FACTBOOK

AO1 and AO3

RESEARCH METHODS

Research Methods

Experimental method

Laboratory experiments are carried out under strictly controlled conditions in an artificial environment, where the independent variable is manipulated and the dependent variable is measured.

In lab experiments participants are required to attend the location of the lab, which is usually within a university or formal setting.

Lab experiments control variables and use standardised procedures which make them easy to replicate. Therefore they have high levels of reliability.

Asch (1951) conducted a lab experiment to investigate his research into majority influence using the line comparison test.

Lab experiments can be ethical in that they usually ask participants to give informed consent.

Lab experiments have an increased risk of demand characteristics, where participants are more likely to discover the aims of the study and alter their behaviour accordingly.

Lab experiments lack ecological validity as the controlled, artificial conditions often do not replicate real life.

Field experiments are carried out under controlled conditions in a natural setting, where the independent variable is manipulated and the dependent variable is measured.

Field experiments are often carried out in natural settings where the researcher can still manipulate the variables. This could be in a public library, cafe or a shopping centre.

Field experiments have high ecological validity as the real life settings are representative of how people would behave in those environments.

Field experiments are difficult to replicate as there is an increased risk of extraneous variables, which makes them lack reliability.

Natural experiments are carried out in a naturalistic environment, where the researcher has little to no control over the variables. In a natural experiment the independent variable cannot be manipulated.

In natural experiments there is less chance of experimental bias as the independent variable is not controlled or manipulated by the researcher. This increases the validity of the study.

If participants in natural experiments do not know that they are part of a study, this raises ethical issues.

Natural experiments have a greater risk up confounding variables influencing the findings, as a result of the lack of control.

Quasi-experiments take a naturally occurring independent variable in order to investigate its effect on the dependent variable. These are variables that already exist either within the participant or within the situation.

An example of a quasi-experiment would be to investigate the reading comprehension in children with autism. In this case, autism is a naturally occurring independent variable.

Quasi-experiments can only be used where a naturally occurring difference between individuals is identified. This can make it harder to control the outcome of the study.

Natural and quasi-experiments have high ecological validity as they are carried out using natural conditions. However, this makes them harder to replicate and lowers their reliability.

Observational techniques

Naturalistic observations are carried out in a naturalistic setting in order to investigate real life behaviour.

Naturalistic observations have high ecological validity as they record real life behaviour and get a true picture of how people would behave in that setting.

Naturalistic observations are extremely difficult to replicate which means that they lack reliability.

Schaffer & Emerson (1964) conducted a naturalistic observation in order to study infant attachment whilst children were in their own homes.

Naturalistic observations can observe behaviour overtly or covertly.

An overt observation is conducted whilst the participants are aware it is taking place.

Zimbardo (1973) used overt observation to record the behaviour of the prisoners and the guards in his simulated prison study.

An overt observation has an increased risk of demand characteristics, as people may not display their true real life behaviour, instead act in a way they think the researchers want.

Overt observations are more ethical as participants will have given full informed consent.

A covert observation is conducted whilst the participants are unaware it is taking place.

A covert observation is likely to give a valid picture of real life behaviour as people will not display demand characteristics.

A covert observation is more unethical as participants are not aware that they are being studied.

Milgram (1963) used covert observation to record the levels of obedience shown by participants in response to a person in authority.

At the end of a covert observation, researchers must fully debrief participants unless they are observed in an environment in which they would expect others to watch their behaviour. For example, if you are in a public place, you would expect others to see you.

Controlled observations are carried out in a structured environment, which is usually in an artificial setting.

Controlled observations allow for more objective data collection, as many extraneous variables can be controlled.

Controlled observations have less ecological validity because the participants are not in their natural environment.

Ainsworth (1970) used a controlled observation to test her method of the 'Strange Situation' in attachment research.

A participant observation is one in which the researcher takes an active role in the observation process. They are often observing from within the group, inside the study.

Participant observations allow the researchers to collect data first hand; however this may be subjective as they are observing from within the group.

A non-participant observation is one in which the researchers take note of behaviour from outside of the study. They will observe participants from a distance.

Non-participant observations are when the researchers remain separate from those that they are studying, in order to record behaviour more objectively.

Observational research is useful as it gives us an insight into how people behave in more realistic settings.

One limitation of the observational method is observer bias. This is when the researchers interpretation of what they are watching may be affected by their own stereotypes and expectations.

In order to reduce observer bias more than one researcher must be used to record the data. On agreement of the category data, inter-observer reliability will increase.

Self-report techniques

Questionnaires collect written responses either on paper or in an online format.

Questionnaires allow opinion based data to be gathered using a set of questions.

Questionnaires use a series of questions to investigate a topic area. These questions can be open-ended or fixed choice.

Open-ended questions allow participants to give answers in any way they choose. They can be short or long and allow for more description.

Open-ended questions collect qualitative data which is more valid as it gives an insight into what people are thinking.

Fixed choice or closed questions limit the responses that participants can give by providing set options or tick boxes.

