

Loftus and Palmer (1974) Reconstruction of automobile destruction. Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in this study?

The method used in this study was **laboratory experiments**. Loftus and Palmer manipulated the words used in the questions and measured the effect of this on recall. They did this in **controlled conditions**, keeping as many other variables as possible the same, thus allowing them to conclude that it was the words used in the questions that caused the differences in recall. This illustrates one of the major strengths of laboratory experiments, control. The more variables you have control over, the easier it becomes to draw conclusions about the effect of the individual variable on the dependent variable. However, laboratory experiments have several weaknesses. The high level of control usually means that you have created an artificial situation which makes it difficult to apply your results to everyday life. This is referred to as **low ecological validity** and is discussed further below. Participants know that they are taking part in a laboratory experiment and this will affect their behaviour in a number of ways. They will be looking for clues as to how to behave (**demand characteristics**) and they will usually want to help the experimenters by giving them the results that they think they want. Demand characteristics may have had a major effect on the results of Loftus and Palmer's first study.

What type of data was collected in this Study?

Loftus and Palmer only collected **quantitative data**, that is speed estimates and numbers of people saying that they had seen a broken headlight. Quantitative data is very useful for making comparisons and it allows statistical analysis to be conducted but it is fairly superficial and does not tell us anything about why people gave the answers that they did.

How ecologically valid was this research?

There are many differences between observing an event like a car crash in real life and observing one on a television screen as part of an experiment. Some of these differences are:

- In an experiment, you are **expecting** something to happen and may be paying more **attention** to what is going on.
- The event is likely to be far **less distressing** on film than it would be in real life. This may affect our memory for the event.
- In an experiment, you may well expect to be asked questions about what you are watching and this may make you attend to the film in a different way.
- In real life, there may be **consequences** arising from the answers that you give and this may put pressure on the witness.

Overall, we can probably conclude that this laboratory experiment had **low ecological validity** and thus may not tell us very much about how people's memories are affected by leading questions in real life.

How representative were the sample of participants?

Loftus and Palmer used **student participants** in both of the studies. It may be that students are **not representative** of the general population and therefore it may be difficult to **generalise** from the results of this study to people in general. Some differences between students and the general population may include:

- Students will usually be **young** and it is possible that people's memories are better when they are young.
- Students are used to taking in lots of information and then being asked questions about it. People who have not studied for many years may be less used to this.
- Students may be **less experienced drivers** than the general population and may be less confident in their ability to estimate speed. This may have led them to be more swayed by the words in the questions.
- Students may be more susceptible to demand characteristics especially if they are students of the researchers conducting the research.

How useful was this research?

Given that it has just been argued that the study had an **unrepresentative sample** and had **low ecological validity**, you might feel that it was not very useful. However, despite the points made above, it is possible to identify a number of **practical applications** that arise from Loftus and Palmer's research. The conclusion that leading questions can affect memory has important implications for **interviewing witnesses**, both by police immediately or soon after an event and also by lawyers in court some time later. Interviewers should avoid leading questions and should be careful to word questions in ways that do not suggest an answer to the person that they are interviewing.

**Deregowski (1972) Pictorial perception and culture.
Thinking like a Psychologist - Evaluating the Core Study**

What are the strengths and weaknesses of the method used in this Study?

This is a complex question to address as there so many different aspects to the article. We will start by considering the strengths and weaknesses of Deregowski's review. **Review** articles are very **useful** as the author brings together several pieces of research and draws some general **conclusions** from them. Often authors of reviews are also evaluating the research conducted by others and again this makes them valuable sources of information about the strengths and weaknesses of the original research. However, Deregowski has not conducted the research he is writing about and may be misrepresenting or misunderstanding the research he is discussing. For example, Banyard (1996) suggests that Deregowski may be slightly **inaccurate** in his reporting of the results from the elephant - antelope picture.

The review also contains some **anecdotal evidence** and this will **lack the objectivity** of controlled research. These reports came from missionaries working in Africa around one hundred years ago and some of the **ethnocentric** nature of their attitudes is evident in the quotes included by Deregowski. Ethnocentrism is discussed further below. There are no references given for these anecdotes and so we cannot research them further, in some cases they **lack detail** and it would be useful to know more about the event that they are describing. What exactly was flashed onto the sheet? Was it a picture or a photograph? If it was a picture, how detailed was it?

What type of data was collected in the Study?

The research conducted by Hudson collected **quantitative data**, that is numbers of people who were or were not able to perceive depth cues in pictures. However, some of the anecdotal evidence is **qualitative**, that is it simply describes experience in words.

Is the research ethnocentric?

Ethnocentrism is defined as the tendency to judge people by the norms and standards of your own cultural group. The research described by Deregowski is ethnocentric in the way that it refers to people of other cultures as being less developed in their ability to perceive pictures. The researchers appear to be failing to consider that for many of their subjects, this would have been the first time they had seen western style drawings

Cross cultural research is often ethnocentric, as researchers are so used to their own cultures that it is difficult for them to be objective about a different culture. It is highly likely that they will interpret what they are seeing in relation to the norms and values of their own culture. This can also affect the way that research is conducted. Some of the subjects used in Hudson's research may have been seeing paper for the first time, or asked to use a pen and paper for the first time. They will have had no previous experience of these types of tasks and it is unlikely that they would have understood the purpose of the research. There may also be ethical issues such as informed consent raised by such research.

What does the research tell us about the nature - nurture debate?

This review tells us that perception of pictures is not an innate skill. Rather, it is the result of **learning and experience**. Someone growing up in a carpentered environment will easily pick up on depth cues in pictures but for those people who live in very different environments these cues are meaningless. This is a very important conclusion as it not only suggests that pictures are not a lingua franca for inter-cultural communication but reveals the extent to which our perceptual abilities are learned ones.

Is the Study useful?

In the sense that we can draw the conclusions outlined above, this review may be considered to be a useful addition to our knowledge about perceptual processes. This review also highlights the dangers of ethnocentrism particularly when conducting cross cultural research. However, we may be able to go further than that and suggest some practical applications that arise from this review. If pictures or images are to be used cross culturally, then it is important that they are designed in ways which will be understood in the same way in all cultures.

Baron-Cohen, Leslie and Frith.(1985) Does the autistic child have a theory of mind? Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in this Study?

This study used a **quasi-experimental design**. This is a design where the researcher does not have control over the independent variable but takes advantage of conditions where the **different conditions of the independent variable occur naturally**. In this study, the researchers could not randomly allocate people to the three groups but had to find autistic, Down's syndrome and developmentally normal children who were able to take part in the research. Quasi experimental designs are **useful** as they allow researchers to investigate naturally occurring variables that can't be manipulated experimentally but they lack some of the **control** of laboratory experiments. It is possible that there are other differences between the groups of participants which are unrelated to the variables being studied.

What type of data is collected in the study?

The data is **quantitative** as it is simply numbers of children that got the belief question correct. This is appropriate data to collect in this study as the researchers were simply proposing that more autistic children would get this question wrong. However, it would have been possible to explore the children's answers further (perhaps with some open ended questions such as 'why will Sally look there?') which would generate qualitative data. However, this would depend on the language skills of the children being tested.

How ecologically valid is the study?

