating strategies and policies within a context that is very uncertain, and hence they will be looking for evidence that demonstrates that their strategies were correct. The logical position that follows from the above argument is that, even if *disconfirmatory* evidence is unpopular, it is certainly both more efficient and more informative than confirmatory evidence. Moreover, if managers adopt the falsification strategy and fail to come up with evidence that disconfirms their current views, then they will be able to have far more confidence in their present positions.

CONTRIBUTING TO THEORY

Good research designs need to have some link to theory. In the case of student projects and dissertations it is generally necessary to *use* theory, whereas for doctoral theses and papers in academic journals it is necessary to demonstrate a *contribution* to theory. This is not as daunting as it might seem, and in this section we elaborate on the types and purposes of theory, and explain how they can be incorporated into research designs.

The term 'theory' often has negative connotations. Someone might report back on a lecture saying, 'It was all a lot of theory!', meaning that it was either difficult to understand or just plain boring. Or someone might react to a new idea saying, 'Well that's all right in *theory*, but …', meaning that although the idea sounds plausible, it would not work in practice. So, in this case theory is seen as the opposite of practice. On the other hand there is the well-known saying, 'There is nothing so practical as a good theory' (Lewin, 1948). In order to unscramble this confusion we offer distinctions between everyday and academic theory, the latter subdividing further into middle-range and grand theories.

Everyday theory refers to the ideas and assumptions we carry round in our heads in order to make sense of everyday observations. For example, if you observe an old man walking down the street arm in arm with a young woman, you might conclude that they were grandfather and granddaughter. In order to reach this conclusion you might hold two assumptions about family relations: that grandparents often live close to their family members, and that grandparents often have very close relations with their grandchildren. If the man is leaning slightly on the woman, then it would strengthen the grandfather–daughter hypothesis; but if the man's walk was very unsteady this might suggest a new theory, that they are patient and nurse. On the other hand, if the man is well dressed and the woman is conspicuously glamorous, an alternative hypothesis might suggest itself: that the man is a wealthy philanderer and the woman is a mistress or 'trophy' wife.

Although everyday theories enable people to make sense out of specific events or situations, **academic theories** tend to look for higher levels of generalization. Following the above example for just a moment, in order to explain what was going on, a sociologist might draw on theories about the power of male patriarchy, palliative care for the elderly, or the evolution of the institution of marriage. The distinction between **middle-range theories** and **grand theories** is a matter of scale and formality. An example of the former would be the key idea of absorptive capacity: that the ability of an organization to absorb new external knowledge depends on whether it already possesses related knowledge (Cohen and Levinthal, 1990). It is middle-range because it is a generalizable proposition that can potentially be tested empirically.

On the other hand, grand theories tend to be more abstract and contain whole edifices of assumptions that are often not testable. The theory of psychoanalysis is one example because it provides a self-contained set of ideas to explain human behaviour. Similarly, personal construct theory (PCT) contains a set of propositions starting with the fundamental postulate that 'A person's processes are psychologically channelled by the way they anticipate events', which is linked to a series of corollaries about human sense-making and communications (Kelly, 1955; Bannister and Fransella, 1971). In the management field, elements of PCT have been used to make sense of group decision making and strategy formulation (see Chapter 7). A number of the integrated philosophies summarized at the end of Chapter 3, such as critical theory or structuration theory, are grand theories in the way we have described them here.

Where researchers are seeking to build theory, this is normally at the level of middle-range theory, and is an incremental process. Thus recent work by Todorova and Durisin (2007) has argued that Cohen and Levinthal's (1990) model of absorptive capacity is too rational and unduly focused around R&D, and consequently more attention needs to be paid to political and systemic processes. This leads to a question about how we can evaluate the quality of theories, or theoretical contributions, and how can we distinguish a good contribution from one that is less good? The answer is that some criteria are fairly obvious: good theories need to be simple, have good explanatory power, and be relevant to issues that need explaining. But beyond this the evaluation of contribution is largely a matter of judgement among people who already know the field quite well, which is why peer review is normally used to evaluate the theoretical contributions of research proposals and academic papers. We will be returning to these issues at various points later in the book, especially in Chapter 12.