Scale questions such as the Likert scale also limit the responses participants can give by providing a set option. For example, a 5-point scale could range from strongly agree to strongly disagree.

Fixed choice and scale questions collect quantitative data which is reliable as scores can be compared more easily and data can be analysed statistically.

Questionnaires can collect large amounts of data as they are relatively easy to administer and can be replicated. This makes them more reliable as a research method.

Although questionnaires tend to be standardised, the conditions in which they are administered may be subject to extraneous variables. This decreases their reliability.

Questionnaires are at risk of social desirability bias which affects the validity of the findings. Participants may respond in ways which are socially desirable and not a true reflection of their opinions.

In questionnaires participants may not understand the question, give untruthful answers or simply guess the correct response. All of these factors reduce the validity of the findings.

Interviews are a method of self-report in which participants answer direct questions from an interviewer.

Interviews are usually conducted face to face and can gather qualitative and quantitative data.

Interviews can be structured, semi-structured and unstructured.

A structured interview is one in which the researchers prepare questions prior to the interview itself.

A structured interview will use a pre-set of standardised questions in order to gather data.

An example of a structured interview would be one that is used by the police when questioning an offender.

Structured interviews are reliable as they are easier to standardise and allow comparisons to be made between participants by asking them the same questions.

Structured interviews are usually carried out in a controlled environment.

An unstructured interview is one in which there is more variation during its procedure. The questions are less formal and more conversational. This allows the researcher to deviate from the questions depending on the responses they receive.

An example of an unstructured interview would be one that is used by a chat show host when questioning a guest.

Unstructured interviews are best used when the research is exploratory and requires depth.

Some interviews can be semi-structured which means they use a pre-set of standardised questions but are more flexible in how they are delivered. The interviewer may allow the participant to deviate from the script.

Interviews are a good way of gathering in depth information about individuals' opinions.

Interviews are valid way of collecting data especially if open-ended questions are used.

Interviews can be a very time consuming research method, as they are mostly conducted on a one to one basis. This can also limit the sample size of the research.

Once data is collected from an interview it needs to be transcribed which is a very time consuming task. This can also be a subjective process.

Interviews are often subject to interpretation bias. This is when the researchers use their own subjective opinion to analyse the data.

Interviews can also risk social desirability bias as participants may respond according to how they think the questions should be answered. This decreases the validity.

Correlations

Correlations analyse the relationship between two or more co-variables.

Correlations look at the association between variables.

A positive correlation is when one variable increases so does the other.

A negative correlation is when one variable increases, the other decreases.

Zero correlation is when there is no relationship between the co-variables.

One difference between correlations and experiments is that correlations do not look for a difference, instead look for a relationship between variables.

In experimental research an independent variable is manipulated in order to measure the effect on the dependent variable, in a correlation there is no such manipulation.

Hypotheses written for correlations are known as alternative hypotheses, not experimental hypotheses as there is no IV or DV in a correlation.

A correlational hypothesis must be operationalised and clearly state the expected relationship between the co-variables.

Correlations can be positive or negative, and strong or weak.

A correlation coefficient of +1 indicates a strong positive correlation.

A correlation coefficient of -1 indicates a strong negative correlation.

A correlation coefficient of 0 indicates no correlation.

A perfect positive correlation is noted as +1 whereas a perfect negative correlation is -1.

Correlations are plotted on a scattergram with variable 1 on the x-axis and variable 2 on the y-axis. Where the two variables meet a cross is drawn. A line of best fit will indicate the type of correlation.

Correlations are when one participant is measured on two scales (testing two different variables).

Correlations show a relationship between two or more variables but it is not possible to establish cause and effect.

Correlations are useful in that they provide a quantifiable measure of how two variables are related.

Correlations can demonstrate useful patterns between variables.

Correlations are quick and cost effective to carry out.

Correlations only tell us how variables are related but not why they are related.

Correlations cannot demonstrate cause and effect between variables.

In some correlations an intervening variable (third variable problem) can be causing the relationship between the other two co-variables.

A third variable problem is a type of confounding in which a third variable leads to a mistaken causal relationship between two others.

Content analysis

A content analysis is a method used to investigate behaviour that is usually measured using qualitative methods.

Content analysis is a method of changing qualitative data into quantitative data (e.g., interview transcripts or video recordings), so that it can be statistically analysed or used descriptively.

Content analysis can be seen as a type of observational research in which people are studied indirectly via the communications that they have produced.

The forms of communication that may be subject to content analysis are wide-ranging and may include spoken interaction, written text or media (books, magazines for TV programmes).

The aim of a content analysis is to summarise and describe the form of communication in a systematic way so conclusions can be made.

A content analysis will look for trends or patterns in the data and convert this to numerical data.

In a content analysis the researcher must decide how to categorise the analysed material.

Coding is the initial stage in a content analysis.

Some data within a content analysis may need to be categorised into meaningful units. This may involve counting the number of times a word or phrase appears in the text.