It is possible to criticise the ecological validity of the study in one important respect. In everyday life, we do not attribute thoughts and beliefs **to dolls** but to other people. Why did the researchers choose to explore theory of mind in this way? If autistic children cannot attribute beliefs to dolls, this does not necessarily mean that they can't attribute beliefs to people. However, Leslie and Frith (1988) replicated the procedure used in the Sally Anne task with real people and found similar results, suggesting that the Sally Anne task is a **valid one** and can be generalised to attributing beliefs to other people.

The use of dolls may be criticised in other ways. Some of the children in this study were eleven years old or more and if the figures quoted earlier are correct, the majority of the autistic children are likely to have been boys. It may be that 'playing with dolls' is not a particularly appropriate task for these children. It can also be argued that using real people in this task is more appropriate than using dolls particularly with autistic children who are highly unlikely to play with dolls usually.

How useful are the results?

There are several points to make here. Firstly there are **four autistic children** from the sample of twenty who did answer the belief question correctly. This does slightly weaken the conclusion drawn by the researchers that autistic children lack a theory of mind. It might be more accurate to say that **most autistic children** lack a theory of mind. A more detailed study of the children who did demonstrate theory of mind might be useful. If they have this ability and are still demonstrating the triad of impairments outlined earlier, then the lack of theory of mind cannot be the 'crucial ingredient' for autism that has been implied.

On a more positive note, there is no doubt that a lack of theory of mind is common in autistic individuals. Now that this has been established, it may be possible to develop **teaching methods** that attempt to encourage theory of mind skills in autistic individuals. For example, many autistic children are unable to recognise emotions from facial expressions in the same way that non autistic individuals do (this is another manifestation of lacking a theory of mind) but can learn to interpret expressions with practice.

Gardner and Gardner (1969) Teaching Sign Language to a Chimpanzee Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the Study?

This is a **case study**. A case study is an **in depth** study of one individual and produces **rich detailed data**. However it is **difficult to generalise** from a case study as you have studied one individual who is unlikely to be representative of the general population. For example, Washoe might have been a very clever chimpanzee who learned far more than other chimpanzees might have done, or she might have been less able than other chimpanzees. Either way, it would be wrong to draw conclusions about the language abilities of chimpanzees from the study of just one.

What type of data was collected in the study?

Data referring to the number of signs that had been learned and the number of times they were produced is **quantitative data**. The strict criteria under which this data was collected is a strength of the study as the researchers did not simply list each sign as learnt when it had been shown once. However, these criteria could be criticised as being arbitrary. Why is a sign not learnt if it only appears once a day for 13 days rather than 15 days?

Some of the other data is **qualitative**, such as the descriptions of Washoe using signs in different contexts. This is very valuable data in a study such as this as it adds significantly to our understanding of what Washoe has learnt.

What does the study tell us about the nature - nurture debate?

The study suggests that chimpanzees are capable of **learning a language**, at least to a rudimentary level. Although in her natural environment, Washoe would not spontaneously produce sign language, it is clear that she is able to adapt to new environments and learn new skills.

What does the study tell us about reinforcement?

A **positive reinforcer** is a **positive consequence** that follows a behaviour. A reinforcer can be anything at all that makes a behaviour more likely to occur again. The study demonstrates the power of **reinforcement to produce new behaviours**. Washoe would not have produced signs if she did not find this reinforcing. The study suggests that reinforcement plays a crucial role in the learning of language in chimpanzees but whether we can generalise this to humans is less clear. Reinforcement is obviously important while children are developing language but as we have already seen, it is unlikely to be enough to explain the very rapid development of language that is observed in children.

Was the study ethical?

Washoe was a wild caught chimpanzee. She would have been caught and taken from her mother while she was still very dependent on her care. This would no doubt have caused extreme distress to both chimpanzees. While in the Gardners care, she was well looked after, but was denied the company of other chimpanzees which might also be considered unethical. There are now countries which would prohibit such research as they have granted the Great Apes the same rights as protect humans.

Was the study useful?

Opinion is usually divided on this question. Direct applications are difficult to identify unless you think that the ability to communicate with animals would be **useful**. The usefulness of this study is evident when you look at how research like this has changed the way that we regard animals. If chimpanzees can use language, even at a basic level, there appear to be fewer and fewer differences between them and humans. As stated above, there are moves towards increased protection particularly from experimentation, for all of the Great Apes and this is a direct result of the achievements of chimpanzees like Washoe.

Samuel and Bryant (1984) Asking only one question in the conservation experiment. Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

This study used an **independent measures experimental design**. The experiment had a high level of **control** as all variables except the crucial one of questioning were kept constant. However, experimental designs sometimes **lack ecological validity** and this is discussed in more detail below. Independent measures means that different children were used in each of the three conditions. This is the most appropriate design to use to test this research question as if the researchers had tested the same children in all three conditions there would definitely have been order effects. However, because there are different children in all three conditions there is the possibility that there are **individual differences** between the groups of children. This is unlikely however, as large numbers of children were tested and they were allocated randomly to the three groups. The study is also a **replication** of previous research, although Samuel and Bryant added the two question and the fixed array conditions to the original design. Replications are very useful in psychological research for confirming previous findings or where slight changes in methodology are used, to investigate the effects of these changes.

Was the sample representative?

Samuel and Bryant tested 252 children which is a **large sample**. They tested children from the age of five to the age of eight which allowed them to draw conclusions about the age at which children started to be able to conserve. They all came from one area of the country (Devon) which might mean that they are not representative of children from other areas of the country. For example, if Devon used different teaching strategies to other parts of the country this might have an effect on the children's cognitive abilities. This is not really a criticism of the study and overall the sample is large enough to allow generalisations to be made.

What type of data was collected in the study?

The data collected was **quantitative**. Samuel and Bryant only give us the numbers of errors made by the children and no other information. The advantages of this type of data is that it can be subjected to statistical analysis and allows easy comparisons between groups. The weaknesses of this type of data is that it tells us nothing about how the children responded to the tasks. Did they ask any questions? Did they ever say they didn't know? Were they sure about their answers or unsure? This is the kind of information that could be illustrated by qualitative data such as quotes from the children

Was the study ethical?

The study used children under sixteen as the participants. Parental consent can only be assumed as the researchers do not give us this information. The task itself does not raise any serious ethical concerns as it unlikely to have distressed the children in any way and no deception was used.

Does the study have high or low ecological validity?

This is a more difficult question. The task itself is quite an artificial one. It is not an everyday occurrence to ask children this type of question, although the skills that are being tested are **everyday skills**. Perhaps a more ecologically valid method would have been to ask children to choose which of two beakers of juice or rows of smarties they would prefer to have. This would be more 'real' to the children as well as demonstrating clearly that they could conserve. There are some difficulties in evaluating the actual question used as the researchers do not tell us the exact wording of the question. Asking a child 'are they the same?' may be a slightly ambiguous question. There are many ways in which this question might have been asked and it is possible that children may have interpreted the question differently.

Was the study useful?

This study has important implications for **teaching children** and for psychologists who question children. The study clearly demonstrates that repeating questions when talking to children makes them think that they need to give a different answer. This has been demonstrated by several other researchers who have tested children's eye witness testimony. If children are asked the same questions repeatedly they will offer different answers to subsequent questions.

Bandura, Ross and Ross (1961) Transmission of aggression through imitation of aggressive models.

