9 Experiments

In this chapter you will learn about:

- what is meant by an experimental research strategy;
- issues to address in planning and performing experimental research;
- how experiment-based research might be applied to the Internet;
- how experiments have been used in IS and computing;
- the advantages and disadvantages of experiment-based research;
- analysing and evaluating experiment-based research.

Defining Experiments

In everyday language, we often speak of doing an experiment when we mean we will try something out and find out what happens. For example, I wondered if taking a different route to work would reduce my journey time. Yesterday I carried out an 'experiment' to find out – I took the different route and timed my journey. It took me 10 minutes less than my normal average journey, so I *might* conclude that my experiment showed the new route was quicker. However, my experiment could be criticized. How can I be certain that there were no other factors affecting yesterday's journey that might have caused my shorter journey time? Did I set off for work at exactly the same time as normal, or was my departure later than usual, so I missed the rush hour traffic and had a quicker journey? Was it the first day of the school holidays, so there was less traffic on the roads? Had an accident occurred on a road feeding into my route, so fewer cars were able to join my route? Had there been an item on the radio warning of possible disruptions on my route and advising drivers to take a different road? Academic researchers would therefore criticize my experiment because I had not paid

attention to the many variables that could affect my journey, and so I am not justified in saying that my new route *caused* my shorter journey time.

In academic research, an experiment is a strategy that investigates cause and effect relationships, seeking to prove or disprove a causal link between a factor and an observed outcome. It is often associated with research in the physical sciences (for example, physics, chemistry and metallurgy) and is at the heart of the scientific method and positivism (see Chapter 19). Researchers start by developing a theory about their topic of interest, which leads to a statement based on the theory that can be tested empirically via an experiment. This statement is of the form 'Factor A causes B', and is known as a *hypothesis*. For example:

- Hypothesis: the Watson–Klein–Bloggs algorithm enables quicker data processing than our current algorithm.
- Hypothesis: smoking cigarettes causes lung cancer.
- Hypothesis: students pay more attention in lectures if you don't give them handouts.

An experiment is then designed to prove or disprove the hypothesis. All factors that might affect the results are excluded from the study, other than the one factor that is thought to cause a particular outcome. The experiment is run and careful measurements or other observations are made of the outcomes. If the researchers are confident that no other factor could have caused the observed results, their hypothesis that factor A causes outcome B is proven. However, even the most carefully designed laboratory-based experiment might have been contaminated by some unrecognized other factor. Good researchers would not therefore draw firm conclusions from experiments until they have been repeated many times by both themselves and other researchers. (This need for *repeatability* of experiments and outcomes is discussed further in Chapter 19.)

A research strategy based on experiments is typically characterized by:

- Observation and measurement. The researchers make precise and detailed observation of outcomes and changes that occur when a particular factor is introduced or removed.
- A process of (1) observation or measurement of a factor; (2) manipulation of circumstances; and (3) re-observation or re-measurement of the factor to identify any changes.
- Proving or disproving a relationship between two or more factors.
- Identification of causal factors. The researchers do not just identify that two factors appear to be linked, that is, they occur at the same time or always happen in sequence. The researchers aim to discover which factor is the cause (called the 'independent variable') and which the effect (called the 'dependent variable(s)').
- Explanation and prediction. The researchers are able to *explain* the causal link between two factors by means of the theory from which their hypothesis was

derived or, in some cases, by means of a new theory that they propose. They are able to *predict* future events if their experiment proves that a factor will always cause a particular outcome.

- Repetition. Experiments are typically repeated many times, and under varying conditions, to be certain that the observed/measured outcomes are not caused by some other factor, for example, faulty equipment.

> **Warning: Terms and definitions**
>
> Note that a few writers use the term 'experiment' to mean *any* piece of field research including surveys and case studies, and so on. They use the term 'formal experiment' to mean the kind of experimental research described in this chapter.
>
> In this book an experiment is defined as a particular kind of research strategy that aims to isolate cause and effect by manipulation of what is thought to be the causal, or independent, variable and measurement of its effect on the dependent variable(s).