A content analysis is reliable as the final set of data is in numerical form and can be analysed statistically.

Content analysis is useful when analysing data that cannot otherwise be analysed using other methods.

Content analyses are useful when studying mental health such as schizophrenia. Patients can be interviewed and their responses can be transcribed in order to look for patterns.

Content analysis is it reasonably ethical research method as most participants give permission for their data to be used.

Content analysis is a flexible technique in that it may produce both qualitative and quantitative data.

Content analysis can often be subjective as individual researchers may interpret the data differently.

There is a danger in content analysis in that the researcher interpreting the data may attribute opinions and motivations that were not initially intended. This increases the subjectivity of the findings.

Content analysis unless objective than other research methods.

Content analyses can be very time consuming as researchers often have a lot of data to work with.

As individuals tend to be studied indirectly as part of a content analysis, their communications are usually analysed outside of the context within which it occurred.

Unless inter-rater reliability is assured, content analyses can lack reliability.

Thematic analysis is a form of content analysis but the outcome is qualitative.

The main process in a thematic analysis involves the identification of key themes. These will be identified by the researcher at the start of the analysis.

In a thematic analysis patterns and themes often emerge once the data has been coded.

Case Studies

A case study is carried out over an extended period of time studying the behaviour of one individual or small group.

Case studies follow individuals over an extended period of time.

Case studies often involve analysis of unusual or rare behaviour.

The case study method will involve researchers constructing a case history of the individual, which may be from observations questionnaires or interviews.

Case studies are a longitudinal research method, which give a true insight into an individual's behaviour.

Case studies can collect data using a variety of research methods focused on the behaviour of one individual.

Case studies allow detailed information to be gathered from individuals in order to explore the unique nature of their behaviour.

Case studies are a valid method of data collection as most focus on collecting qualitative data from an individual.

Case studies are useful as they offer rich, detailed insights which may explain unique behaviour or atypical patterns.

The case study of H.M was significant in memory research as it helped to demonstrate the existence of separate stores in STM and LTM.

Freud (1909) conducted a case study of Little Hans, the five year old with a phobia.

Case studies may generate hypotheses for future research or may lead to the revision of explanations or theories.

Generalisation from the findings of case studies is limited as sample sizes are too restrictive.

Case studies are subjective and may have bias from the interpretation of the researcher. This lowers their validity.

Case studies are often reliant on self-report data and accounts from family and friends, these may be inaccurate lowering the validity of the findings.

Scientific Processes

Aims

An aim is a general purpose of intent, which is written at the start of an investigation.

Aims in psychological research are usually developed from theories. They are general statements that describes the purpose of an investigation.

An aim is not a prediction of the outcome of the results. It simply states the general purpose of the investigation.

Hypotheses

A hypothesis is a clear and precise prediction about the outcome of the results.

A hypothesis is a testable statement written at the outset of any study.

A hypothesis is stated at the start of an investigation following the aim.

A hypothesis can look for a difference or a relationship between the variables.

Hypotheses can be directional and non-directional.

A directional hypothesis is also known as one-tailed hypothesis.

A directional hypothesis states the direction of the difference or relationship.

A directional hypothesis states the outcome of the results by predicting which one of the conditions or variables will have an effect.

An example of a directional hypothesis (one-tailed) – It is predicted that females will recall significantly more words on the memory test out of 20 compared to males.

A non-directional hypothesis is also known as a two-tailed hypothesis.

A non-directional hypothesis does not state the direction of the difference or relationship.

A non-directional hypothesis is unsure of the outcome of the results, so predicts there will be a difference/ relationship but the direction is uncertain.

An example of a non-directional hypothesis (two-tailed) – It is predicted that there will be a significant difference between males and females and the number of words they can recall on the memory test out of 20.

Hypotheses in Psychology must be operationalised, this means that the independent and dependant variable must be stated explicitly.

Operationalisation is clearly defining the variables in terms of how they can be measured.

Research conducted using experimental methods will write an experimental hypothesis. This can be one-tailed or two-tailed.

Research conducted using correlational methods will write an alternative hypothesis.

Hypotheses are predictions written at the start of an investigation and aim to look for a significant difference or relationship.

A research hypothesis predicts that the outcome of the results will be due to the experimenter manipulation.

A null hypothesis predicts that the outcome of the results were not due to experimental manipulation but instead due to chance.

A null hypothesis predicts that there will be no significant difference between the conditions/ variables and the outcome is due to chance.

An example of a null hypothesis – It is predicted that there will be no significant difference between males and females and the number of words they can recall on the memory test out of 20. All results will be due to chance.

At the end of an investigation researchers must accept or reject the research hypothesis. They will also accept or reject the null hypothesis.

Sampling

A sample in Psychology is a small group collected from the target population.

A sample is a group of people who take part in research who are drawn from a population presumed to be representative.

A target population is a group of people who are the focus of the researcher's interest, from which a smaller sample is drawn.

Target populations can be large groups of individuals who the researcher is interested in studying. For example students attending colleges in the northwest of England.

