Research methods: Psychological enquiry

- **Experimental method**
  The factors to bear in mind when designing and running a psychology experiment.

- **Quasi-experiments**
  Studies that use the experimental method with pre-existing groups or natural situations.

- **Correlational studies**
  Looking for links between various factors.

- **Naturalistic observation**
  Examining behaviour by watching it take place naturally.

- **Case studies**
  Studying one individual over a period of time.

- **Interviews**
  Obtaining data from participants by asking them questions.

- **Discourse analysis**
  What people say is influenced by the context in which they say it.

In common with other sciences, psychology is concerned with theories and with data. A theory provides a general explanation or account of certain findings or data. It also generates a number of experimental hypotheses, which are predictions or expectations about behaviour based on the theory. For example, someone might propose a theory in which it is argued that some people are more hostile than others. This theory could be used to produce various hypotheses or predictions, such as the following: hostile people will express anger more often than non-hostile ones; hostile people will react more strongly than non-hostile ones to frustrating situations; hostile people will be more sarcastic than non-hostile people.

© 2004 Psychology Press Ltd
Psychologists spend a lot of their time collecting data in the form of measures of behaviour. Data are collected in order to test various hypotheses. Most people assume that this data collection involves proper or true experiments carried out under laboratory conditions, and it is true that literally millions of laboratory experiments have been carried out in psychology. However, psychologists make use of several methods of investigation, each of which has provided useful information about human behaviour.

As you read through the various methods of investigation, it is natural to wonder which methods are the best and the worst. In some ways, it may be more useful to compare the methods used by psychologists to the clubs used by the golf professional. The driver is not a better or worse club than the putter, it is simply used for a different purpose. In similar fashion, each method of investigation used by psychologists is very useful for testing some hypotheses, but is of little or no use for testing other hypotheses. However, as we will see, the experimental method provides the best way of being able to make inferences about cause and effect.

**EXPERIMENTAL METHOD**

The method of investigation used most often by psychologists is the experimental method. In order to understand what is involved in the experimental method, we will consider a concrete example.

**Dependent and independent variables**

Suppose that a psychologist wants to test the experimental hypothesis that loud noise will have a disruptive effect on the performance of a task. As with most hypotheses, this one refers to a dependent variable, which is some aspect of behaviour that is going to be measured. In this case, the dependent variable is some measure of task performance.

Most experimental hypotheses state that the dependent variable will be affected systematically by some specified factor, which is known as the independent variable. In the case we are considering, the independent variable is the intensity of noise. More generally, the independent variable is some aspect of the experimental situation that is manipulated by the experimenter.

We come now to the most important principle involved in the use of the experimental method: the independent variable of interest is manipulated, but all other variables are controlled. It is assumed that, with all other variables controlled, the one and only variable that is being manipulated is the cause of any subsequent change in the dependent variable. In terms of our example, we might expose one group of participants to very intense noise, and a second group to mild noise. What would we need to do to ensure that any difference in the performance of the two groups was due to the noise rather than any other factor? We would control all other aspects of the situation by, for example, always using the same room for the experiment, keeping the temperature the same, and having the same lighting.

**Confounding variables**

Another way of expressing the essence of the experimental method is that it is of fundamental importance to avoid any confounding variables. These are variables that are manipulated or allowed to vary systematically along with the independent variable. The presence of any confounding variables has grave consequences, because it prevents us from being able to interpret our findings. For example, suppose that the participants exposed to intense noise performed the task in poor lighting conditions so that they could hardly see what they were doing, whereas those exposed to mild noise enjoyed good lighting conditions. If the former group performed much worse than the latter group, we would not know whether this was due to the intense noise, the poor lighting, or some combination of the two.
You might think that it would be easy to ensure that there were no confounding variables in an experiment. However, there are many well-known published experiments containing confounding variables. Consider, for example, a study by Jenkins and Dallenbach (1924). They gave a learning task to a group of participants in the morning, and then tested their memory for the material later in the day. The same learning task was given to a second group of participants in the evening, and their memory was tested the following morning after a night’s sleep.

What did Jenkins and Dallenbach find? Memory performance was much higher for the second group than for the first. They argued that this was due to there being less interference with memory when people are asleep than when they are awake. Can you see the flaw in this argument? The two groups learned the material at different times of day, so time of day was a confounding variable. Hockey, Davies, and Gray (1972) discovered many years later that the time of day at which learning occurs is much more important than whether or not the participants sleep between learning and the memory test.

**Participants and settings**

Proper use of the experimental method requires careful consideration of the ways in which the participants are allocated to the various conditions. A detailed account is given in the Research methods: Design of investigations chapter, so we will focus here only on experiments in which there are different participants in each condition. Suppose that the participants exposed to intense noise were on average much less intelligent than those exposed to mild noise. We would then be unable to tell whether poorer performance by the former participants was due to the intense noise or to their low intelligence. The main way of guarding against this possibility is by means of **randomisation**, in which the participants are allocated at random to the two conditions.

Numerous studies are carried out using students as participants. This raises the issue of whether students are representative of society as a whole. For example, it is possible that students would be less distracted than other people by intense noise because they are used to studying over long periods of time in conditions that can be noisy, such as halls of residence.

The experimental method is used mainly in laboratory experiments. However, it is also used in **field experiments**, which are experiments carried out in natural settings such as in the street, in a school, or at work. Some of the advantages of the experimental method are common to both laboratory and field experiments, whereas other advantages and limitations are specific to one type of experiment. We will consider the common advantages next, with more specific advantages and limitations being discussed after that.

In many studies, use is made of pre-existing groups of people. For example, we might compare the performance of males and females, or that of young and middle-aged people.
individuals. Do such studies qualify as genuine experiments? The answer is “No”. Use of the experimental method requires that the independent variable is manipulated by the experimenter, but clearly the experimenter cannot decide whether a given person is going to be male or female for the purposes of the study!