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

The method used was a **laboratory experiment with observational measures**. Laboratory experiments allow for **high levels of control** which are evident in the study, for example, the way that the model produced exactly the same behaviour for each child and the fact that each child was observed in the same room with the same toys. Laboratory experiments often **lack ecological validity** and you may feel that the situations that the children were put in were highly artificial and do not bear much relation to everyday life. Experiments such as these raise **ethical issues** and this is discussed further below. The children were rated by their nursery teacher and an experimenter for pre-existing levels of aggression and these ratings were used to match the children. This is another example of experimental control although the ratings may be slightly subjective and the teacher may have known some children better than others. The children were also observed playing with their toys and the experimenters used a system of categories to code their behaviour. There may be some error or observer bias here and there is an interesting issue here. All the children were observed using the same categories and there are results for 'imitative aggression' recorded for the children in the non-aggressive model condition and the control condition. Even if the children did display the behaviours shown by the aggressive model it cannot be correct to call these behaviours 'imitative' when the children did not see the model performing them.

Was the sample representative?

The sample consisted of 72 children from the same nursery. Although 72 children is a reasonably sized sample, when you look at how many children are in each group this number reduces significantly. There are **six children** in each of the experimental conditions and it is very hard to argue that generalisable conclusions can be drawn from six children. There were all from the same nursery and therefore represent a particular group of children who have all had similar educational / play experiences and therefore may not be representative of all children that age.

What type of data was collected in the study?

The main data discussed by the authors is **quantitative**, that is we are simply given numbers to represent the amount of aggressive / non-aggressive behaviours shown by the children. This data is useful and allows for some straightforward comparisons to be drawn between the groups. The number of children in each group would make statistical analysis difficult and the quantitative data does not really give us a clear picture of the children's actions. Two children may have exhibited the same numbers of aggressive acts but in different ways or to different extents. There is also qualitative data included such as the comments made by the children and these are very useful in illustrating the children's experiences and feelings although they are difficult to analyse statistically.

Was the study ethical?

The answer to this question has to be 'no'. There is no information given in the study about parental consent and it is highly likely that parents would not have consented to their children being subjected to this experience. Some of the children were upset and distressed by the aggressive model's behaviour, all of the children were told that some toys were not for them but for some other children and throughout the experiment the children were on their own with an experimenter that they did not know very well.

What does the study tell us about the nature - nurture debate?

The study concludes that boys showed more imitative physical aggression than girls and also suggests that children were more upset by the female aggressive model's behaviour. Although this suggests that boys may be more inclined to copy aggressive behaviours this does not really allow us to conclude anything about the nature-nurture debate. It is equally possible that the boys have been reinforced for copying adult male behaviour more than the girls have and that boys may have been reinforced for aggressive behaviour (or girls' not reinforced) previously.

What does the study tell us about reinforcement?

Although the study considers modelling (or imitation) rather than reinforcement, we can draw some tentative conclusions. There is the suggestion that imitating a same sex model is more acceptable and this is likely to be because this has been reinforced in the past.

Was the study useful?

The study has practical applications and is often used as support for those people who consider that **violence on television** will be imitated by children. However this study did not test the effect of televised models and there may be many differences between the way children understand the behaviour of 'real' people compared to characters on television.

Hodges and Tizard (1989) Social and family relationships of ex-institutional adolescents Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

This study is a **longitudinal** study. This means that the researchers have collected data from the adolescents at regular intervals since they were very young. Longitudinal studies are **valuable sources of data** in developmental psychology as they allow change to be monitored. However, they are obviously very **time consuming** as well as expensive and they suffer from **subject attrition** which is a loss of subjects between each data collection. The adolescents may not want to take part any more or may have moved and be untraceable. This can lead to quite large initial samples reducing drastically at later stages of the research. To counter these problems, some researchers prefer to use **cross sectional designs** where participants of different ages are tested all at once. This allows you to draw conclusions more quickly, but does have the problem of subject variability since you are not testing the same people in each age group.

This study also uses a variety of **self-report and questionnaire / interview methods**. These are valuable tools for collecting data and if they include open-ended questions will allow the researcher to explore issues in detail. However these methods are subjects to several problems, the most important of which is **social desirability**. Can we be sure that participants are telling us the truth when we ask them about this kind of personal information?

Was the sample representative?

The sample is quite small by this point in the research and some of the sub-groups such as those restored to their natural parents is very small indeed. It is likely that a much larger sample would be more representative. However the matching of the ex-institutional children with the comparisons was done on a number of variables and this makes comparisons between the two groups valid ones.

What type of data was collected in the study?

Both quantitative and qualitative data was collected in this study. Much of the interview material would be **qualitative**. Qualitative data has the advantage of reflecting the participants own thoughts and feelings (or at least the thoughts and feelings that they wish you to know) whereas a more quantitative 'forced choice' approach may mean people are being pushed into categories that they would not have put themselves into. However, if a study like this collected only qualitative data it would almost impossible to draw any general conclusions from the research due to the problems associated with coding the material that had been collected. Asking people to respond using predetermined categories may not be ideal but it does allow easy statistical analysis of the data which means comparisons can be made.

What does the study tell us about the nature-nurture debate?

The study strongly indicates a **negative effect** of early institutionalisation, although it also suggests that many of these effects can be overcome. These findings strongly support the 'nurture' side of the nature nurture debate as it is clear that the social relationships of all the adolescents questioned in this study were the result of the type of experiences they had had during their lives.

Was the study useful?

The study was very useful in highlighting the problems experienced by ex-institutional children. These results could be applied to suggest how children in care should be looked after. For example, the importance of an **attachment figure** would appear to be crucial for social development. If children are cared for by huge numbers of temporary carers, this is likely to have a detrimental effect on their long term development. The study also suggests that adoption is highly successful and this again could have useful implications for those who are involved in making decisions about children in care.

Freud (1909) Analysis of a phobia of a five year old boy Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

Case studies have both strengths and weaknesses. They allow for detailed examinations of individuals and often are conducted in clinical settings so that the results are applied to helping that particular individual as is the case here. However Freud also tries to use this case to support his theories of child development generally and case studies should not be used to make generalisations about larger groups of people.

There are several other **weaknesses** with the way that the data was collected in this study. Freud only met Hans once and all of his information came from Hans father. We have already seen that Hans' father was an admirer of Freud's theories and tried to put them into practice with his son. This means that he would have been biased in the way he interpreted and reported Hans' behaviour to Freud. There are also examples of leading questions in the way that Hans' father questioned Hans about his feelings. See the extract on the previous page for an example of this.

Finally, there are problems with the conclusions that Freud reaches. He claims that Hans recovered fully from his phobia when his father sat him down and reassured him that he was not going to castrate him and one can only wonder about the effects of this conversation on a small child! More importantly, is Freud right in his conclusions that Hans' phobia was the result of the Oedipus complex or might there be a more straightforward explanation? Hans had seen a horse fall down in the street and thought it was dead. This happened very soon after Hans had attended a funeral and was beginning to question his parents about death. A behaviourist explanation would be simply that Hans was frightened by the horse falling over and developed a phobia as a result of this experience. Gross cites an article by Slap (an American psychoanalyst) who argues that Hans' phobia may have another explanation. Shortly after the beginning of the phobia (after Hans had seen the horse fall down) Hans had to have his tonsils out. After this the phobia worsened and it was then that he specifically identified white horses as the ones he was afraid of. Slap suggests that the masked and gowned surgeon (all in white) may have significantly contributed to Hans' fears.