A sample is a small group taken from the target population. For example, 50 students attending one sixth form college in the northwest of England.

Sampling techniques are methods used to gather a sample from the target population.

In ideal circumstances a sample will be representative of the target population it is taken from.

Having a representative sample reduces the bias and allows generalisations from the findings, back to the target population.

In practise it is often difficult to represent full target populations within a given sample.

Bias in the context of sampling is where certain groups are over or under-represented within the sample itself. This limits the extent to which generalisations can be made back to the target population.

Generalisation is the extent to which findings and conclusions from a particular investigation can be applied back to the wider population.

Generalisation is possible if the sample of participants is representative of the target population.

There are several sampling techniques used in Psychology when gathering participants to form a sample.

A random sample is one in which all members of the target population have an equal chance of being selected into.

In order to conduct a random sample, researchers must first obtain a list of all the members of the target population. These names are placed in a hat or random number generator from which the sample is then selected.

A random sample is potentially unbiased however it can be difficult and time-consuming to conduct.

A random sample may still result in an unrepresentative sample group, as individuals with similar characteristics could be selected at random. For example, a random sample of 20 students could all be male.

A random sample is likely to produce a more representative sample compared to other techniques.

A random sample may also result in participant attrition, if those selected do not wish to participate in the research.

A systematic sample is when every nth member of the target population is selected. For example a researcher may select every 5th person from the list.

In a systematic sample a sampling frame is often produced using the people from the target population.

In a systematic sample researchers will select a sampling frame such as every 10^{th} person taken from a list of the target population.

A systematic sample is free from bias as the researcher has no influence over who is selected.

Systematic sampling is an objective technique compared to other sampling methods.

Systematic sampling can be a time consuming process and participants may still refuse to take part.

A stratified sample comprises of individuals from certain subgroups in equal proportion to that of the target population.

A stratified sample reflects the proportions of people in certain subgroups (strata) within the target population.

A stratified sample is a mini replica of the target population in terms of the exact demographics of that group.

If a target population comprises of 60% female and 40% male, a stratified sample would reflect these exact proportions.

To carry out a stratified sample the researcher must identify the different subgroups that make up the target population. The exact proportions are calculated and represented in the final sample.

A stratified sample is the most representative sampling technique as every subgroup from the target population is represented proportionately.

Stratified samples allow for generalisations as they represent the target population.

Stratified samples will have issues with individual differences based on those individuals and their personalities from within the subgroups.

Stratified sampling is a time consuming task, especially when gathering the data from the individual subgroups in the target population.

Opportunity sampling is where participants are selected by chance, as they are available at the time.

Opportunity samples are often gathered in naturalistic settings where researchers simply ask anyone available to participate.

Opportunity sampling takes a chance to ask those who are around at the time of the study. This could be in a public place such as a cafe or library.

Opportunity sampling is cost effective.

Opportunity sampling is quick and easy to conduct. Researchers do not need to gain a list of the members of the target population before they begin.

Opportunity sampling is subject to researcher bias and often those selected do not represent the wider population.

In opportunity sampling there is bias as the sample is unrepresentative of the target population, as they are mostly drawn from one specific area. This limits the generalisation.

In opportunity sampling the researcher has complete control over the selection of participants, and in doing so may exclude some individuals on the basis that they do not look approachable. This results in a biased sample.

Volunteer sampling involves participants selecting themselves to be part of the sample.

Volunteer sampling is also known as self-selected sampling.

To gain a volunteer sample a researcher may place an advert in a newspaper or online. Participants then self-select to take part in the investigation.

Volunteer sampling is easy to conduct and requires minimal input from the researcher.

Volunteer sampling is likely to be more ethical as participants are giving their consent to participate.

Volunteer bias can be a problem when asking for a volunteer sample. Individuals with particular personality traits may put themselves forward.

Volunteer samples are at greater risk of social desirability bias or demand characteristics.

Volunteer samples are not representative of the wider population, as only certain types of individuals put themselves forward and volunteer.

Samples may limit the generalisation of the findings as they may not represent the target population.

Pilot Studies

A pilot study is a small scale version of an investigation that takes place before the real investigation is conducted.

A pilot study is like a pre-run of an investigation in order to iron out any problems or issues with the methodology.

A pilot study will use a small sample size to test the procedure and check that the investigation runs smoothly.

Pilot studies can be used with all research methods and are a good way to ensure internal validity.

The aim of a pilot study is to check that procedures, materials and measuring scales all work effectively prior to an investigation.

A pilot study allows researchers to make changes or modifications before their investigation begins.

Pilot studies are a part of the design process and allows researchers to make changes before they begin gathering data.

Pilot studies are likely to result in more reliable findings.

Experimental Designs

Experimental designs are ways in which participants are allocated to conditions.

Experimental designs are the different ways in which participants can be organised in relation to the experimental conditions.

A repeated measures design is an experimental technique in which all participants experience both conditions of the experiment.