**Common advantages**

**Causal relationships**

What is generally regarded as the greatest advantage of the experimental method is that it allows us to establish cause and effect relationships. In the terms we have been using, the independent variable in an experiment is often regarded as a cause, and the dependent variable is the effect. Philosophers of science have argued about whether or not causality can be established by experimentation. However, the general opinion is that causality can only be inferred. If $y$ (e.g. poor performance) follows $x$ (e.g. intense noise), then it is reasonable to infer that $x$ caused $y$.

We can see why findings from studies based on the experimental method do not necessarily establish causality from the following imaginary example. An experiment on malaria is carried out in a hot country. Half of the participants sleep in bedrooms with the windows open, and the other half sleep in bedrooms with the windows closed. Those sleeping in bedrooms with the windows open are found to be more likely to catch malaria. It would obviously be wrong to argue that having the window open caused malaria. Having the window open or closed is relevant to catching the disease, but it tells us nothing directly about the major causal factor in malaria (infected mosquitoes).

**Replication**

The other major advantage of the experimental method concerns what is known as replication. If an experiment has been conducted in a carefully controlled way, it should be possible for other researchers to repeat or replicate the findings obtained from that experiment. There have been numerous failures to replicate using the experimental method, but the essential point is that the chances of replication are greater when the experimental method is used than when it is not.

**Laboratory vs. field experiments**

Laboratory and field experiments both involve use of the experimental method, but they differ in that field experiments are carried out in more natural settings. As an example of a field experiment, let us consider a study by Shotland and Straw (1976; see PIP, p. 675). They arranged for a man and a woman to stage an argument and a fight fairly close to a number of bystanders. In one condition, the woman screamed, “I don’t know you”. In a second condition, she screamed, “I don’t know why I ever married you!”. When the bystanders thought the fight involved strangers, 65% of them intervened, against only 19% when they thought it involved a married couple. Thus, people were less likely to lend a helping hand when it was a “lovers’ quarrel” than when it was not. The bystanders were convinced that the fight was genuine, as was shown by the fact that 30% of the women were so alarmed that they shut the doors of their rooms, turned off the lights, and locked their doors.

The greatest advantage of laboratory experiments over field experiments is that it is generally easier to eliminate confounding variables in the laboratory than in the field. The experimenter is unlikely to be able to control every aspect of a natural situation.

Another clear advantage of laboratory experiments over field experiments is that it is much easier to obtain large amounts of very detailed information from participants in the laboratory. For example, it is hard to see how information about participants’ physiological activity or speed of performing a range of complex cognitive tasks could be obtained in a field experiment carried out in a natural setting. There are two main reasons why field experiments are limited in this way. First, it is not generally possible to introduce bulky...
equipment into a natural setting. Second, the participants in a field experiment are likely to realise they are taking part in an experiment if attempts are made to obtain a lot of information from them.

One of the advantages of field experiments over laboratory experiments is that the behaviour of the participants is often more typical of their normal behaviour. However, the greatest advantage of field experiments over laboratory experiments is that they are less artificial. The artificiality of laboratory experimentation was emphasised by Heather (1976, pp. 31–33):

Psychologists have attempted to squeeze the study of human life into a laboratory situation where it becomes unrecognisably different from its naturally occurring form ... Experiments in psychology ... are social situations involving strangers, and it might be suggested that the main kind of knowledge gleaned from years of experimentation with human subjects is information about how strangers interact in the highly artificial and unusual setting of the psychological experiment.

The effects of being observed

An important reason why laboratory experiments are more artificial than field experiments is because the participants in laboratory experiments are aware that their behaviour is being observed. As Silverman (1977) pointed out, “Virtually the only condition in which a subject [participant] in a psychological study will not behave as a subject [participant] is if he does not know he is in one.” One consequence of being observed is that the participants try to work out the experimenter’s hypothesis, and then act accordingly. In this connection, Orne (1962) emphasised the importance of demand characteristics, which are “the totality of cues which convey an experimental hypothesis to the subjects.” Orne found that the participants in one of his studies were willing to spend several hours adding numbers on random number sheets and then tearing up each completed sheet. Presumably the participants interpreted the experiment as a test of endurance, and this motivated them to keep going.

Another consequence of the participants in laboratory experiments knowing they are being observed is evaluation apprehension. Rosenberg (1965) defined this as “an active anxiety-toned concern that he [the participant] win a positive evaluation from the experimenter or at least that he provide no grounds for a negative one.”

Sigall et al. (1970) contrasted the effects of demand characteristics and evaluation apprehension on the task of copying telephone numbers. The experimenter told participants doing the task for the second time that he expected them to perform it at a rate that was actually slower than their previous performance. Adherence to the demand characteristics would have led to slow performance, whereas evaluation apprehension and the need to be capable would have produced fast times. The participants actually performed more quickly than they had done before, indicating the greater importance of evaluation apprehension.

This conclusion was strengthened by the findings from a second condition, in which the experimenter not only said that he expected the participants to perform at a slower rate, but also told them that those who rush are probably obsessive-compulsive. The participants in this condition performed the task slowly, because they wanted to be evaluated positively.

Another way in which laboratory experiments tend to be more artificial than field experiments was identified by Wachtel (1973). He used the term implacable experimenter to describe the typical laboratory situation, in which the experimenter’s behaviour (e.g. instructions) affects the participant’s behaviour, but the participant’s behaviour does not influence the experimenter’s behaviour. There are two serious problems with experiments using an implacable or unyielding experimenter. First, because the situation (including the experimenter) is allowed to influence the participant but the participant isn’t allowed to affect the situation, it is likely that the effects of situations on our behaviour are overestimated. Second, because much of the richness of the dynamic interactions between individual and situation has been omitted, there is a real danger that seriously oversimplified accounts of human behaviour will emerge.
Artificiality

How much does it matter that laboratory experiments are artificial? As Coolican (1998) pointed out, “In scientific investigation, it is often necessary to create artificial circumstances in order to isolate a hypothesised effect.” If we are interested in studying basic cognitive processes such as those involved in perception or attention, then the artificiality of the laboratory is unlikely to affect the results. On the other hand, if we are interested in studying social behaviour, then the issue of artificiality does matter. For example, Zegoit et al. (1975) found that mothers behaved in a warmer and more patient way with their children when they knew they were being observed than when they did not.