What type of data was collected in the study?

The data is all **qualitative** and has been subject to **interpretation** by first Hans' father and then Freud. The reports of Hans' conversations with his father are fascinating, especially when you remember that Hans was only four or five years old during most of this time. The leading questions would suggest that Hans' father 'put words in his mouth' on several occasions.

Was the study useful?

This is a difficult question to answer. It offers us a fascinating insight into Freud's theories and raises a number of questions about the origins of phobias. Freud's work is still highly influential despite some major revisions to his ideas. The notion of the unconscious mind affecting the physical health of a person and the need for interpretation of material by a therapist are widely accepted notions.

Milgram (1963) Behavioural Study of Obedience

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used?

The method used by Milgram was the **laboratory experiment**. The main advantage that Milgram had with this method was the amount of **control** he had over the situation. He controlled what the participants saw, heard and experienced and was able to manipulate their behaviour through what they were exposed to. This method also allowed accurate measurement of variables and the clear standardised procedures meant that replication was possible.

The disadvantages of this method include **low ecological validity** and the influence of **demand characteristics** on the participants and it could be argued that they were behaving in the way that they thought was expected of them rather than producing natural behaviour. Milgram has also been heavily criticised regarding the **ethics** of this study (see below) .

Was the sample representative?

Milgram's sample was a **self-selected sample** of 40 males obtained through **advertising**. This could be regarded as being a biased sample as they were all male American citizens. They were also **volunteers** and the majority of the population is unlikely to volunteer to take part in research and those who do may be atypical of the target population in some way. Hence there may be problems generalising from these results.

What type of data was collected?

The data collected was **quantitative** in that it involved measuring participants' obedience level, numerically, in terms of how far up the voltage scale they were prepared to go. This type of data has the advantage of being easy to compare and statistically analyse. However, Milgram included no qualitative descriptions of **why** the participants obeyed or how they felt during the experiment although there are a few brief descriptions of participants' behaviour during the experiment.

Does this study have high or low Ecological Validity?

As with all laboratory experiments there are problems with Milgram's study regarding its ecological validity. It involved an extremely unusual task carried out under **very artificial conditions** and as such, is likely to have produced very unnatural behaviour from the participants. This has implications for the extent to which we can generalise from these results to real life situations and it can be argued that the study tells us nothing about obedience in everyday life but simply shows us how obedient *these* people were, in *this* environment, performing *this* task.

Was the study ethical?

Milgram's study was probably one of the most unethical pieces of psychological research ever conducted. It can be criticised in terms of almost all the British Psychological Society Ethical Guidelines including **informed consent, deception, right to withdraw** and **protection from harm**. However, in Milgram's defence, we can argue that he did not expect the participants to obey to the extent that they did or to find the task so stressful. He also conducted a thorough debriefing and follow-up monitoring of his participants. A survey conducted one year later revealed that 84% of the participants were **glad to have taken part** in the study and psychiatric examinations of them showed that none had suffered long term harm.

What does this study tell us about Individual and Situational Explanations of behaviour?

The individual explanation for the behaviour of the participants would be that it was something about them as people that caused them to obey, but a more realistic explanation is that the **situation** they were in influenced them and caused them to behave in the way that they did. Some of the aspects of the situation that may have influenced their behaviour include the formality of the location, the behaviour of the experimenter and the fact that it was an experiment for which they had volunteered and been paid.

How useful is this research and to what extent can it be applied to everyday life?

As stated previously, the stimulus for this study was the Holocaust and this study has contributed significantly to the discussions regarding the behaviour of the Germans at this time. In particular, it provides strong evidence against the 'Germans are different' hypothesis. It also gives a valuable insight into the power of situations and of authority. The results suggest that we have a **natural tendency** to obey authority figures even when we feel that what we are being asked to do is morally wrong. However, the applications are restricted by the methodological limitations such as low ecological validity and an unrepresentative sample.

Haney, Banks and Zimbardo (1973) A study of prisoners and guards in a simulated prison Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used?

The method used by Haney, Banks and Zimbardo is not clear-cut. Zimbardo refers to it as an **experiment** with the independent variable of assignment to either prisoner or guard and the hypothesis that assignment to guard or prisoner will result in significant differences in behaviour. However, it could also be described as a **controlled observation** as it involved recording spontaneously occurring behaviour under conditions contrived by the researcher. The main advantage of either of these methods is the researcher's control over the variables. The weaknesses can include **low ecological validity** and the influence of **demand characteristics**, which are discussed further below. There can also be **ethical problems**.

Was a representative sample used?

The sample consisted of **24 volunteers** who were predominantly white, middle class, male students. This is clearly a biased sample as all the participants are the same gender, age, ethnic group and of similar educational and social backgrounds. Hence it would be difficult to generalise the results of this study to other, different groups in society.

What type of data was collected?

The data collected was both **qualitative** and **quantitative**. The qualitative data included details of transactions between groups of participants observed directly and recorded on video and audiotapes. Also collected were daily guard shift reports and post- experimental interviews. The quantitative data was gained from personality and mood questionnaires. Hence the researchers had numerical data that they could compare and analyse and also detailed, descriptive information.

Does the study have high or low Ecological validity?

There are two arguments regarding the ecological validity of this study. One is that it has high ecological validity because the prison environment in which the study takes place is very realistic and that the participants believed in the situation and displayed realistic behaviour. The other side of the argument is that the study lacks ecological validity because it is not a real prison, but an artificial environment created for an experiment and that the participants were simply responding to demand characteristics and acting out a role.

What does the study tell us about Individual and situational explanations of behaviour?

The **individual explanation** for the behaviour of the participants would be that the guards behaved in the way that they did because they were **naturally cruel and sadistic people** and that the prisoners were **naturally subservient and weak**. However the fact that they were all initially **screened** and found to be similar in terms of mental and physical health and stability argues against this explanation, as does the fact that they were **randomly allocated** to the roles of prisoner and guard. Hence a more convincing explanation is that they behaved in the way that they did because of the **situation** they were in. This would support the initial hypothesis proposed by Zimbardo that the social environment created in prisons is what has the negative and destructive effect on its inhabitants.

What does the study tell us about Reinforcement?

A further explanation for the behaviour of the participants can be described in terms of reinforcement. The escalation of aggression and abuse by the guards could be seen as being due to the **positive reinforcement** they received both from fellow guards and intrinsically in terms of how good it made them feel to have so much power. Similarly the prisoners could have learnt through negative reinforcement that if they kept their heads down and did as they were told they could **avoid** further unpleasant experiences.

How useful is this research and to what extent can it be applied to everyday life?

As the aim of this study was to examine the effect of prison on its inhabitants the obvious applications are to the **organisation and administration of real prisons**. Following this research, Zimbardo proposed changes to prisons and to guard training but his suggestions were not taken up and, in fact, prisons in the USA have been radically reformed in the last 25 years to make them **less humane!** However, testimony about the research influenced Congress to change one law so that juveniles accused of federal crimes cannot be housed before trial with adult prisoners because of the likelihood of violence against them. The study also gives a valuable insight into the power of situations and roles on behaviour.