Planning and Conducting Experiments

When designing an experiment you have to think about the hypothesis to be tested, the variables to be controlled and measured, and internal and external validity. There are also different kinds of experiments: true experiments, quasi experiments and uncontrolled trials. All of these are explained in this section.

Hypotheses

An experiment is based on a hypothesis to be tested. A hypothesis is a statement that has not yet been tested empirically (that is, by gathering field data) but for which it is *possible* to devise empirical tests that will provide clear evidence to support it or reject it. It is written as a kind of prediction, for example:

- When factor A occurs B will happen.
- An increase in D causes a decrease in C.
- Water will freeze when the temperature drops to 32 degrees Fahrenheit (0 degrees centigrade).

A hypothesis must be testable and it must always be possible to disprove, or falsify, the hypothesis. For example, the statement 'All squares have four sides of equal length and contain four right angles' is not falsifiable, (because something only gets called a square if it meets that definition) and so is not a hypothesis.

Independent and dependent variables

It is important to distinguish between *dependent* and *independent* variables. The independent variable affects one or more dependent variables: its size, number, length or whatever exists independently and is not affected by the other variable. A dependent variable changes as a result of changes to the independent variable (see the box below). An experiment will be based on manipulation of the independent variable to observe the changes in the dependent variable(s). You might introduce an independent variable into a situation, or remove one, or change the size of one. For example, the introduction of a new teaching and learning approach might lead to improved examination results. This could be investigated via an experiment, where the independent variable is the new approach and the dependent variable is the exam results.

Sometimes an experiment is designed to manipulate more than one independent variable. However, the more independent variables that are measured, the more complex the statistical analysis required. Unless you are already highly skilled in statistics, I recommend you stick to one independent variable, two at most.

Variables in experiments

Independent variable	Dependent variable
Cause	*Effect*

Controls

The aim of a research strategy based on experiments is to show that one factor *only* causes an observed change. The researcher therefore tries to control all the variables, either all at once or in a sequence of experiments so that in the end just one factor remains as the only viable cause of the observed change. For example, in the experiment mentioned above, to examine the link between a new teaching approach and examination results, it would be necessary for the examination and the marker to be exactly the same as before the introduction of the new approach, otherwise it is possible that an easier examination or a more generous marker caused the improved results. Ways of controlling variables that might affect your outcomes include:

- **Eliminate the factor from the experiment**: for example, if students are to be used as subjects in an experiment about how people can be helped to learn programming skills, it might be necessary to exclude from the study any students who have previous computing experience, to ensure that all the participants have comparable previous experience.
- **Hold the factor constant**: this is used if it is not possible to eliminate a factor from a situation. For example, if the age of the people participating in an experiment might have a possible effect on the results, choose participants who are all the same age.
- **Use random selection of subjects**: if sufficient research subjects are chosen on a random basis, any factors associated with individuals that might interfere with the results (age, weight, gender, beliefs, and so on) should cancel each other out across the whole group.
- **Use control groups**: two groups are set up where the members are equally balanced: numbers, gender, age, previous relevant experience, and so on. For one group, the control group, no manipulation of the independent variable occurs. For the other, the independent variable is manipulated. Outcomes are measured for each group. Providing the groups have been chosen properly and the experiment designed and conducted properly, any variation in the outcomes between the two groups must be attributable to the variable that was manipulated.
- **Make the researchers and subjects blind**: don't panic – only metaphorically speaking. This approach tries to ensure people don't influence the results because of their expectations. When testing a new drug, medical researchers make sure that neither the researchers nor the patient subjects know who is receiving the drug under trial and who is receiving a placebo, a sugar-pill with no effect. In IS and computing, a computer program rather than a researcher could allocate test materials to groups, and researchers could measure outcomes and statistically analyse them without knowing which group was the control group and which received the manipulation.

Observation and measurement

The experiment strategy involves observations, making measurements of the dependent variables and observing change. Typical things which are observed and measured include:

- project data, for example, number of person-hours, cost, time to completion;
- self-report responses (for example, the subjects of the experiment complete a questionnaire about their feelings or how they rate a concept);
- behavioural counts – the number of times a certain kind of behaviour occurs (for example, number of times someone asks for help or accesses an online help system);
- number of bugs in a piece of code;
- time to process a file of data.