Carlsthirt, Ellsworth, and Aronson (1976) drew a distinction between mundane realism and experimental realism. Mundane realism refers to experiments in which the situation is set up to resemble situations often found in everyday life. In contrast, experimental realism refers to experiments in which the situation may be rather artificial, but is sufficiently interesting to produce full involvement from the participants. Milgram’s (1974) research on obedience to authority is a good example of experimental realism (see PIP, Chapter 20). The key point is that experimental realism may be more important than mundane realism in producing findings that generalise to real-life situations.

Ethical issues

What we will do here is to discuss a few ethical issues that are of special relevance to laboratory or field experiments. So far as laboratory experiments are concerned, there is a danger that the participants will be willing to behave in a laboratory in ways they would not behave elsewhere. For example, Milgram (1974) found in his work on obedience to authority that 65% of his participants were prepared to give very intense electric shocks to someone else when the experiment took place in a laboratory at Yale University. In contrast, the figure was only 48% when the same study was carried out in a run-down office building. Thus, participants are often willing to do what they would not normally do in the setting of a prestigious laboratory.

Another ethical issue that applies especially to laboratory experiments concerns the participant’s right to withdraw from the experiment at any time. It is general practice to inform participants of this right at the start of the experiment. However, participants may feel reluctant to exercise this right if they think it will cause serious disruption to the experimenter’s research.

So far as field experiments are concerned, the main ethical issue relates to the principle of voluntary informed consent, which is regarded as central to ethical human research. By their very nature, most field experiments do not lend themselves to obtaining informed consent from the participants. For example, the study by Shotland and Straw (1976) would have been rendered almost meaningless if the participants had been asked beforehand to give their consent to witnessing a staged quarrel! The participants in that study could reasonably have complained about being exposed to a violent quarrel.

Another ethical issue with field experiments is that it is not possible in most field experiments to tell the participants that they have the right to withdraw at any time without offering a reason.

Summary

The respective strengths and weaknesses of laboratory experiments and field experiments can be summed up with reference to two different kinds of validity: internal validity and external validity. Internal validity refers to the validity of an experiment within the confines of the context in which it is carried out, whereas external validity refers to the validity of an experiment outside the research situation itself. Laboratory experiments tend to be high in internal validity but low in external validity, whereas field experiments are high in external validity but low in internal validity.

KEY TERMS

Mundane realism: the use of an artificial situation that closely resembles a natural situation.
Experimental realism: the use of an artificial situation in which the participants become fully involved.
Internal validity: the validity of an experiment in terms of the context in which it is carried out.
External validity: the validity of research findings outside the research situation.

Research activity: In small groups, devise a table to summarise the advantages and limitations of the two different approaches: laboratory experiments and field experiments. Choose a selection of topics that psychologists might wish to investigate. Which approach would most suit each example?
QUASI-EXPERIMENTS

“True” experiments based on the experimental method provide the best way of being able to draw causal inferences with confidence. However, it is often the case that there are practical or ethical reasons why it is simply not possible to carry out a true experiment. In such circumstances, investigators often carry out what is known as a quasi-experiment. Quasi-experimental designs “resemble experiments but are weak on some of the characteristics” (Raulin & Graziano, 1994). There are two main ways in which quasi-experiments tend to fall short of being true experiments. First, the manipulation of the independent variable is often not under the control of the experimenter. Second, it is usually not possible to allocate the participants randomly to groups.

There are numerous hypotheses in psychology that can only be studied by means of quasi-experiments rather than true experiments. For example, suppose that we are interested in studying the effects of divorce on young children. We could do this by comparing children whose parents had divorced with those whose parents were still married. There would, of course, be no possibility of allocating children at random to the divorced or non-divorced parent groups! Studies in which pre-existing groups are compared often qualify as quasi-experiments. Examples of such quasi-experiments would be comparing the learning performance of males and females, or comparing the social behaviour of introverted and extraverted individuals.

Natural experiments

The natural experiment is a type of quasi-experiment in which a researcher makes use of some naturally occurring event for research purposes. An example of a natural experiment is a study by Williams (1986) on the effects of television on aggressive behaviour in Canadian children aged between 6 and 11 years. Three communities were compared: one in which television had just been introduced, one in which there was only one television channel, and one in which there were several channels. The children in the first community showed a significant increase in verbal and physical aggression during the first two years after television was introduced, whereas those in the other two communities did not. This was not a true experiment, because the children were not allocated randomly to the three conditions or communities.

Adams and Adams (1984) carried out a natural experiment following the eruption of the Mount St Helens volcano in 1980. As the volcanic eruption had been predicted, they were able to assess the inhabitants of the small town of Othello before and after it happened. There was a 50% increase in mental health appointments, a 198% increase in stress-aggravated illness, and a 235% increase in diagnoses of mental illness.

Advantages and limitations

What are the advantages of natural experiments? The main one is that the participants in natural experiments are often not aware that they are taking part in an experiment, even though they are likely to know that their behaviour is being observed. Another advantage of natural experiments is that they allow us to study the effects on behaviour of independent variables that it would be unethical for the experimenter to manipulate. For example, Adams and Adams (1984) were interested in observing the effects of a major stressor on physical and mental illness. No ethical committee would have allowed them to expose their participants deliberately to stressors that might cause mental illness, but they were able to take advantage of a natural disaster to conduct a natural experiment.

What are the limitations of natural experiments? The greatest limitation occurs because the participants have not been assigned at random to conditions. As a result, observed differences in behaviour between groups may be due to differences in the types of participants in the groups rather than to the effects of the independent variable. Consider, for example, the study by Williams (1986) on television and aggression. The children in the community...
that had just been exposed to television might have been naturally more aggressive than the children in the other two communities. However, the children in the three communities did not differ in their level of aggression at the start of the study.