Piliavin, Rodin and Piliavin (1969) Good Samaritanism: an underground phenomenon? Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used?

The method used by Piliavin was the **field experiment**. The main strength of this method is that it has **high ecological validity** and demand characteristics are unlikely to influence the participants. The weaknesses include the **lack of control** over the environment and the possibility of bias from extraneous variables. The lack of control also makes replication difficult. The field experiment can also have specific ethical problems, which are discussed in relation to this study below.

Was a representative sample used?

The sample consisted of the 4450 American passengers using that particular train, 45% of which were black and 55% white. This is a good-sized sample that is likely to be fairly representative of the American public. However the sample is restricted to the people who were using *that* train at *that* time.

What type of data was collected?

The data gathered was both **qualitative** and **quantitative**. The quantitative data included the number and type of passengers who helped as well as the time taken to offer assistance. The qualitative data came from the spontaneous comments made by the passengers. Both types of data are valuable in building up a full picture of what happened and why. The quantitative data allowed for comparisons and statistical analysis and the qualitative data provided some of the thoughts and feelings of the people involved including perhaps providing explanations for why they did or did not help.

Does the Study have high or low Ecological Validity?

As the study took place in a real life environment and the participants were unaware that they were part of a study, the ecological validity is very **high**. Hence, the behaviour displayed by the participants is likely to be their natural behaviour and the way in which they would behave in other situations.

Was the study ethical?

The ethical problems with this study include the fact that the participants' consent was not gained and that they were deceived regarding the staged collapse and the identities of the other researchers on the train. There is also a problem regarding invasion of privacy in that the participants were unaware that they were being observed and that data was being recorded about their behaviour.

What does the study tell us about Reliability?

Reliability refers to the **consistency** of a measure and one aspect of reliability relevant to studies involving observations is how consistent different observers are when recording information on the same event i.e. inter-rater reliability. The reliability of this study was increased by the fact that there were two independent researchers observing and recording data. Hence they were able to measure **inter-rater reliability**.

How useful is this research and to what extent can it be applied to everyday life?

This study is a good example of where psychological research can be used to explain behaviour in everyday life. The fact that it is a field experiment with high ecological validity using a large sample makes the results highly applicable to other situations and environments, and useful in terms of explaining and predicting how people are likely to behave when faced with an emergency situation in everyday life. The study also provides strong evidence against the Diffusion of Responsibility theory, which was supported by previous bystander research.

Tajfel (1970) Experiments in Intergroup Discrimination

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used?

Tajfel used the **laboratory experiment** method of research which had the advantage of enabling him to **control** the environment in terms of what the participants experienced including the information and instructions given to them and ensure that no other factors could influence their behaviour. Manipulation of the environment in this way enables **cause and effect relationships** to be indicated and the use of standardised procedures makes replications possible.

The main disadvantage of this method is the **lack of ecological validity** which is discussed further below.

Was a representative sample used?

The sample could be considered to be a **biased one** as the participants were all male, of similar age and from one particular school in one particular area of the country. It is difficult to see how these results could be generalised to other groups such as females, other age groups or people from other geographical areas. It could also be argued that these teenage boys were simply displaying the competitiveness typical of boys of this age and not discrimination.

What type of data was collected?

The data generated was **quantitative** in that it involved calculating the number of participants who selected the different options in the matrices. This data allowed for comparisons to be made and statistical analysis to be carried out. However the study provided no qualitative data describing how the boys behaved or why they made the choices they did.

Does this study have high or low Ecological validity?

The study can be criticised in terms of its **low ecological validity**. It involved an **unusual task** performed in an **artificial environment** and it could be argued that it produced unnatural behaviour on the part of the participants. There is also a strong possibility that they were influenced by the **demand characteristics** of the situation and acted in the way that they thought was expected of them.

What does the study tell us about Reductionism?

Discrimination is clearly a very complicated human behaviour with many factors influencing it, and a criticism of Tajfel's study is that he has reduced this complex psychological phenomenon down to a very simple level, focusing just on minimal groups and performance of a simple experimental task.

What does the study tell us about Ethnocentrism?

Ethnocentrism refers to the tendency to interpret human behaviour from the viewpoint of ones own ethnic, social or other group and the belief that ones own group is superior to all others. It is clearly an issue that is relevant to this study as Tajfel's aim was to examine group behaviour and in particular, behaviour towards out-groups. The participants were clearly seen to favour their own group perhaps because of the ethnocentric belief that their group was superior to all others.

How useful is this research and to what extent can it be applied to everyday life?

This study deals with a destructive and anti-social, but very common feature of society in examining the causes of prejudice and discrimination. As such it can be seen as a very useful piece of research that could be used to improve everyday life. However, the applications are restricted by the methodological limitations such as low ecological validity, an unrepresentative sample and the reductionist principles adopted.

Schachter and Singer (1962) Social, cognitive and physiological determinants of emotional state Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

The method used in this study was a **laboratory experiment**. This method allows for a **high level of control** over variables as well as **random allocation to conditions**. However, laboratory studies often **lack ecological validity**.

Is the sample representative?

The sample used in this study were **184 undergraduate students** who received course credits for participating in the research. Firstly, undergraduate students may not be **representative** of the general population and secondly, the offer of **course credits** may increase the likelihood that the students behaved in ways which they thought they experimenters wanted them to behave (demand characteristics).

What type of data was collected?

The data collected in this study was **quantitative data** although there were two different methods of data collection. The first measure of emotion was taken from **self reports** where participants had to rate their feelings on a 0-4 scale. Self reports of this type have several weaknesses. Firstly different participants may **interpret the scale in different ways** and this may mean that participants with very similar levels of emotion actually rate themselves very differently. Secondly, participants may **not tell the truth** when they asked questions like these. This may be due to **demand characteristics or to social desirability bias**.

The second measure of emotion was **behavioural observation**, Observers watched the participants through one way mirrors and recorded their behaviours using a number of pre-determined categories. Observation can be a useful method of data collection especially if people do not realise that they are being observed (although this may raise ethical issues) but the categories need to be **carefully defined** and the **observers trained** in order to ensure that observations are reliable.

It could be argued that collecting the two different measures of emotion in this study is a strength as they provide some kind of comparison. If the self reports and the behavioural observations agree, then this increases the **reliability** of the conclusions.

Was the study ethical?

This study broke several ethical guidelines. The major ethical concern raised by this study is the issue of **deception**. Participants were deceived in several ways during this study: they were told that the injection was a vitamin when it was epinephrine, they were misled (to various degrees) about the effects that the injection would have on them, they were led to believe that the 'stooge' was another participants and they were not aware that their behaviour was being observed through one -way mirrors. The study could also have caused **distress** to the participants, particularly those in the anger conditions, who may have been very **offended** by the questions that they were asked.

Does the study offer an individual or a situational explanation for behaviour?

The authors conclude that emotional experiences are a result of cognitive appraisal of the situation we find ourselves in. In this sense they are offering a **situational explanation** for behaviour (emotions) although as we will see below, this conclusion may have been slightly overstated.

Was the study useful?