There should be a 'before' and 'after' measurement, (called *pre-test* and *post-test*) otherwise no change, which might be attributable to the manipulation of the independent variable, can be observed. For example, if a new computerized system for customer orders is introduced, it is not sufficient to measure customers' satisfaction with the system by means of a post-implementation questionnaire and then say that their satisfaction was *caused* by the new system. Customer satisfaction questionnaires would need to be completed before and after the introduction of the new system, to see if any change (hopefully improvement) is detected. (Of course, customer satisfaction could really be dependent on the quality of goods, speed of delivery, match to catalogue description, and so on.)

Normally quantitative data is used, because we need to measure change and use statistical analysis. Sometimes, we have to find a variable to measure that only approximates to what we really want to measure. For example, most IS and computing researchers do not have access or the expertise to use medical equipment that can measure stress levels in humans (for example, heart rate, blood pressure, skin conductance). Instead, we might have to ask people to fill out a questionnaire where they rate their stress levels on a scale of 1–10.

In IS and computing research, many outcome measures are related to each other. For example, productivity rates, code defects and delivery times. An experiment could be designed, for instance, to show that a new development method decreases the time taken to deliver the finished product. However, as well as the delivery time, it would be important to measure other factors, such as code defect rates or analyst stress, to see if they are adversely affected by adoption of the new method.

Results are usually analysed using statistics. It is important that you decide *before* carrying out an experiment exactly what will be measured and what statistical tests you will use on the results (some help is available in Chapter 17 on quantitative data analysis).

Internal validity

Your experiment has good *internal validity* if the measurements you obtain are indeed due to your manipulations of the independent variable, and not to any other factors. Common threats to the internal validity of experiments include:

- **Differences between the experimental and control group**: if they were different to start with, any differences you subsequently measure might not be attributable to your manipulation of the experimental group.
- **History**: events you do not notice interfere between your pre-test and post-test observations. For example, a researcher monitors stress levels in a group of systems analysts in an organization. After implementation of a new timesheet system the researcher measures again and finds a huge increase in stress. However, what the researcher does not know is that between the two tests, the organization

announced a large number of redundancies and plans to relocate their main offices to another country. Many of the analysts are concerned about their future, which is likely to have affected their stress levels.

- **Maturation:** participants change between tests. An obvious example is any experiment involving children. You could test 4-year-olds' competency at some task, and test them again at 6 years old – their performance is likely to have improved regardless of any manipulation you had done during the past 2 years. Maturation can also be a problem with experiments involving adults. The first test might give them practice, or cause boredom or fatigue, or alert them to the purpose of the study. Any of these effects might influence their performance in the second test.
- **Instrumentation:** faulty instruments used to measure the dependent variable will affect the results. This could be inaccurate measuring devices, or interviewers becoming more bored, or more practised, with experience.
- **Experimental mortality:** some subjects might drop out before a study is complete (not necessarily through death). They might leave the company, or be reassigned to another project, or use their right to withdraw from your research. Such drop-outs often lead to bias in your remaining sample of participants. For example, if all the people who drop out are those who are most familiar with a particular technology in a company, you can no longer compare the use of that technology with any new technology. Similarly, some systems development projects may also 'die' – they are abandoned before completion. If you are collecting data about the historical costs of projects, to compare with current or future costs, how will you decide the costs of incomplete projects? Will you take their costs up to the time of abandonment, or estimate costs likely to be incurred if they had continued? It is always important to note who or what dropped out from your study, to see if they bias any subsequent analysis.
- **Reactivity and experimenter effects:** people might change their behaviour as a reaction to being tested. For example, someone tested for speed of data input might subsequently decide to undertake keyboard skills training, and show a much better result in later tests. Participants also often want to help the researcher, or to look good, and so try to give what they hope is the 'right' data – a particular concern when the researcher has more power than the subjects, for example, a teacher carrying out experiments on her students. Experimenters can also unconsciously bias the results they obtain, because of the way they interact with their participants. The psychology literature also discuses the ways in which experimenters' age, race or sex may affect the results they obtain from participants.