It is usually possible to check whether the participants in the various conditions are comparable. For example, they can be compared with respect to variables such as age, sex, socio-economic status, and so on. If the groups do differ significantly in some respects irrelevant to the independent variable, then this greatly complicates the task of interpreting the findings of a natural experiment.

The other major limitation of natural experiments involves the independent variable. In some natural experiments, it is hard to know exactly what aspects of the independent variable have caused any effects on behaviour. For example, there is no doubt that the eruption of Mount St Helens was a major stressor. It caused stress in part because of the possibility that it might erupt again and produce more physical devastation. However, social factors were also probably involved. If people in Othello observed that one of their neighbours was highly anxious because of the eruption, this may have heightened their level of anxiety.

**Ethical issues**

It can be argued that there are fewer ethical issues with natural experiments than with many other kinds of research. The reason is that the experimenter is not responsible for the fact that the participants have been exposed to the independent variable. However, natural experiments can raise various ethical issues. First, there can be the issue of informed voluntary consent, in view of the fact that the participants are often not aware that they are taking part in an experiment. Second, experimenters carrying out natural experiments need to be sensitive to the situation in which the participants find themselves. People who have been exposed to a natural disaster such as a volcanic eruption may resent it if experimenters start asking them detailed questions about their mental health or psychological well-being.

**CORRELATIONAL STUDIES**

Suppose that we were interested in the hypothesis that watching violence on television leads to aggressive behaviour. One way of testing this hypothesis would be to obtain two kinds of information from a large number of people: (1) the amount of violent television they watched; and (2) the extent to which they behaved aggressively in various situations. If the hypothesis is correct, then we would expect that those who have seen the most violence on television would tend to be the most aggressive. In technical terms, this study would be looking for a **correlation**, or association, between watching violent programmes and being aggressive. Thus, the closer the link between them, the greater would be the correlation or association.

One of the best-known uses of the correlational approach is in the study of the role of nature and nurture in intelligence. What is done is to assess the intelligence of pairs
of identical or monozygotic and fraternal or dizygotic twins. After that the degree of similarity in intelligence within pairs is worked out by means of a correlation. Identical twins are more alike genetically than fraternal twins. As a result, their levels of intelligence should be more similar than those of fraternal twins if heredity plays an important role in determining intelligence. As predicted, the correlation indicating the degree of similarity in intelligence is nearly always higher for identical twins than for fraternal twins. However, it has proved hard to provide a detailed interpretation of the findings.

Advantages and limitations

Correlational designs are generally regarded as inferior to experimental designs, because it is hard (or impossible) to establish cause and effect. In our example, the existence of an association between the amount of television violence watched and aggressive behaviour would certainly be consistent with the hypothesis that watching violent programmes can cause aggressive behaviour. However, there are other possible interpretations of the data. The causality may actually operate in the opposite direction. In other words, aggressive individuals may choose to watch more violent programmes than those who are less aggressive. There may be a third variable which accounts for the association between the variables of interest, i.e. watching violent programmes and aggressive behaviour. For example, people in disadvantaged families may watch more television programmes of all kinds than those in non-disadvantaged families, and their deprived circumstances may also cause them to behave aggressively. If that were the case, then the number of violent television programmes watched might have no direct effect at all on aggressive behaviour.

CORRELATIONAL STUDIES

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allows study of hypotheses that cannot be examined directly</td>
<td>Interpretation of results is difficult</td>
</tr>
<tr>
<td>More data on more variables can be collected more quickly than in an experimental set-up</td>
<td>Cause and effect cannot be established</td>
</tr>
<tr>
<td>Problems of interpretation are reduced when no association is found</td>
<td>Direction of causality is uncertain</td>
</tr>
<tr>
<td>Even when strong correlations are found it may be obvious that no causal relationship exists</td>
<td>Variables other than the one of interest may be operating</td>
</tr>
</tbody>
</table>

In spite of the interpretive problems posed by the findings of correlational studies, there are several reasons why psychologists continue to use this method. First, many hypotheses cannot be examined directly by means of experimental designs. For example, the hypothesis that smoking causes a number of physical diseases cannot be tested by forcing some people to smoke and forcing others not to smoke! All that can be done is to examine correlations or associations between the number of cigarettes smoked and the probability of suffering from various diseases.

Second, it is often possible to obtain large amounts of data on a number of variables in a correlational study much more rapidly and efficiently than would be possible using experimental designs. Use of a questionnaire, for example, would permit a researcher to investigate the associations between aggressive behaviour and a wide range of activities (such as watching violent films in the cinema; reading violent books; being frustrated at work or at home).

Third, interpretive problems are much reduced if there is no association between two variables. For example, if it were found that there was no association at all between the amount of violent television watched and aggressive behaviour, this would provide fairly strong evidence that aggressive behaviour is not caused by watching violent programmes on television.

Fourth, the interpretive problems with associations or correlations between two variables are often not as great as in the example of violent programmes and aggression.
Suppose, for example, we discover a correlation between age and happiness, in which older people are generally less happy than younger people. Although it would not be possible to offer a definitive interpretation of this finding, we could be entirely confident that unhappiness does not cause old age!

**Ethical issues**

Correlational analyses are used very widely. As a result, it is not possible to identify any particular ethical issues that apply to most studies in which such analyses are carried out. However, correlational analyses are often used in socially sensitive research, which raises political and/or social issues. For example, consider the correlational evidence suggesting that individual differences in intelligence depend in part on genetic factors. Some people have argued mistakenly that this implies that race differences in intelligence also depend on genetic factors. The key ethical issue here (and in many other correlational studies) is for the researcher to be fully aware of the social sensitivity of the findings that he or she has obtained.

Another ethical issue is raised by the real possibility that the public at large will misinterpret the findings from correlational studies. For example, the finding that there was a correlation between the amount of television violence watched by children and their level of aggression led many influential people to argue that television violence was having a damaging effect. In other words, they mistakenly supposed that correlational evidence can demonstrate a causal relation. Television companies may have suffered from such over-interpretation of findings.