It could be argued that the conclusions drawn from the study were useful in resolving the debate between the James Lange and the Cannon Bard theories of emotion. Schachter and Singer concluded that there are not specific patterns of arousal associated with each emotion but a **general pattern of arousal** which is combined with **cognitive appraisal** to determine the emotion. Therefore physiological arousal is necessary, but not sufficient. However, there are **problems** with this conclusion. Firstly it has proved **impossible to replicate** this study (and today it would be considered too unethical to replicate). Secondly, the conclusions may be overstated. It is possible that participants are **simply more suggestible** when under the influence of adrenaline and therefore imitate the stooge more. This would go some way to explaining the fact that the self reports and the observer ratings are in fact, different (see above). Finally, the initial analysis of the data had revealed only one difference (between the observer ratings of emotion between Epi-Ignorant and Placebo groups in the Anger condition). It was only when the researchers excluded participants who they decided had worked out the aim of the study that they found differences in the observer ratings for the euphoria conditions. All of these criticisms would suggest that Schachter and Singer may have **overstated their conclusions**.

Dement and Kleitman (1957) The relation of eye movement during sleep to dream activity

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

This study was a **laboratory experiment**. Laboratory studies have **high levels of control** over variables but may **lack ecological validity** (see below). The study used the **electroencephelograph (EEG)** to measure brain activity and this allows **precise measurements** to be taken.

Was the sample representative?

The sample was very small indeed. Only **nine** participants were studied in total and only five of these were studied intensively. This is a very small number of participants to **generalise** from. You could argue that physiological **processes are likely to be the same in all people** but this may not be the case. It is possible that sleep patterns and in particular, relationships between eye movements and dreaming, vary from person to person and the conclusions drawn from Dement and Kleitman's research would be strengthened if the same relationships were established in a **larger sample of people**.

What type of data was collected?

The data collected in this study by mostly **quantitative**. In relation to the first two research questions, the data is simply numbers of participants who could or could not recall dreams or could or could not accurately estimate the length of their dreams. Quantitative data is relatively easy to collect and can be **analysed statistically**. However, it **reduces** complex qualitative phenomena to numbers and this often results in a **lack of detail** in the data. The data collected in relation to the third question was more **qualitative** as here participants were asked to describe the contents of their dreams. This gives us **richer and more interesting** data but such data is difficult to use for comparisons and is hard to analyse statistically.

Was the study ecologically valid?

The research studied participants who went to sleep in a laboratory with electrodes stuck to their head. It is unlikely that this bears much relation to sleep in a normal environment! It is possible that being in such an artificial condition meant that their sleep was disturbed and if this was the case, the researchers would **not** have been studying **normal sleep patterns**. The participants were also woken up several times during the night and asked about their dreams. Again, this is unlikely to happen normally and may have had an affect on the way the participants slept. However, research conducted outside of the controlled conditions of the laboratory would have been unable to measure brain activity and eye movements in the ways.

Was the study useful?

The conclusions that were reached by Dement and Kleitman have been **replicated** by many other researchers. However, there is one methodological issue that should be considered. In relation to the first research question, Dement and Kleitman conclude that dreaming takes place in REM rather than non-REM sleep. What they have actually demonstrated is that dreams are recalled more often from REM rather than non-REM sleep and it may be that dreaming does occur in non-REM sleep and it much harder to recall them.

When the research was first conducted, very little was known about the relationship between eye movements and dreaming and so Dement and Kleitman's research really did add **new information** to what was known about sleep. It is difficult, nearly fifty years later, to understand what a major breakthrough this study represented. The use of EEG to record brain activity whilst sleeping was also relatively new and it was not until research like this Core Study, that it became clear that dreams could be studied in an objective way. Dement and Kleitman's research **generated very many other studies into sleep and dreaming** and there have been many useful findings.

Sperry (1968) Hemisphere deconnection and unity in consciousness Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used by Sperry?

Sperry conducted a **quasi-experimental** design. One weakness of this type of design is that it does **not** give the researcher **full control** over the independent variable (whether someone had their corpus callosum severed). However, quasi-experimental designs allow researchers to investigate variables that are **not able to be investigated in strict laboratory experiments**. Sperry's research can be seen as similar to a **collection of detailed case studies** conducted with **highly controlled and objective laboratory equipment and procedures**. The major strength of Sperry's work is definitely **the techniques** that he developed which allowed the functions of the two hemispheres to be studied in ways which had previously been impossible.

Was the sample representative?

Sperry had **eleven participants** who had already undergone surgery to sever their corpus callosum. Eleven would be considered quite a small number in other types of studies but it would not have been possible to find large numbers of people who had had this operation. There are two questions that need to be considered. Firstly, are these eleven people **representative of everyone who has had this operation**, and secondly, **can the results of this study be used to tell us anything about the functioning of the 'normal' brain?** The results suggest that all the eleven participants experienced **very similar effects** to each other and so it would be safe to conclude that anyone who had this operation would experience these effects. However, as Sperry did not control the independent variable in this study, he was not able to test these participants prior to the operation. It is possible that their brain functioning may have been **atypical** (different from the norm) before the operation and this would make drawing conclusions about the functions of the hemispheres in non separated brains more difficult.

What type of data was collected in this study?

There are examples of both **quantitative and qualitative** data in this study. The majority of the data is quantitative as Sperry simply records whether something could be identified or not. This is the **important data** in this study as it is from this that Sperry is able to draw his conclusions about the different functioning of the two hemispheres. However, the results are **illustrated** with some **revealing qualitative** examples of the experiences of the split-brain patients and this **adds significantly to our attempt to understand the experiences of these people**. For example, the description of patients giggling at nude photographs presented to their right hemisphere whilst denying that they have seen anything is far more revealing than simply reporting how many responded and under what conditions.

Was the research ecologically valid? / Can the results be applied to real life?

In a sense this research has **very little ecological validity** as the techniques that Sperry developed artificially separate the visual and tactile information received by the individual. It is difficult to think of a situation where this would happen in real life, and as we saw in the Core Study, split brain patients have a number of **simple strategies** for coping in the real world that they were unable to use in the laboratory conditions. On the other hand, ecologically valid research should be studying **real problems** and Sperry's research is looking at a naturally occurring variable (split brain as a result of the operation) and trying to understand exactly what effects this operation has on the individuals concerned.

Was the research useful?

The research was **extremely useful**. Sperry's work revealed facts about the **lateralisation of functions** between the two hemispheres that had only been suggested by previous studies. There are many ways in which this knowledge could be applied to helping people cope with the effects of brain damage.

Raine, Buchsbaum and LaCasse (1997) Brain abnormalities in murderers indicated by positron emission tomography
Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

This study was a **highly controlled laboratory (quasi) experiment** using very sophisticated equipment. However, there are some **problems with P.E.T. scans** and these include the fact that the generation of the images is a very complex process which has scope for errors to be made. P.E.T. scans identify 'hotspots of activity' but these could be excitatory nerves (the 'on' switch) or inhibitory nerves (the 'off' switch). As the brain becomes practised at a task, the amount of activity declines and this means that activity seen on a scan may simply represent the brain processing something new. All of the above means that we should treat the data produced from a P.E.T. scan with caution. Finally, P.E.T. scans require an injection into the bloodstream of a radioactive marker. The dose is tiny but for safety, no-one is allowed to have more than one scanning session (twelve scans) per year.

Was the sample representative?