External validity

Your experiment has good *external validity* if your results are not unique to a particular set of circumstances but are *generalizable*, that is, the same results can be predicted for subsequent occasions and in other situations. Experimenters seek high external validity. The best way of demonstrating generalizability is to repeat the experiment many times in many situations. You can also design your experiment carefully so that it is likely to have high external validity. The main threat to external validity comes from non-representativeness – using a test case or sample of participants that is not typical of those found elsewhere. For example:

- **Over-reliance on special types of participants:** researchers often use students as the subjects of their experiments. But students are often younger and better educated than the general population, with different personal values and motivations. An experiment's results might be generalizable to other students, but not to people in general. Researchers have sometimes used students in experiments as substitutes for managers, but students have not had the same experiences as most managers and are not subject to the same work pressures and politics as managers, so again such experiments' results are unlikely to be generalizable to managers in organizations.

  Researchers often ask for volunteers to take part in their experiments, but the psychology literature shows that volunteers have certain characteristics that differentiate them from the general population (for example, more intelligent, more sociable, higher respect for scientific researchers). Again results found from an experiment involving volunteer participants may not be generalizable to a wider population.
- **Too few participants:** sometimes researchers do not have enough participants, so that it is impossible to show that a result is statistically significant.
- **Non-representative participants:** we have already seen that students and volunteers may not be representative of a wider population. You need to make sure that your group of participants is typical of the population that you wish to make statements about. You do this by selecting a representative sample by, for example, random sampling or stratified sampling (see Chapter 7 on surveys).
- **Non-representative test cases:** similarly, a researcher investigating, say, the efficiency of a new data processing algorithm, would need to ensure that data files on which an experiment is based are typical of the kinds of data files used in a real-life context.

Quasi- or field experiments

So far our discussion has concentrated on true experiments involving manipulation of the independent variable, pre- and post-test measurement of the dependent variable(s), and control or removal of all other variables. Such experiments usually have to take place in a laboratory so that all variables can be carefully controlled. However, a laboratory is, by its very nature, an artificial situation and may not be typical of situations found elsewhere. This is a particular problem for many IS researchers who are interested in the social context in which information systems are designed and used, and which is usually not replicable in a lab. But if researchers move to a more natural setting, they lose control over the variables and find it far harder to manipulate an independent variable. For example, IS researchers could not deliberately engineer the take-over of one business by another to monitor how that affected information flows within an organization. Some researchers therefore use *quasi-experiments* (quasi means 'as if'). They try to remain true to the spirit of classic laboratory experiments, but concentrate on observing events in real-life settings where there is a 'naturally occurring' experiment. Supposing, for example, a new user input device is invented, to replace the mouse and keyboard. Laboratory experiments may show that users find it comfortable and efficient. However, only long-term use will show whether it causes an increase in cases of repetitive strain injury (RSI). A quasi-experiment (also called a *field experiment*) could therefore be designed. Researchers examine how many cases of RSI were found over the previous five years and then monitor how many cases are found once the new input device is put into use. Any increase in RSI cases *could* be attributable to the new device. However, other causes might be possible, such as an increase in the amount of daily inputs required, or a change in the workforce with less attention given to training employees in how to sit and hold their arms correctly to avoid RSI. In quasi-experiments, there are often variables that the researchers cannot control and that might have caused the measured effect. They can never, therefore, establish cause and effect as conclusively as a true experiment.

The uncontrolled trial

Novice IS and computing researchers sometimes talk of the introduction of a new computer system or a new method as being an experiment. They are making a change and want to see what happens. However, such activity does not meet the tenets of either a true or a quasi-experiment. A systems developer could, for example, introduce a new systems development method. But how will they know the effect of the introduction of the new method? They need to measure something before the introduction of the new method, and again afterwards, and they need to measure and account for other things in the environment that might also cause any change (for example, other variables, the passing of time and so on). Without these measurements and analysis, the developer cannot say that the new method has caused any effect, so they are *not* conducting an experiment, they are conducting an uncontrolled trial, from which little can be concluded. (If you do want to explore what happens when a new computer system or a new method is introduced, case study and action research are often more appropriate strategies. See Chapters 10 and 11.)