**NATURALISTIC OBSERVATION**

Naturalistic observation involves methods designed to examine behaviour without the experimenter interfering with it in any way. This approach was originally developed by the ethologists such as Lorenz and Tinbergen. They studied animals in their natural habitat rather than in the laboratory, and discovered much about their behaviour. An example of the use of naturalistic observation in human research is the work of Brown, Fraser, and Bellugi (1964). They studied the language development of three children (Adam, Eve, and Sarah) by visiting them at home about 35 times a year.

One of the key requirements of the method of naturalistic observation is to avoid intrusion. Dane (1994, p. 1149) defined this as “anything that lessens the participants’ perception of an event as natural.” There are various ways in which intrusion can occur. For example, there will be intrusion if observations are made in an environment that the participants regard as a research setting. There will also be intrusion if the participants are aware that they are being observed. In many studies, the experimenter is in the same room as the participants, and so they are almost certain to realise they are being observed. When this is the case, the experimenter may try to become a familiar and predictable part of the situation before any observations are recorded.

The participants in naturalistic observation often display a wide range of verbal and non-verbal behaviour. How can observers avoid being overloaded in their attempts to record this behaviour? One approach is to focus only on actions or events that are of particular interest to the researcher; this is known as event sampling. Another approach is known as time sampling, in which observations are only made during specified time periods (e.g. the first 10 minutes of each hour). A third approach is point sampling, in which one individual is observed in order to categorise their current behaviour, after which a second individual is observed.

In considering the data obtained from naturalistic observation, it is important to distinguish between recording and interpretation or coding. For example, an observer may record that the participant has moved forwards, and interpret that movement as an
aggressive action. In practice, however, observers typically only focus on interpreting or coding the participants’ behaviour. For example, Bales (1950) developed the interaction process analysis, which allows observers to record inferred meanings for the forms of behaviour shown by members of a group (e.g. “offers suggestion”).

There have been various attempts to develop ways of categorising people’s behaviour in naturalistic observation without interpreting it. For example, McGrew (1972) devised a detailed and comprehensive recording system to place the social interactions of children at nursery school into 110 categories.

**Advantages and limitations**

What are the advantages of naturalistic observation? First, if the participants are unaware that they are being observed, then it provides a way of observing people behaving naturally. When this happens, there are no problems from demand characteristics, evaluation apprehension, the implacable experimenter, and so on. Second, many studies based on naturalistic observation provide richer and fuller information than typical laboratory experiments. For example, participants’ behaviour may be observed in a range of different social contexts rather than on their own in the laboratory. Third, it is sometimes possible to use naturalistic observation when other methods cannot be used. For example, the participants may be unwilling to be interviewed or to complete a questionnaire. In the case of participants being observed at work, it may be impossible to obtain permission to disrupt their work in order to carry out an experiment.

### NATURALISTIC OBSERVATION

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>People tend to behave naturally</td>
<td>Experimenter has no control over the situation</td>
</tr>
<tr>
<td>Information that is gathered is rich and full</td>
<td>Participants can be aware of being watched and</td>
</tr>
<tr>
<td>Can be used where other methods are not possible</td>
<td>this can affect behaviour</td>
</tr>
<tr>
<td></td>
<td>Problems of reliability due to bias or imprecise</td>
</tr>
<tr>
<td></td>
<td>categorisation of behaviour</td>
</tr>
<tr>
<td></td>
<td>Problems of validity due to observers’ or coders’</td>
</tr>
<tr>
<td></td>
<td>assumptions</td>
</tr>
<tr>
<td></td>
<td>Replication is not usually possible</td>
</tr>
</tbody>
</table>

What are the limitations of naturalistic observation? These are some of the major ones:

- The experimenter has essentially no control over the situation; this can make it very hard or impossible to decide what caused the participants to behave as they did.
- The participants are often aware that they are being observed, with the result that their behaviour is not natural.
- There can be problems of reliability with the observational measures taken, because of bias on the part of the observer or because the categories into which behaviour is coded are imprecise. Attempts to produce good reliability often involve the use of very precise but narrow categories, leading to much of the participants’ behaviour simply being ignored. Reliability can be assessed by correlating the observational records of two different observers. This produces a measure of inter-rater reliability.
- The fact that observations are typically interpreted or coded prior to analysis can cause problems with the validity of measurement. For example, it may be assumed invalidly that all instances of one child striking another child represent aggressive acts, when in fact many of them are only playful gestures. Thus, great care needs to be taken in operationalisation, which is a procedure in which a variable (e.g. aggressive act) is defined by the operations taken to measure it.
- There are often problems of replication with studies of naturalistic observation. For example, the observed behaviour of children in a school may depend in part on the fact that most of the teachers are very lenient and fail to impose discipline. The findings might be very different at another school in which the teachers are strict.

**Key Term**

**Operationalisation:** a procedure in which variables of interest are defined by the operations taken to measure them.
Ethical issues

Naturalistic observation poses ethical problems if the participants do not realise that their behaviour is being observed. In those circumstances, they obviously cannot give their voluntary informed consent to be involved in the study. There can also be problems about confidentiality. Suppose, for example, that naturalistic observation takes place in a particular school, and the published results indicate that many of the children are badly behaved. Even if the name of the school is not mentioned in the report, many people reading it will probably be able to identify the school because they know that the researchers made detailed observations there.

CASE STUDIES

The great majority of studies in psychology have involved the use of experimental or correlational methods on groups of participants. These approaches permit the use of statistical techniques providing information about the extent to which the results obtained from a given sample can be generalised to some larger population.

There are often good reasons why it is not feasible to use numerous participants in a study. For example, a brain-damaged patient may have a very unusual pattern of impaired performance, and there may not be other patients having the same pattern. Another example might be a therapist who has a patient with a rare mental disorder, but there is no possibility of him or her collecting data from other patients with the same disorder. In such circumstances, it can be very useful to carry out a case study, in which one individual is investigated thoroughly and over a period of time.