This is a difficult question to answer! Were the sample representative of murderers? Were the sample representative of murderers claiming Not Guilty by Reason of Insanity? Would it even be possible to establish this? The matching by age and sex (and schizophrenia) is a **strength** of the study but further matching by mental illness could have been used. It might also have been interesting to compare the brain activities of murderers claiming Not Guilty by Reason of Insanity with murderers not claiming this although this would raise very many ethical and practical issues. Finally, 41 participants in the experimental and control groups is a reasonable size sample for this kind of research.

What type of data was collected in the study?

The data collected in the study was **quantitative** as the colours in the PET scans represent different levels of metabolic activity in the brain. PET scans will produce **reliable data** (in the sense that the PET scanner will measure metabolic activity **consistently**) but as we have already seen there is **scope for error** in the interpretation of the data. The data represented metabolic activity during a Continuous Performance Task and while it is certainly interesting to discover that murderers and non-murderers show different levels of activity, it is difficult to see how activity during a CPT can tell us anything about the reasons for someone committing a murder.

Was the study reductionist?

The study '**reduces**' the behaviour of murder to a **measure of metabolic activity** in the brain. As Raine would no doubt agree, murder is a complex act which is likely to be affected by a number of **social and individual variables** and cannot be explained simply in terms of metabolic activity. In support of the study, it could be argued that Raine has identified a **very important difference** between the brains of murderers and non-murderers which will no doubt generate more research.

What does the study tell us about the nature - nurture debate?

It is difficult to draw conclusions about the causes of murder from this study. It is not possible to conclude that the differences in brain activity did cause the individuals to commit murder and the results definitely **do not** suggest that **violence is determined by biology alone**. Neither do the results tell us what caused the differences in brain activity in the first place and this could be genetic, biological or environmental.

What are the ethical implications of the research?

The research itself does not appear to break any ethical guidelines but there are **ethical implications** raised by this research. If the results were taken by some, to suggest that people with a particular pattern of brain activity were likely to commit murder, then all kinds of 'science fiction' scenarios could be imagined. This would overlap with the issue of **social control**.

Was the research useful?

The ability to collect new types of data from different groups of people is always useful. Raine's work has established a potentially significant difference in the brains of murderers who are claiming that they are Not Guilty by Reason of Insanity. The comments made above clearly suggest that the results should be treated with caution but there is no doubt that this kind of research will continue.

Gould (1982) A nation of morons

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in this study?

Review articles are slightly different from other methods that are used in psychological research. The author of a review does not collect data but **comments** on research conducted by another researcher. Gould's review offers a fresh perspective on what may be regarded as a shameful episode in psychology's history. He highlights many problems both with the **collection and interpretation** of the data and with the benefit of hindsight, is able to show how the **conclusions** drawn from Yerkes' work had **major social and political consequences**. The review also brings to our attention some of the strengths and weaknesses of the use of **psychometric tests** more generally. There is no doubt that many psychometric tests are valuable tools and allow for the testing and comparisons of large numbers of people relatively quickly and cheaply. If the tests are reliable and valid (see below) then the tests will be more objective than an individual opinion as to the personality characteristics or intelligence of a job applicant for example. However, there are also many problems associated with the way tests are administered and scored and results should often be treated with caution.

Was the sample representative?

This is a difficult question to answer. The sample of men tested by Yerkes totalled over 1.75 million army recruits. This has to be the largest sample of participants ever used in psychological research! However, the fact that there are so many questions over the way the tests were administered and the content of the tests, makes it very unwise to generalise Yerkes' conclusions on intelligence to any other groups.

What type of data was collected in this study?

The data was **quantitative**. All Yerkes collected was the IQ scores as measured by his tests. Gould, however, does offer us some descriptions of the testing conditions which could be described as **qualitative data** and these descriptions add considerably to our understanding of the problems in Yerkes' research

What does the study tell us about the reliability and validity of intelligence tests?

Gould argues convincingly that the tests used by Yerkes were **neither reliable nor valid**. They were unreliable due to the many different ways in which people were tested and the fact that so many were given the wrong tests. They were not valid because they did not test native intellectual ability but cultural knowledge which could not possibly have been acquired by people who had only recently arrived in the country.

What does the study tell us about ethnocentrism?

The review highlights the dangers of ethnocentrism. Yerkes and his colleagues were apparently '**blind**' to the flaws in their research and their reasoning. These flaws look very obvious to us nearly a century later but as we will see in the introduction to the Core Study by Hraba and Grant, American society was a very different place 100 years ago. The **unchallenged ethnocentric attitudes** of the researchers contribute to '**scientific racism**' where unsubstantiated arguments are presented as scientific and have damaging effects to particular groups of people. This is discussed further below.

What does the study tell us about the nature - nurture debate?

People supporting a **eugenicist argument** would no doubt quote research like Yerkes' as support for their points of views. However Gould's review shows us clearly that whatever Yerkes was measuring was the result of **nurture** rather than nature. People scored highly on the tests if they spoke fluent English and were familiar with American culture. If not, they scored badly. This cannot be taken to support any argument that proposes the intellectual superiority of white people over black people as the tests were biased in favour of the White American recruits.

What does the study tell us about social control?

The study highlights the dangers of **scientific racism**. Massive **political consequences** resulted from the work conducted by Yerkes and huge numbers of lives were no doubt lost as a result of the Immigration Act. This illustrates the power of research and the dangers of taking such research at face value.

Was the study useful?

Gould's review is very useful. It is important that people continue to question and reassess research that has taken place in the past. It also shows the importance of considering original research rather than simply accepting the brief summaries of results of studies that are often found in textbooks. It shows us the importance of being **critical** and always looking for details of how research was conducted.

Hraba and Grant (1970) Black is beautiful; A re-examination of racial preference and identification

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in the study?

The method is **quasi-experimental** as the researchers did not have control over the independent variable (the race of the children). Quasi-experimental designs are suitable for research where the independent variable varies naturally and cannot be manipulated by the experimenter. The research was also a **replication** of a previous study. Replications are useful in **confirming previous findings or in highlighting differences** between the original study and the replication. In this case, the replication suggests the massive social changes that had taken place between 1939 and 1969 are responsible for the differences in the way the children responded to the questions. It is now more than 30 years since Hraba and Grant conducted this study and it is possible that a **further replication** would reveal more differences in racial preferences.

Was the sample representative?

Hraba and Grant tested a **large sample** of children (160) drawn from five different schools in Lincoln, Nebraska. 89 of the children were black and these children represented 60% of the black children in the age group studied in these schools. This would suggest that Hraba and Grant had a representative sample of black children aged between 4 and 8 in public school in Lincoln Nebraska. Whether it can be assumed that these children are representative of black children in other parts of America is not so clear. It could be argued that differences between the area where the children lived in the original study and Lincoln, may be large enough to account for the differences found. Children living in areas where there are far bigger black populations may also have very different attitudes to their own race. Lastly, the research has focussed on simple distinctions between black and white and this does not in any way represent the vast diversity of race.

What type of data was collected in this study?

The data was quantitative. We simply know how many children chose the black doll and the white doll in response to each question. It would have been interesting to know a little more in relation to these questions. For example, why did children choose a particular doll, what made them think that the doll looked bad, or nice, or what exactly about the doll did they think looked like them? Did they have a strong preference for one doll or were they just picking one because they had been asked to?

Can the study be applied to real life?