Experimental designs

Many different kinds of experimental designs are possible. A particular design is chosen by researchers to ensure that any change in the dependent variable can only

be attributable to the independent variable and nothing else. Some possible designs include:

- **One group, pre-test and post-test:** the participants' performance is measured (for example, data modelling ability), the researchers then apply some treatment (for example, a new method of teaching data modelling), they then measure the participants' performance again. By comparing the before and after scores, the researchers can assess the effects of the treatment. However, they cannot know whether time had an effect – the participants might just have got better over time anyway, without the researchers doing anything.
- **Static group comparison:** the participants are divided into two groups. The researchers apply the treatment to one group and do nothing to the other group. (For example, one class is taught by a traditional method, another via a new method of data modelling.) The performance of both groups is then measured. Differences in outcomes between the groups could be explained by the treatment. However, if participants were not assigned randomly to the two groups, any difference *might* be caused by other factors than the treatment – perhaps all the good students were in just one of the groups.
- **Pre-test/post-test control group:** participants are randomly assigned to one of two groups. Performance is measured before intervention – both groups should perform in the same way if the random assignment has been carried out correctly. Treatment is given to one group. Performance of both groups is again measured. Any difference is assumed to be caused by the treatment. However, pre-testing of the participants might have affected their subsequent performance.
- **Solomon four-group design:** this design controls for the possibility of pre-testing affecting subsequent performance, but it is expensive because of the number of participants needed. Participants are randomly assigned to four groups:
  - Group A is pre-tested, treated and post-tested.
  - Group B is pre-tested, receives no treatment, and post-tested.
  - Group C receives the treatment and is post-tested.
  - Group D receives no treatment and is tested.

Many more designs are possible – see, for example Field and Hole (2003).

The Internet and Experiments

Experiments are possible on how users interact with particular websites, that is, the experiments focus on human–computer interaction. You would need to be certain, however, that each user was seeing the website in the same way – remember different browsers can display websites differently, and some users will have configured their own browsers to their own preferences. It is often best to place the users in a laboratory where you could ensure that all saw the same.

Web-based experiments *are* possible. Usually participants are invited to visit a webpage where they are asked to follow a set of instructions. Their responses and actions are monitored using cookies or the log of a web-server. This means that much of the experiment is automated, and so needs low maintenance and is little affected by the researcher-experimenter. Such web-based experiments are useful where many participants are required, more than could be reached by the researcher in their own neighbourhood, or where specific kinds of subjects are needed, that are geographically dispersed.

Experiments involving computer programs that run over the Internet are, however, difficult because the Internet is not under the control of the researcher – connections, network availability, upload and download speeds, and so on. A researcher cannot even be certain that two emails sent almost simultaneously will take the same route over the Internet to their intended recipient. The difficulties are compounded for researchers who want to experiment on people via the Internet because these users have an offline existence as well as an online existence, and that offline existence cannot be monitored or controlled by the researcher. At best, therefore, researchers will probably have to content themselves with quasi-experiments, and recognize the limitations and possibility of alternative explanations for any cause–effect explanation they offer. For example, a researcher might monitor the number of incidents of flaming (abusive messages) in an online discussion group before and after the posting of a set of 'netiquette' rules. Any decrease might be attributable to the posting of the rules, but the researcher would have to look for other causes such as the main flamers not taking part in the discussion during the time of the experiment (on holiday? not interested? too busy?), or the topic under discussion being one where there was little disagreement anyway among the participants.

Examples of Experiments in IS and Computing Research

Experiments have been used extensively in IS and computing research to investigate human–computer interaction. For example, when the computer mouse was first introduced, a number of studies investigated users' speed of data input, comparing those using a mouse and drop-down menus and those using keyboard control keys.