In a repeated measures design participants take part in all the experimental conditions. They will repeat a process/ procedure.

In a repeated measures design participants experience condition A followed by condition B. Their scores in each condition can be compared. This means each participant is compared to themselves.

A strength of using a repeated measures design is that participant variables can be controlled which results in higher validity.

Repeated measures designs require fewer participants, therefore are more cost effective and less time consuming.

In a repeated measures design there can be issues with the participants taking part in two or more conditions. Participants may perform better second time round or may perform worse second time round. This is known as order effects.

Order effects are a disadvantage of a repeated measures design. These effects can be due to practise (participants getting better) or fatigue (participants getting worse).

Counterbalancing is a technique which aims to control order effects in repeated measures designs.

Counterbalancing splits participants into two groups, half do condition A followed by condition B, whereas the other half do condition B followed by condition A.

Order effects can arise because repeating two tasks could create boredom or fatigue. This might result in a deterioration in performance on the second task. Alternatively participants performance may improve through the effects of practise.

Demand characteristics are risk factors in repeated measures designs, as participants may alter their behaviour in condition two.

An independent groups design is when two separate groups of participants experience different conditions of the experiment.

The independent variable usually indicates whether an independent groups design is being used. For example, one group could be females and another group males. These are two separate groups of participants.

In an independent groups design separate groups of participants usually take part in the same procedure, measuring the same dependent variable.

In independent groups designs participants are allocated to different groups where each group represents one experimental condition.

Performance of two separate groups of participants can be compared easily in an independent groups design.

One problem with an independent groups design is that participants are all different which increases the risk of individual differences.

Independent group designs have a higher risk of participant variables, this reduces the validity of the findings.

If random allocation is used in an independent groups design the results are believed to be more valid.

Random allocation is an attempt to control for participant variables in an independent groups design. This aims to ensure each participant has the same chance of being in either of the two conditions.

There are no order effects in an independent groups design and there is less chance of demand characteristics affecting the results.

Independent measures designs require more participants and can be more time consuming.

Matched pairs designs take one participant in condition A and match them on key characteristics with one participant in condition B.

In a matched pairs design participants are paired together based on key variables such as age, gender and social class.

Participants in a matched pairs design all take part in the same procedure in order to compare their performance on the task.

In some cases pre-tests are conducted in an attempt to control for participant variables, when a matched pairs design is used.

In a matched pairs design participants only take part in a single condition, so order effects and demand characteristics or not a problem.

Matched pairs designs use different participants in each condition, which results in a higher risk of participant variables.

Matched pairs designs can be very time consuming and difficult to conduct, especially given all participants must be matched with a partner of similar characteristics.

Matched pairs designs are less economical than other designs particularly if a pre-test is required.

Observational Design

When designing an observation behavioural categories are created in order to code behaviour that is seen.

Behavioural categories are when target behaviours are broken up into smaller components that are observable and measurable.

Behavioural categories give researchers a structured record of what they observe. These categories are sometimes referred to as a behavioural checklist.

When behavioural categories are identified they become operationalised, which means they can be measured reliably.

It is important that behavioural categories are exclusive and do not overlap with each other, this is to ensure reliability.

In order for behavioural categories to be seen as reliable the target behaviours must be precisely defined and made observable and measurable.

Prior to an observation, a researcher should ensure that all possible target behaviours are stated in their behavioural checklist.

Some researchers may want to record behaviour by simply writing down everything they observe. This means lots of detailed information can be gathered which gives a true, valid insight into what is being seen.

In some cases there may be too much information for the researcher to record in a single observation. It is often recommended that more than one observer takes part.

In observations, data is easier to analyse as it is mostly recorded quantitatively in a tally chart of behavioural categories. This makes it a reliable technique.

It is recommended that researchers do not conduct observational studies alone. Single researchers may miss important details or may only notice events based on their own perspective. This introduces bias in the research.

To make data recording more objective, observations should be carried out by at least two researchers. Their data can be checked for consistency and if in agreement gives the findings high inter-observer reliability.

To ensure inter-observer reliability researchers should familiarise themselves with the behavioural categories and observe the behaviour at the same time. They will then be required to compare their findings and on agreement reliably state the outcomes.

Event sampling is an observational technique where all the behaviours observed are recorded by the researcher. This could be during a full one-hour lunch break.

Event sampling is when a target behaviour or event is recorded every time it occurs.

Event sampling gathers a large amount of data which often gives a good valid picture of what is being observed.

Event sampling is useful when the target behaviour or the event happens quite infrequently and could be missed if time sampling was used.

One problem with event sampling is that data may be missed especially if there is a lot going on during the observation.

Event sampling can be a subjective technique as researchers will decide what information is recorded, sometimes its only data that can be feasibly recorded.

Time sampling is an observational technique where all the behaviours are recorded at set time intervals. These intervals are pre-agreed at the start of the observation. This could be behaviour which is recorded every 5 minutes.

Time sampling is when a target behaviour is recorded in a fixed time frame.

Time sampling is an effective way of gathering observational data when there are less researchers.