Some researchers have argued that the study of individual cases can be more fruitful than the study of groups of participants. One of the most convincing statements of that argument was put forward by Gordon Allport (1962):

Why should we not start with individual behaviour as a source of hunches ... and then seek our generalisations but finally come back to the individual not for the mechanical application of laws but for a fuller and more accurate assessment than we are now able to give? ... We stop with our wobbly laws of generality and seldom confront them with the concrete person.

Some of those who have favoured single-case studies have been of an anti-scientific persuasion. However, a prominent experimentalist who advocated the use of single-case studies was the behaviourist B.F. Skinner. In a discussion of research on operant conditioning, Skinner (1966) argued, “instead of studying a thousand rats for one hour each, or a hundred rats for ten hours each, the investigator is likely to study one rat for a thousand hours.”

There are several types of case studies, and they are carried out for various reasons, some of which will be considered here. One reason is to test a current theory. For example, Atkinson and Shiffrin (1968) argued that information only enters the long-term memory store via the short-term memory store (see PIP, Chapter 9). As a result, a brain-damaged individual with impaired short-term memory should also have impaired long-term memory. Evidence that seemed to be inconsistent with this theory was reported in a case study on KF, who was involved in a motorcycle accident (Shallice & Warrington, 1970). He had very poor short-term memory for words and digits, but his long-term learning and recall were unaffected.

Case studies can also be used to refine theories. Baddeley and Hitch (1974) argued that people possess an articulatory loop which is used in the rehearsal of verbal information (see PIP, Chapter 9). It used to be assumed that rehearsal within the articulatory loop requires use of the speech muscles. However, Baddeley and Wilson (1983) carried out a case study on a student, GB. He suffered from anarthria, which meant that he could not use his speech muscles and was unable to speak. In spite of this disorder, GB was able to make use of the articulatory loop.

**KEY TERM**

Case study: detailed investigation of a single individual.
Some case studies are based on very unusual individuals. For example, there was Chris Sizemore, who was the central character in the film *The Three Faces of Eve* (see PIP, p. 109). Some of the time she was Eve White, a well-behaved and inhibited woman. At other times, she was Eve Black, who was promiscuous and impulsive. At still other times, she was Jane, who was more stable than either of the other two personalities. The existence of individuals with multiple personalities raises issues about our usual assumption that everyone has one personality and one self.

**Advantages and limitations**

What are the advantages of case studies? First, as we have seen, a single-case study can provide good evidence that a particular theory is in error. Of course, it is then desirable to find and test other individuals to check the findings from the first case study. Second, a case study can help to refine our theoretical understanding. Third, case studies can provide rich information that is used by the researcher or therapist to develop new theoretical ideas. An example of this is the case of Dr Schreber, which was discussed earlier. Fourth, case studies can provide information about exceptional types of behaviour or performance that had been thought to be impossible.

What are the limitations of case studies? The greatest limitation is their typically low reliability. The findings that are obtained from one unusual or exceptional individual are unlikely to be repeated in detail from another individual. Thus, it is often very hard to generalise from a single-case study. Second, many case studies involve the use of lengthy, fairly unstructured, interviews. Such case studies share the limitations that are identified for interviews in the next section. Third, researchers generally only report some of the data they obtained from their interviews with the participant. They may be unduly selective in terms of what they choose to report or to omit.

**Ethical issues**

Case studies with clinical patients can pose important issues about confidentiality. A therapist such as Sigmund Freud may want to publish details of his case studies because they seem to support his theoretical position. However, the patient may be very unwilling for personal information about him or her to be published. Case studies with brain-damaged patients can pose ethical issues about voluntary informed consent. For example,
patients with severe language impairments may find it hard to understand what will be involved in a case study, and so they cannot give proper informed consent.

INTERVIEWS

As Coolican (1994) pointed out, there are various kinds of interview which vary enormously in terms of the amount of structure they contain. In what follows, we will make use of his categorisation of different types of interview.

Non-directive interviews

Non-directive interviews possess the least structure, with the person being interviewed (the interviewee) being free to discuss almost anything he or she wants. The role of the interviewer in non-directive interviews is to guide the discussion and to encourage the interviewee to be more forthcoming. This type of interview is used very often in psychotherapy, but has little relevance to research.

Informal interviews

Informal interviews resemble non-directive interviews, in that the interviewer listens patiently and focuses mainly on encouraging the interviewee to discuss issues in more depth or detail. However, informal interviews differ in that there are certain general topics that the interviewer wishes to explore. One of the best-known examples involving informal interviews was a large-scale study of workers at the Hawthorne works of Western Electric. The aim of this study was to explore industrial relations via a series of interviews. What emerged from informal interviews was that the relatively minor issues initially raised by the workers generally reflected deeper and more serious worries (Roethlisberger & Dickson, 1939; see p. 20 of the Research methods: Design of investigations chapter).

Guided interviews

Informal but guided interviews possess a little more structure than informal interviews. The interviewer identifies beforehand the issues to be addressed, but how and when to raise those issues is decided during the course of the interview. Structured but open-ended interviews use a formal procedure in which all interviewees are asked precisely the same questions in the same order. Such a procedure prevents the interviewee from side-tracking the interview and taking control of it away from the interviewer. The interviews are open-ended, in the sense that the questions that are asked allow plenty of scope for various kinds of answers (e.g. “How do you see your career developing?”).

Clinical interviews

The clinical interview or clinical method resembles the structured but open-ended interview. In essence, all of the interviewees or participants are asked the same questions, but the choice of follow-up questions depends on the answers that are given. Piaget made much use of the clinical method in his research on cognitive development in children (see PIP, Chapter 15). Piaget understood that children might perform poorly on a task because they did not understand fully what the experimenter wanted them to do. One way of trying to avoid this problem was by giving the experimenter the flexibility to ask questions in various ways. In spite of this, critics of Piaget have argued that the children he studied often failed to solve problems because of the complex language used by the experimenter.