More recent examples of the use of experiments in IS and computing research include:

- Zendler, Horn, Schwaertzel and Ploedereder (2001) investigated how an expert programmer retrieved reusable software components from a library. They compared different representation techniques used for software components, in order to see whether one representation technique was more effective in aiding retrieval than the other types. They only had one experimental subject, but in their article argue that such a single-case design can be a valid experimental approach.
- Smith, Keil and Depledge (2001) investigated the problem of 'runaway' information systems, where people in an organization are reluctant to report negative news about a project, so it continues, running over time and budget. They used an experiment strategy to find out what factors influence people's willingness to report bad news about a project. This experiment involved role-playing by business students – the researchers discuss whether they think this limits the generalizability of their findings.
- Fencott, van Schaik, Ling and Shafiullah (2002) investigated how users move around in a virtual environment. Their experiments were based in a computer laboratory where some psychology undergraduates were observed navigating around a virtual environment. Each subject was faced with four doors in the virtual environment and had free choice over which door to go through. The doors were identical, except that each had a single black circle moving across its surface in one of four directions: vertically, horizontally, diagonally up or diagonally down. Via experiment the researchers investigated whether a particular type of movement caused more people to use that particular door. A second experiment investigated whether the types of 'reward' found in the different rooms beyond each door induced people to stay longer in that room.

Evaluating Experiment-based Research

Advantages of experiments as a research strategy include:

- They are a well-established strategy, seen by many as the most 'scientific' and therefore most acceptable approach. Where people have not received any formal research methods training, this is often the only research strategy they know.
- They are the only research strategy that can prove causal relationships.
- Laboratory experiments permit high levels of precision in measuring outcomes and analysing the data.
- Laboratory experiments allow researchers to remain at their normal place of work, without the time and costs incurred in visiting field sites.

Disadvantages of experiments include:

- Laboratory experiments often create artificial situations, which are not comparable with real-world situations.
- It is often difficult or impossible to control all the relevant variables.

- It is often difficult to recruit a representative sample of participants.
- It may be necessary to conceal from the participants the purpose of the research, so that they do not skew the results by, for example, performing in the way they think you want them to. However, deception of participants is normally viewed as unethical (see Chapter 5).

Use the 'Evaluation Guide' below to help you analyse and evaluate experiment-based research. If there is insufficient information for you to answer some questions, you should not completely reject the research report, but you should treat the findings with caution and be wary about relying on the report for evidence.

EVALUATION GUIDE: EXPERIMENTS

1. Was a hypothesis or predicted outcome of the experiment clearly stated in the introduction to the research?
2. Was the research a true experiment, a quasi-experiment or an uncontrolled trial?
3. What information is given about the independent and dependent variables manipulated or measured in the study? What additional information would you like?
4. What information is given about any participants and how they were found? What additional information would you like?
5. What information is given about how representative the sample is of the wider population about which conclusions are drawn? Are you satisfied that the sample is representative?
6. What information is given about the apparatus and the process the researchers used to make measurements? What additional information would you like?
7. What limitations in their experiment strategy do the researchers recognize?
8. Given your current state of knowledge, can you identify other flaws or omissions in the researchers' reporting of their experiment?
9. Assuming their statistical analysis is correct, have the researchers convinced you that they have demonstrated cause and effect?
10. Overall, how effectively do you think the experiment strategy has been reported and used?

Experiments involve observations and measurements and analysis of quantitative data. You should therefore also read Chapters 14 ('Observations') and 17 ('Quantitative Data Analysis'). If your experiments involve people as participants, there are clearly limits to what you can ask them to do as compared to what you might do to a computer or a program. You must therefore read Chapter 5 on research participants and ethics.

PRACTICAL WORK

1. Study the following statements and write down the independent and dependent variables in each:
   a. Frequent computer crashes create frustration.
   b. Children learn more from interactive multimedia CDs than from books.
   c. Men are better programmers than women.
   d. Novice programmers spot fewer bugs than experienced programmers.
2. Study the following statements and decide which ones are falsifiable and could therefore be used as hypotheses:
   a. It always rains on the first day of the school holidays.
   b. Women are better at multi-tasking than men.
   c. Either it is sunny or it is not sunny.
   d. All triangles have three sides.
   e. The number of bugs already found in a program is related to the number of bugs still to be discovered.
   f. Companies without a website will suffer reduced sales.
3. Design an experiment (true, quasi- or uncontrolled trial) to investigate each of the following:
   a. Whether a lecturer's use of PowerPoint® slides in a lecture increases students' recall of the subject matter.
   b. Whether managers' use of email increases or decreases when stress levels are high.
   c. Whether object-oriented models are produced quicker than entity-relationship diagrams.
   d. Whether the use of a computer-based knowledge management system leads to better management decision-making.
4. Practise analysing and evaluating experiment-based research. Study a piece of research that used experiments. (Concentrate on the approach used rather than on any statistical analysis of the data.) Answer the questions in the 'Evaluation Guide' above, using the material in this chapter to help you.