Time sampling is effective in reducing the number of observations that have to be made.

One problem with time sampling is that essential/ important behaviours could be missed during the set intervals.

Questionnaire Construction

Questionnaire construction will often depend on the type of data required to gather, either qualitative or quantitative.

Qualitative data will require the construction of open-ended questions in order to gather true opinions from individuals.

Open-ended questions are those for which there is no fixed answer choice and respondents can answer in any way they wish.

Quantitative data will require the construction of fixed choice questions or rating scale questions.

Fixed choice questions have set responses from which people must select an answer.

Rating scale questions use numerical or qualitative scores to gather their opinion based data.

A rating scale asks respondents to identify a value that represents the strength of their feeling about a particular topic. The options are usually set out numerically.

Likert scales ask respondents to indicate their agreement with a statement.

In a Likert scale people are required to state a number that represents their opinion, this is often from strongly agree (1) to strongly disagree (5).

Likert scales are often based on a 5-point or 7-point scale.

Fixed choice questions and scale questions are easier to analyse as they often result in quantitative data. This makes them a more reliable form of data collection.

Open-ended questions take longer to analyse as they often result in qualitative data. However, this makes them a more valid form of data collection.

When designing question is it is essential the questions are clear and do not confuse people.

Technical terms and jargon should be avoided when designing questionnaires. The best questions are simple and easily understood.

Emotive and socially sensitive questions should be avoided in questionnaires, especially if this is likely to have an effect on the respondents.

When designing a questionnaire, researchers should avoid questions which result in 'maybe' or 'sometimes' options. This reduces the validity of the results.

Design of Interviews

When designing an interview, researchers will create an interview schedule. This is a list of questions they intend to cover.

An interview schedule should be standardised to reduce the effect of interviewer bias.

Data collection in interviews can be gathered manually in a notebook or can be recorded and analysed afterwards.

Interviews usually involve an interviewer and a single participant; however, group interviews may also be appropriate.

When conducting interviews researchers should ensure they are in a quiet room with little to no distraction.

When designing interviews researchers may benefit from asking neutral questions at the start in order to establish rapport with the individual.

Data collection via interview can be both qualitative and quantitative.

Interviews can collect data from open-ended questions or fixed choice questions.

Interviews aim to collect first-hand information from people in order to gather a true insight into their thoughts and opinions. However, the researcher must make it known that all of their responses will be kept confidential.

Variables

Variables are anything that can vary or change within an investigation. Variables are generally used in experiments to determine if changes in one thing result in changes to another.

Variables are manipulated and measured in psychological research.

There are two fundamental variables in experimental research; the independent variable and the dependent variable.

An independent variable (IV) is an aspect of an experimental situation which is manipulated by the researcher.

An independent variable is one which the researcher changes or manipulates.

An independent variable is the one thing that is different between the groups. This could be gender - males and females.

A dependent variable (DV) is a variable which is measured by the researcher and any effects seen acting on the DV should be caused by the IV.

The dependent variable is the one which is measured by the researchers.

The dependent variable is usually the performance of the participants in the experimental conditions. This could be the number of correctly recalled words in a memory test.

Both the IV and DV in experimental research should be operationalised in order to make them testable.

Operationalisation is clearly defining the variables in terms of how they can be manipulated or measured.

In experimental research there should be strict control of variables to ensure reliability of the findings.

Extraneous variables are anything other than the independent variable which may have an effect on the dependent variable. These variables should be controlled by the researchers at the outset of a study.

Extraneous variables will affect the internal validity of the results if they are not controlled for.

Extraneous variables can be anything extra in the environment which may have an effect on the results.

Extraneous variables are unwanted and should be identified at the start of an investigation, where a researcher takes steps to minimise their influence.

Extraneous variables can be participant related such as age, but most are environmental such as noise/ distraction or poor lighting in the lab.

Extraneous variables are often seen as nuisance variables that make it harder to detect a true result.

Confounding variables are a kind of extraneous variable which can change with the independent variable.

Confounding variables often crop up at the end of an investigation when analysing the results.

A confounding variable is an extraneous variable that varies systematically with the IV so we cannot be sure of the true source of the change to the DV.

Confounding variables make it difficult to tell if the change seen in the DV is due to the IV or something else.

An example of a confounding variable is individual memory recall. In an experiment measuring the number of correctly recalled words, the data could be skewed if it is later identified that one participant had a photographic memory.

It is important to identify all possible confounding variables and consider their impact on the research in order to ensure the internal validity of the study.

There are several ways in which confounding variables can be controlled. For example using random allocation as part of the design or using a repeated measures design to ensure participants take part in all conditions of the IV.

Confounding variables can be minimised by controlling all other variables as tightly as possible. Also to include participants with similar characteristics or demographic information.

Control

Random allocation is an attempt to control for participant variables usually in independent groups designs. This reduces bias and ensures that each participant has the same chance of being allocated into each condition.

Random allocation is when participants are randomly allocated to the experimental conditions in the study.