Fully structured interviews

Finally, there is the fully structured interview. In this type of interview, a standard set of questions is asked in the same fixed order to all of the interviewees, and they are only
allowed to choose their answers from a restricted set of possibilities (e.g. “Yes”; “No”; “Don’t know”). As Coolican (1994, pp. 121–122) pointed out, “this approach is hardly an interview worth the name at all. It is a face-to-face data-gathering technique, but could be conducted by telephone or by post.”

Advantages and limitations

What are the advantages of the interview method? As might be expected, the precise advantages depend on the type of interview. Relatively unstructured interviews have the advantage that they are responsive to the personality, interests, and motivations of the interviewee. In principle, they can perhaps reveal more about the interviewee than is likely to be the case with more structured interviews. One of the advantages of fairly structured interviews is that it is easy to compare the responses of different interviewees, all of whom have been asked the same questions. Another advantage is good reliability, in that two different interviewers are likely to obtain similar responses from an interviewee when they ask exactly the same questions in the same order. A further advantage is that there is a reasonable probability of being able to replicate or repeat the findings from a study using structured interviews. Finally, structured interviews have the advantage that it is usually fairly easy to analyse the data obtained from them.

What are the limitations of the interview method? So far as unstructured interviews are concerned, there is the problem that the kinds of information obtained from different interviewees vary in an unsystematic way. As a result, the data from unstructured interviews tend to be hard to analyse. A further limitation with unstructured interviews is that what the interviewee says is determined in a complex way by the interaction between him or her and the interviewer. In other words, the personality and other characteristics of the interviewer typically influence the course of the interview, and make it hard to work out which of the interviewee’s contributions are and are not affected by the interviewer. Finally, the fact that the information obtained from interviewees in unstructured interviews is influenced by the interviewer means that the data obtained can be viewed as unreliable.

One of the main limitations with structured interviews is that what the interviewee says may be somewhat constrained and artificial because of the high level of structure built into the interview. Another limitation is that there is little or none of the flexibility associated with unstructured interviews.

Finally, we need to consider three limitations that are common to all types of interview. First, there is the issue of social desirability bias. Most people want to present a favourable impression of themselves to other people, and this may lead them to distort their answers to personal questions. For example, people are much more willing to admit that they are unhappy when filling in a questionnaire anonymously than when being interviewed (Eysenck, 1990). Second, interviews can only extract information of which the interviewee is consciously aware. This is a significant limitation, because people are often unaware of the reasons why they behave in certain ways (Nisbett & Wilson, 1977). Third, there is the limitation that many interviewers lack some of the skills necessary to conduct interviews successfully. Good interviewers are able to make an interview seem natural, they are sensitive to non-verbal cues, and they have well-developed listening skills (Coolican, 1994).

Ethical issues

Interviews (especially clinical interviews) are often concerned with personal issues about which the interviewee is sensitive. This clearly raises the issue of confidentiality. There are various ways in which confidentiality can be broken. For example, Coolican (1994) discussed a study by Vidich and Bensman (1958) in which direct quotations from interviewees
in Springdale in the United States were published. Made-up names were used, but the people of Springdale were able to identify the actual individuals on the basis of what they said.

Confidentiality can also be broken if a detailed written account or video recording of an interview falls into the wrong hands. Finally, of course, the interviewer himself or herself may disclose sensitive personal information about the interviewee to other people.

There is another ethical issue that is of particular importance with structured interviews. Interviewees may be aware that several other interviewees are being asked the same questions, and that their answers will be compared. As a result, some interviewees may feel that they must answer embarrassing questions in order not to spoil the experiment.

**DISCOURSE ANALYSIS**

According to Potter and Wetherell (1987), discourse analysis is concerned with “all forms of spoken interaction, formal and informal, and written texts of all kinds.” The basic underlying assumption is that the ways in which we use language are greatly affected by the social context. Thus, for example, when politicians give speeches, it would be naive to assume that what they say simply reflects their genuine beliefs and views. It is generally accepted that what they say is designed to have certain effects on their audience, on other politicians, and on the public.

There is much evidence to indicate that people do adjust what they say or write to fit the circumstances. For example, consider studies using the bogus pipeline. The participants are wired up to an impressive-looking machine (the bogus pipeline), and informed that it can detect any lies they produce. Most white participants express more negative attitudes towards black people when wired up to the bogus pipeline than under standard conditions. The implication is that the attitudes that people express normally are constructed so as to be socially acceptable to other people.

Gilbert and Mulkay (1984) carried out a discourse analysis based on interviews with 34 scientists. The importance of social factors in discourse was revealed by comparing what these scientists said during the interviews with their academic publications. The general pattern was for scientists to be much more confident about the meaning of their findings when interviewed than they were in their writings.

As Curtis (1997, p. 24) pointed out, “The idea that there is one way to perform discourse analysis is both naive and illusory.” Nevertheless, he identified seven features that are often found in discourse analysis:

1. Select some written or spoken material that is relevant to the issues you want to study.
2. Read or listen to the discourse several times, trying to decide how it has been constructed. Account needs to be taken of the social context in which it was produced.

**Key Term**

Discourse analysis: a qualitative form of analysis applied to language productions in spoken or written form.
3. Develop a qualitative coding system focusing on the functions or purposes that seem to be served by the discourse.

4. Produce some tentative hypotheses about the purposes served by the discourse, but be willing to modify these hypotheses if subsequent analysis indicates that they are inadequate.

5. How has the person producing the discourse tried to legitimise or make persuasive his or her version of events?

6. Examine the discourse for evidence of extreme case formulations. People often use extreme terms (e.g. always; never) to make their preferred interpretation seem more persuasive.

7. Examine the discourse carefully to see whether the purposes or functions it serves vary from one part to another.

Advantages and limitations

One of the advantages of discourse analysis is that it is based on the correct assumption that our use of language is often much influenced by the social context. This is true of how we remember events in our lives as well as our expressed attitudes and beliefs. As Coolican (1994, p.178) pointed out:

*When we remember and attribute in real life, as opposed to the psychology experiment, our accounts attend to blame, defence, accountability, explanation and so on. What we often do is to present rememberings as facts when they are really constructions.*

For example, the way you describe events in your life is likely to vary depending on whether you are talking to your parents, to your best friend, or to an acquaintance.