The study attempted to measure racial preference and racial awareness by using dolls. It could be argued that doll choice does not necessarily tell us anything about the child's attitudes to their own race. Hraba and Grant point out that there was no relationship between doll preference and choice of best friends for either the black or the white children. We are not given details about the dolls except that for were used, two black and two white and 'identical in all other respects'. How were they dressed? Were they meant to be male or female? The collection of qualitative data (see above) might have revealed more about each child's reason for picking a particular doll. There may also have been some demand characteristics in the procedure of the study. Children were asked eight questions and they may have made the assumption that they needed to choose a different doll in response to each of the questions.

What does the study tell us about ethnocentrism?

The study shows us that both white and black children were ethnocentric in the sense that they were more likely to choose a doll of their own race as the 'nice doll' and a doll of the other race as the 'bad doll'. In fact, the study perhaps reveals more about the **white children's ethnocentric attitudes** than the black children's racial preferences. The study also strongly suggests that these attitudes are social in nature and will change in response to social changes.

Was the research useful?

The study highlights a number of useful points. Firstly, it demonstrates the fact that the results of psychological research have a '**lifespan**'. Findings from the 1930s may not have anything useful to tell us about behaviours today. This suggests that we should be aware of the **social and political context**, as well as the culture, in which a piece of research is conducted. Secondly, the research suggests that racial preferences are learnt and this might lead to suggestions about how negative attitudes could be 'unlearned'.

Rosenhan (1973) On being sane in insane places

Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in this study?

Field experiments have the major advantage of being conducted in a real environment and this gives the research **high ecological validity**. However it is not possible to have as many controls in place as would be possible in a laboratory experiment. **Participant observation** allows the collection of highly detailed data without the problem of demand characteristics. As the hospitals did not know of the existence of the pseudopatients, there is no possibility that the staff could have changed their behaviour because they knew they were being observed. However this does raise serious **ethical issues** (see below) and there is also the possibility that the presence of the pseudopatient would change the environment in which they are observing.

Was the sample representative?

Strictly speaking, the sample is the twelve hospitals that were studied. Rosenhan ensured that this included a range of old and new institutions as well as those with different sources of funding. The results revealed **little differences** between the hospitals it this suggests that it is probably reasonable to generalise from this sample and suggest that the same results would be found in other hospitals.

What type of data was collected in this study?

There is a huge variety of data reported in this study, ranging from the **quantitative data** detailing how many days each pseudopatient spent in the hospital and how many times pseudopatients were ignored by staff through to **qualitative descriptions** of the experiences of the pseudopatients. One of the strengths of this study could be seen as the **wealth of data** that is reported and there is no doubt that the conclusions reached by Rosenhan are well illustrated by the qualitative data that he has included.

Was the study ethical?

Strictly speaking, no. The staff were **deceived** as they did not know that they were being observed and you need to consider how they might have felt when they discovered the research had taken place. Was the study justified? This is more difficult as there is certainly no other way that the study could have been conducted and you need to consider whether the results justified the deception. This is discussed later under the heading of usefulness.

What does the study tell us about individual / situational explanations of behaviour?

The study suggests that once the patients were labelled, the label stuck. Everything they did or said was interpreted as typical of a schizophrenic (or manic depressive) patient. This means that the situation that the pseudopatients were in had a powerful impact on the way that they were judged. The hospital staff were not able to perceive the pseudopatients in isolation from their label and the fact that they were in a psychiatric hospital and this raises serious doubts about the **reliability and validity** of psychiatric diagnosis.

What does the study tell us about reinforcement and social control?

The implications from the study are that patients in psychiatric hospitals are 'conditioned' to behave in certain ways by the environments that they find themselves in. Their behaviour is shaped by the environment (nurses assume that signs of boredom are signs of anxiety for example) and if the environment does not allow them to display 'normal' behaviour it will be difficult for them to be seen as normal. Labelling is a powerful form of social control. Once a label has been applied to an individual, everything they do or say will be interpreted in the light of this label. Rosenhan describes pseudopatients going to flush their medication down the toilet and finding pills already there. This would suggest that so long as the patients were not causing anyone any trouble, very little checks were made.

Was the study useful?

The study was certainly useful in highlighting the ways in which hospital staff interact with patients. There are many suggestions for improved hospital care / staff training that could be made after reading this study. However, it is possible to question some of Rosenhan's conclusions. If you went to the doctor falsely complaining of severe pains in the region of your appendix and the doctor admitted you to hospital, you could hardly blame the doctor for making a faulty diagnosis. Isn't it better for psychiatrists to err on the side of caution and admit someone who is not really mentally ill than to send away someone who might be genuinely suffering? This does not fully excuse the length of time that some pseudopatients spent in hospital acting perfectly normally, but it does go some way to supporting the actions of those making the initial diagnosis.

Thigpen and Cleckley (1954) A case of multiple personality Thinking like a Psychologist - Evaluating the Core Study

What are the strengths and weaknesses of the method used in this study?

This study is a **Case Study**. This means that it is a **detailed analysis** of one individual. The depth of analysis and the richness of the data is a strength of case studies and there is no doubt that this study represents an enormous amount of research time.

The study also uses a range of **psychometric and projective tests**. Psychometric tests are simple to use and allow broad comparisons to be made and it is certainly interesting that Eve White and Eve Black show some distinct differences in their results. Projective tests are more **subjective** as they rely on the therapists interpretation of the responses given by the patient and are therefore much more prone to bias and the effect of expectations. With both of these measures, it should be remembered that personality is not stable and any individual might achieve a different score if tested on several different occasions when in different moods.

Is the sample representative?

A case study is never representative of anyone other than the individual being studied. We can only really apply the results of this case study to helping the individual concerned. However, clinical case studies are very useful for other therapists particularly when cases are highly unusual.

What type of data was collected in the study?

Both **qualitative** and **quantitative data** was collected and they complement each other. If we simply had the scores from the intelligence and memory tests, the case that the authors are making for a diagnosis of multiple personality disorder would not be as convincing. If the article simply contained the descriptions of the differences (fascinating as they are) people might still be sceptical. Taken together, it can be argued that the authors present a **strong case** for the existence of more than personality in Eve White.

What does the study tell us about the nature-nurture debate?

The study suggests that **environmental factors**, particularly **traumatic events**, can profoundly affect our mental health. Whilst there may be convincing evidence that there is a **genetic predisposition** to certain mental health disorders, there is no doubt that they are triggered by events in people's lives. In material published later than this study, information is given about a number of traumatic events that Eve had experienced. She had seen a man drown, and another man cut into pieces by a machine at a lumber mill, and had witnessed her mother very badly injured. There is little doubt that such events would have an effect on the mental health of a child even if you do not accept the diagnosis of MPD.

Are there any ethical implications raised by this study?

Does a therapist have the **right to judge** which personality is more deserving of survival than others? Some of the comments made about Eve Black are extremely judgemental and probably both sexist and dated. Therapists have a great deal of power over their patients and a responsibility to use this power wisely.

Was this study useful?

Case Studies are primarily useful in a **clinical** sense and this has been described above. However, this case study was the first fully documented case of multiple personality and it has no doubt led to the disorder being more easily recognised than at the time of this case. However, as was briefly mentioned in the introduction, this is a highly controversial diagnosis and there is some concern that patients are being labelled incorrectly as suffering from this disorder.