CONTRASTING VIEWS ON VALIDITY AND RELIABILITY

There is an underlying anxiety among researchers of all persuasions that their work will not stand up to outside scrutiny. This is very understandable since research papers and theses are most likely to be attacked on methodological grounds, and one of the key justifications for doing 'research' is that it yields results that are more accurate and believable than common everyday observations.

The technical language for examining this problem includes terms such as 'validity', 'reliability' and 'generalizability'. But as we have indicated above, these mean different things within different research traditions. In Table 4.7 we therefore summarize how these terms are discussed from the philosophical viewpoints of positivism, relativism and constructionism.

The implication of Table 4.7 is fairly obvious: that depending upon where people stand on the epistemological continuum, they are likely to use different criteria for judging the quality of research. This will affect how they design and conduct their own studies and how they assess the quality of others' work, particularly when they are acting as examiners, reviewers or just colleagues.

TABLE 4.7 Four perspectives on validity, reliability and generalizability

Viewpoint	Strong Positivist	Positivist	Constructionist	Strong Constructionist
Validity	Has the design excluded all rival hypotheses?	Does the design make it possible to eliminate plausible alternative explanations?	Have a sufficient number of perspectives been included?	Does the study clearly gain access to the experiences of those in the research setting?
Reliability	Do the measures correspond closely to reality?	Do the measures used provide a good approximation to the underlying concepts of interest?	Will similar observations be reached by other observers?	Is there transparency about data collection and interpretation?
Generalizability	Does the study confirm or contradict existing findings in the same field?	Are the patterns observed in the sample data consistent with findings from other studies?	Is the sample sufficiently diverse to allow inferences to other contexts?	Do the concepts and constructs derived from this study have any relevance to other settings?

RESEARCH DESIGN TEMPLATE

We have argued throughout this chapter that research designs should take account of epistemology, and hence formal research designs need to focus on different issues. In Table 4.8 we list some of the main headings that need to be covered within each epistemology. The key point about this table is that a research proposal will need to consider different issues and to use different language according to where the researchers, and more importantly any external assessors, stand. The way to use this template is to decide which epistemology is most appropriate to your research study and then follow the questions down the relevant column.

TABLE 4.8 Research design template

Epistemology	Strong Positivist	Positivist	Constructionist	Strong Constructionist
Background	What is the theoretical problem and what studies have been conducted to date?	What is the theoretical problem and what studies have been conducted to date?	What are the ongoing discussions among researchers and practitioners?	What are the ongoing discussions among researchers and practitioners?
Rationale	What is the main gap in existing knowledge?	What are the main variables, and how are they related to one another?	What perspectives have been covered and what are missing?	What are the limitations in the discussions so far?
Research Aims	Specify testable hypotheses.	List main propositions or questions.	Identify the focal issue or question.	Explain how the research will add to the existing discussion.
Data	Define variables and determine measures.	Define dependent and independent variables and determine measures.	Explain and justify a range of data collection methods.	Identify main sources of data. How will interviews be recorded/ transcribed, etc.?
Sampling (see Chapter 8)	Explain how group selection and comparison will eliminate alternative explanations.	Justify sample size and explain how it reflects the wider population.	How will the sample enable different perspectives to be included?	Explain sampling strategy. Will it be opportunistic, emergent, comparative, etc.?
Access (see Chapter 4)	How are experimental subjects to be recruited?	How can responses to questionnaires etc. be assured?	What is the strategy for gaining access to individuals, organizations?	How will insights from co-researchers be combined?
Ethics (see Chapter 4)	Is participation voluntary?	Could results be used to harm any participants?	Will the interests of individuals and organizations be protected?	How 'open' is the research? Will there be any deception?
Unit of Analysis	Differentiate between control, experimental groups, etc.	Specify whether individuals, groups, events or organizations.	How will units/ cases be compared with each other?	What are the entities that are to be compared with each other?
Analysis (see Chapters 7, 9 and 10)	Statistical procedures for examining differences between groups.	Statistical procedures for examining relationships between variables.	Arrangements for coding, interpreting and making sense of data.	How will co-researchers be involved in sense-making?