FURTHER READING

Field and Hole (2003) is an easy to read, practical and humorous guide to designing and conducting experiments. The book is aimed at psychology students but is useful for researchers in other disciplines too. It covers different experimental designs, how to analyse and interpret the data using statistics, and how to write up the experimental research.

Jarvenpaa, Dickson and DeSanctis (1985) discuss common methodological problems in experimental IS studies, particularly those concerning internal validity. They provide some useful guidelines for experimental IS researchers. Similarly Kitchenham et al. (2001, 2002) argue that empirical studies in software engineering often have poor experimental design, ▶

inappropriate use of statistical techniques and conclusions that do not follow from the results. They therefore offer guidelines for empirical research in software engineering, concentrating on experiments and statistical analysis and drawing from advice written for medical researchers. These guidelines are relevant to other branches of IS and computing too. Tichy (1998) discusses several arguments *against* the use of experiments in computer science, and refutes each of them. Finally, Reips (2002) summarizes what is known so far about the conduct of Internet-based experiments, and gives recommendations on how they can be used effectively.

References

Fencott, C., van Schaik, P., Ling, J., & Shafiullah, M. (2002). The effects of movement of attractors and pictorial content of rewards on users' behaviour in virtual environments. *Interacting with Computers, 15*, 121–140.

Field, A., & Hole, G. (2003). *How to design and report experiments*. London: Sage.

Jarvenpaa, S.L., Dickson, G., & DeSanctis, G. (1985). Methodological issues in experimental IS research: Experiences and recommendations. *MIS Quarterly, 9*(2), 141–156.

Kitchenham, B.A., Pfleeger, S.L., Pickard, L.M., Jones, P.W., Hoaglin, D.C., El-Emam, K., et al. (2001). *Preliminary guidelines for empirical research in software engineering*. Ottawa: National Research Council of Canada.

Kitchenham, B.A., Pfleeger, S.L., Pickard, L.M., Jones, P.W., Hoaglin, D.C., El-Emam, K., et al. (2002). Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering, 28*(8), 721–734.

Reips, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology, 49*(4), 243–256.

Smith, H.J., Keil, M., & Depledge, G. (2001). Keeping mum as the project goes under: Towards an explanatory model. *Journal of Management Information Systems, 18*(2), 189–227.

Tichy, W.F. (1998). Should computer scientists experiment more? *IEEE Computer, 31*(5), 32–40.

Zendler, A., Horn, E., Schwaertzel, H., & Ploedereder, E. (2001). Demonstrating the usage of single-case designs in experimental software engineering. *Information and Software Technology, 43*, 681–691.

10 Case Studies

In this chapter you will learn about:

- what is meant by a case study research strategy;
- issues to address in planning and undertaking case study research;
- how case study research might be applied to the Internet;
- how case studies have been used in IS and computing;
- the advantages and disadvantages of case study research;
- analysing and evaluating case study research.

Defining Case Studies

A case study focuses on one instance of the 'thing' that is to be investigated: an organization, a department, an information system, a discussion forum, a systems developer, a development project, a decision, and so on. This one instance, or case, is studied in depth, using a variety of data generation methods (interviewing, observation, document analysis and/or questionnaires – see Chapters 13–16). The aim is to obtain a rich, detailed insight into the 'life' of that case and its complex relationships and processes.

If using a survey (see Chapter 7), a researcher is able to take a wide but only shallow view of lots of instances of the phenomenon under investigation and is unlikely to obtain much information about the *context* of the phenomenon for each instance. If using an experiment (see Chapter 9), the researcher must divorce a phenomenon from its context, in order to establish that the measured outcomes can only have been caused by the researcher's manipulation of an independent variable and not by