Another advantage is that discourse analysis focuses on the ways in which language is used in real-life settings. As such, it avoids much of the artificiality of most experiments. In addition, the claims of those who favour discourse analysis that language is the primary mode of communication among human beings are correct.

There are several limitations of discourse analysis, many of which were discussed by Burman and Parker (1993). A major limitation is that the validity of discourse analysis is open to considerable doubt, and that procedures for assessing validity are lacking. For example, if two researchers interpret a given piece of discourse in very different ways, we cannot be sure which of them has produced the more valid interpretation.

A further limitation is that we often have little information about the reliability or consistency with which the discourse analysis has been carried out. When such information is available, it frequently indicates that reliability is low (Coolican, 1994).

Another limitation is that what emerges from discourse analysis may be unduly influenced by the views and beliefs of the researcher. As Human (1992) expressed it, discourse analysis is sometimes simply “a researcher’s ideas with examples.” A key reason why this can happen is because there are so few constraints on the researcher as he or she tries to make sense of any given written or spoken discourse.

A final limitation is that discourse analysis is based solely on the analysis of language in its various forms. However, language is by no means the only means of communication open to people. Account needs to be taken of non-verbal communication of various kinds (e.g. body language).

Ethical issues

It is often important for the researcher to make sure that anyone whose discourse is to be analysed has given their permission for it to be used for that purpose. However, that ethical issue may not arise if the discourse is in the public domain (e.g. a speech or television interview given by a politician).
There can also be ethical issues if the researcher’s proposed interpretation is likely to offend those who provided the discourse. For example, Wetherell and Potter (1988) carried out discourse analysis on interviews with white New Zealanders. They concluded that those interviewed had racist attitudes towards the Maoris, although they did not directly say so in the interviews. In such circumstances, it is important for the researcher to discuss his or her proposed interpretation with the participants before the results of the study are published or made generally available.

A final ethical issue stems from the fact that discourse analysis often involves detailed analysis of an individual’s discourse. As a result, it is sometimes impossible to adhere to the ethical principles that the information provided by participants should be confidential and that individuals should not be identifiable.

**PERSONAL REFLECTIONS**

- One of the key developments in psychology over the past 50 years has been the gradual increase in the range of research methods used by psychologists. Indeed, the change has been so marked that it is hard to believe that much research used to consist of studying rats in mazes and Skinner boxes! Some psychologists have tried to start a controversy between “traditional” researchers and those who favour the newer methods such as discourse analysis. My view is that both approaches have significant advantages: the key issue is to use whatever research method is best suited in the particular circumstances.

**SUMMARY**

**Experimental method**

The key principle of the experimental method is that the independent variable is manipulated (with all other variables controlled) in order to observe its effect on some dependent variable. In other words, it is important to avoid confounding variables. The experimental method is used in laboratory and field experiments. Use of the experimental method allows us to infer causality, and it often permits replication. Laboratory experiments have various advantages over field experiments: it is usually easier to eliminate confounding variables, and to obtain detailed behavioural and physiological information. The greatest advantage of field experiments over laboratory experiments is that they are less artificial, and suffer less from factors such as demand characteristics, evaluation apprehension, and the implacable experimenter.

**Quasi-experiments**

Quasi-experiments fall short of true experiments either because the experimenter has not manipulated the independent variable or because the participants are not allocated at random to conditions. Natural experiments are quasi-experiments involving some naturally occurring event. Advantages of natural experiments include the possibility that the participants will not be aware they are taking part in an experiment and the opportunity to study the effects of very stressful events. Limitations include problems of interpreting the findings due to a lack of randomisation or to the use of complex independent variables.

**Correlational studies**

Correlational designs are inferior to experimental designs, because they do not permit inferences about causality. However, many issues can only be studied by assessing correlations or associations between variables. It is often possible to obtain large amounts of data very rapidly in correlational studies. The problems of interpretation are much reduced if there is no correlation or association between two variables.
Naturalistic observation involves the use of methods designed to assess behaviour without the experimenter interfering in any way. Methods of data collection include event sampling, time sampling, and point sampling. We should distinguish between data recording and interpretation or coding. Naturalistic observation can provide rich and full information from people who are unaware that they are being observed. However, the experimenter has essentially no control over the situation, the participants are often aware they are being observed, and there can be problems with the reliability and validity of measurement.

A single individual is investigated thoroughly in a case study. Case studies can be carried out to test a current theory, to refine a theory, to permit the development of new theoretical ideas, and to reveal the exceptional characteristics of certain individuals. Case studies generally have very low reliability, and this makes it hard to generalise from a single-case study. Case studies based on interviews often suffer from the limitation that what the participant says is determined in part by the interviewer or researcher, who may then be too selective in what he or she reports of the interview.

There are several types of interview ranging from the unstructured to the totally structured. Unstructured interviews are responsive to the personality, interests, and motivations of the interviewee, but the data obtained tend to be unreliable. In contrast, structured interviews permit comparisons among interviewees, and they tend to be fairly reliable, but what the interviewee says can be constrained and artificial. All types of interviews can produce problems due to social desirability bias, and interviewees can only provide information of which they are consciously aware.

Discourse analysis is based on the assumption that our use of language is much affected by the social context. It involves a careful analysis to identify the underlying purposes of the person who produced the discourse, using a qualitative coding system. Limitations of discourse analysis include low validity and reliability, and the danger that the views of the researcher will influence the findings too much. Ethical issues arise unless the permission of anyone providing discourse for analysis is obtained, and it can be hard to maintain confidentiality of the data.

**FURTHER READING**


**REVISION QUESTIONS**

Sample questions on research methods are given at the end of the Research methods: Data analysis chapter.

**REFERENCES**