Epistemology	Strong Positivist	Positivist	Constructionist	Strong Constructionist
Process	Explain stages in the research process.	Explain stages in the research process.	Explain what can be pre-planned and what can be open-ended.	Provide realistic timing including adequate provision for contingencies.
Practicalities (see Chapters 6 and 8)	How will groups be recruited? Where will experiments take place?	Who will gather data? How will it be recorded/ stored? Who will analyse it?	How will researchers share observations? Who will do transcriptions, etc.?	How will co-researchers be engaged?
Theory	How will hypotheses be tested?	In what ways will the results add to existing theories?	Will the research build on existing theory or develop new concepts?	Will the research build on existing theory or develop new concepts?
Outputs (see Chapter 11)	Where will the research results be published?	What is the dissemination strategy?	What is the dissemination strategy?	How will insights be shared with colleagues and collaborators?

CONCLUSION

In this chapter we have discussed some of the key philosophical debates underlying research methods in the social sciences, and we have looked at the implications these have for the design of management research. Some key points are:

- There is a clear dichotomy between the positivist and social constructionist worldviews, but the practice of research involves a lot of compromises.
- Each position has its own language and criteria for evaluating research designs.
- There is considerable diversity of methods and designs, especially within the constructionist research tradition.
- Differences in opinion about research methods are often underpinned by ontological differences.

The worldview held by an individual researcher or institute is an important factor, which affects the choice of research methods. But there are other factors, too. Senior academics can exert pressure on junior colleagues and students to adopt methods that they favour. Governments, companies and funding organizations can exert pressure on institutions to ensure that the aims and forms of research meet with their interests. The politics of research are complex, and researchers neglect them at their peril. That is why we have chosen to devote the next chapter to a discussion of these issues.

EXERCISE 4.6

Discussion questions (for small groups in class)

Classify the following according to whether you consider them to be ontologies, epistemologies, methodologies or methods: grounded theory; unobtrusive measures; narrative; case method; ethnography; critical realism; participant observation; experimental design; falsification; theoretical saturation. If it is a weak association put * into the corresponding box, ** for a moderate association, and *** for a strong association. Explain your reasoning. (Note: many of them could be more than one thing.)

	Ontology	Epistemology	Methodology	Method
Grounded theory				
Unobtrusive measures				
Narrative				
Case method				
Ethnography				
Critical realism				
Participant observation				
Experimental design				
Falsification				
Theoretical saturation				

FURTHER READING

Charmaz, K. (2014) *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*, 2nd edn. London: Sage.
A good textbook that introduces how to design and conduct a grounded theory study. Charmaz is a prominent advocate for constructionist approaches to grounded theory, distancing herself from the more positivist leanings of the founders of grounded theory.

Creswell, J.W. (2003) *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*, 2nd edn. Thousand Oaks, CA: Sage.
A great beginner's overview of the three different paradigms.

Locke, K. (2001) *Grounded Theory in Management Research*. London: Sage.
This is an excellent overview of the origins of grounded theory including the differences of opinion between Glaser and Strauss, the key methods and approaches as currently practised, and the specific adaptations that may be required when conducting organizational or management research.

Miller, D.C. and Salkind, N.J. (2002) *Handbook of Research Design & Social Measurement*, 6th edn. Thousand Oaks, CA: Sage.
A useful, albeit slightly dated, handbook for further reference with excellent chapters on how to develop a research design, formulate a research problem and compose a research proposal. It also includes a section on applied and evaluation research.

Shadish, W.R., Cook, T.D. and Campbell, D.T. (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin.
An updated version of the classic book on experimental forms of social research.

Teddlie, C. and Tashakkori, A. (2009) *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. London: Sage.
Chapter 7 of this authoritative text on mixed methods research considers research designs for studies which combine quantitative and qualitative methods.

Want to know more about this chapter? Visit the SAGE edge™ website at **https://edge.sagepub.com/easterbysmith** to access practice questions, selected journal articles, useful web links and much more.