Random allocation can be done manually by placing names in a hat or more systematically using a name/ number generator.

Random allocation of participants to experimental and control conditions is an extremely important process in research. Random allocation greatly decreases systematic error, so individual differences or ability are far less likely to affect the results.

Counterbalancing is an attempt to control for the effects of order in a repeated measures design.

Counterbalancing splits the participants into two groups and varies their exposure to the experimental conditions. Half of the group do condition A followed by condition B, whereas the other half of the group do condition B followed by condition A.

Counterbalancing is a useful way of minimising order effects such as practise and fatigue.

Randomisation uses chance to control for bias when designing materials and deciding the order of experimental conditions.

Randomisation is a technique used by researchers to minimise the effect of extraneous or confounding variables.

Randomisation can be used to randomly allocate participants into conditions.

Randomisation is used in the presentation of trials in an experiment to avoid any systematic errors that might occur as a result of the order in which the trials take place.

Randomisation is likely to result in more internally valid results.

Within an investigation, all participants should be subject to the same environment, information and experience.

Standardisation is the extent to which all procedures are kept the same for all participants.

Standardisation is necessary to ensure replication gives all participants the same experience.

Standardisation is controlled by the researcher and can be given to participants through standardised instructions.

Having standardisation in a study ensures that any changes in data can be attributed to the IV. In addition, it is far more likely that results will be successfully replicated on subsequent occasions.

Failure to include standardisation in research will result in less reliability.

Demand characteristics are the result of any cue from the researcher or the environment that is interpreted by the participants as revealing the purpose of the investigation. This may lead to a change in their behaviour.

Demand characteristics are when participants change or alter their behaviour in order to please the researcher.

Demand characteristics are when clues in the experimental situation may help a participant to guess the experiment is intentions or the aims of the study.

Presence of demand characteristics in a study suggest that there is a high risk that participants will change their natural behaviour in line with their interpretation of the aims of a study.

Participants may also look for clues in a situation which tell them how they should behave. This will result in unnatural responses.

Participants may act in a way that they think is expected and over perform to please the researcher. This is known as the 'please-u effect'.

Participants may act in a way which deliberately sabotages the results of a study, by underperforming. This is known as the 'screw-u effect'.

Demand characteristics are more likely to be seen in lab experiments rather than natural experiments.

Demand characteristics reduce the internal validity of a study.

Demand characteristics are more likely to be seen when participants have volunteered to take part in a study (self-selected sample).

A repeated measures study design is more likely to present the problem of demand characteristics, as participants will be take part in all conditions of the experiment.

Investigator effects are any conscious or unconscious effect on the research outcome (DV) from the investigators behaviour.

Investigator effects unwanted influences from the investigator on the research outcome.

Investigator effects occur when a researcher unintentionally, or unconsciously influences the outcome of any research they are conducting.

Investigator effects can be evident through nonverbal communication. The researcher can communicate their feelings about what they are observing without realising that they have done so.

Investigator effects can be evident through the physical characteristics of the researcher. The appearance of the researcher and such physical characteristics as their gender will influence the behavioural response of the participant.

Investigator effects mean that the behaviour all participants is a product of the situation because of the researcher and therefore may not be reliable or valid.

Investigator effects can produce biased outcomes which are subjective. A researcher can affect the results reported from a piece of research by interpreting the data in a biased way.

Investigator effects reduce the validity of the findings within an investigation.

Ethics

The British Psychological Society (BPS) has its own set of ethical guidelines outlined in its code of conduct.

Researchers have a professional duty to observe the ethical guidelines outlined in the code of conduct when performing any type of psychological research.

The role of the British Psychological Society's code of ethics is to provide psychologists with a set of guidelines to adhere to when dealing with participants in research.

Ethical guidelines are implemented by an Ethics Committee who sit in research institutions and use a cost-benefit approach to determine whether particular research proposals are ethically acceptable.

Most psychological studies will submit a research proposal to the Ethics Committee outlining the purpose of their research and the methods they intend to use.

Ethical committees review proposals to assess if the potential benefits of the research are justifiable in light of the possible risk of physical or psychological harm.

Ethical committees may request researchers make changes to the study's design or procedure or in extreme cases, deny approval of the study altogether.

There are separate ethical guidelines for human participants in research and animals in research. However, they are both essentially set out to protect the participants from harm during the research process.

Ethical researchers prioritise respect for the rights and dignity of participants in their research and also consider legitimate interests of stakeholders such as funders, institutions and sponsors.

Ethical guidelines are necessary to clarify the conditions under which psychological research can take place.

The ethical code of conduct is built around four major principles; respect, competence, responsibility and integrity.

Informed consent, right to withdraw and confidentiality are all guidelines in the principle of respect.

With the principle of respect, psychologists should respect individual, cultural and social differences between people. This can include factors such as age, gender and ethnicity.

Psychologists should respect the knowledge, insight, experience and expertise of participants and avoid practices that are unfair or prejudiced.

With the principle of competence, psychologists should value their continuing development and maintain high standards throughout their professional work.

Competence suggests that the researchers must be qualified and able to conduct the research they set out to test.

Protection of participants from harm and debriefing are guidelines in the principle of responsibility.

It is the responsibility of the psychologists during research to ensure that participants are protected from physical and mental harm, and that they leave the study in the same state they started.

Deception is a key factor in the principle of integrity. Psychologists should value honesty, accuracy, clarity, and fairness in their interactions with all those involved in the research process.

Psychologists must conduct a cost-benefit analysis prior to the start of a research study. They must consider the benefits to society which could be gained by testing new theories and the costs to the participants within the research.

The cost-benefit analysis may produce conflict about treating the participants ethically. For example, if informed consent is obtained and no deception is used in the study the participants are being treated ethically however, there is an increased risk of demand characteristics.

In the Zimbardo (1973) study where participants were asked to play the role of prisoner or guard, the benefits outweighed the cost. However, Zimbardo did not expect the experiment to be terminated after just 6 days.

Ethical issues arise when a conflict exists between the rights of participants in research and the goals of the study to produce valid and worthwhile data.

Ethical issues can have implications for the safety and well-being of the participants in the research.

In order to maintain ethical standards psychologist must follow the guidelines throughout the whole research process.

At the start of psychological research participants must give their full informed consent as agreement to take part in the research.

Informed consent involves making participants aware of the aims of the research, the procedures and how their data will be used. They should also be given information about their rights.

Participants should be told the nature, purpose, and anticipated consequences of any research participation, and ideally the researcher should gain informed consent at the beginning of research.

If participants are under the age of 16, informed consent needs to be gained from parents or guardians.

Participants should be issued with a consent letter or form detailing all relevant information that might affect their decision to participate. Assuming the participant agrees, this is then signed.

Participants should always be asked to sign a form giving their agreement and full informed consent.

Most lab-based research gains informed consent more easily as participants go into the lab to take part in the research.

Researchers should make it clear to participants they have the right to withdraw from the investigation at any time irrespective of payment or other credits.

In Milgram's (1963) study on obedience the experimenter used four verbal prods to encourage participants to continue, before they were allowed to withdraw. This is unethical practise.

If a participant decides to withdrawal from the research, they have the right to demand their own data and recordings to be destroyed. This can even be at the end on completion of all tasks or activities.

Participants can be given the right to withdraw in the briefing statement outlined at the start of the research. This may be before they give their full informed consent.

Participants have the right to control any information about themselves. This is their right to privacy. If this is invaded then confidentiality should be protected.

Confidentiality refers to the right under the Data Protection Act to have any personal data protected.

All results or information gathered relating to specific individuals must be kept confidential. Participant names or details should not be released or published. Participants should be made aware where any breech of confidentially may occur.

In most psychological studies researchers do not record any personal details and so participants remain anonymous. In many studies participants are referred to using numbers or initials.

In case study methods psychologist often use initials when describing the individuals involved. This is to protect their anonymity. For example, in the case of H.M.

Participants should be told in the briefing and the debriefing that their data will be protected throughout the research process and will not be shared with anyone else.

Participants privacy should be respected and in the case of observational research only recorded where they would expect to be observed in a 'public place'.

The right to privacy extends to the area where the study took place such as institutions or geographical locations and these must not be named.

Invasion of privacy should be avoided in research methods such as questionnaires or interviews. Psychologist should refrain from asking personal questions that may be of a sensitive nature.

Researchers must not cause any physical or psychological harm to participants. They should leave a study in the same state that they entered.

Participants in research should not be placed in any more risk than they would in their daily lives. They should be protected from physical and psychological harm at all times.

Psychological harm includes any stress or anxiety that participants are placed under as a result of the procedures in the study. All participants should leave the research in the same condition in which they entered.

Many classic studies in Psychology broke ethical guidelines. The BPS only introduced a code of conduct in 1985.

Psychological harm can include being made to feel embarrassed, inadequate or being placed under undue stress or pressure.

Most psychological studies avoid physical and psychological harm and will offer a follow up service at the end of the research.

In some psychological studies support or services will be offered following completion of the research. This could be a counselling service.

Milgram (1963) debriefed all his participants straight after his electric shock experiment and disclosed the true nature of the experiment. Participants were assured that their behaviour was common and Milgram also followed the sample up a year later and found that there were no signs of any long-term psychological harm.

Deception means deliberately misleading or withholding information from participants at any stage of the investigation.

Participants who have not received adequate information when they agreed to take part cannot said to have given full informed consent. This is a form of deception.

Intentional deception such as lying to participants, misleading them about the aims or other aspects involved must be avoided as much as possible unless deception is necessary in exceptional circumstances to preserve the integrity of research.

In some circumstances deception is necessary to reduce demand characteristics.

There are some occasions when deception can be justified as long as participants receive a full debriefing at the end of the study.

A way to overcome breaking ethical guidelines after a piece of unethical research is to debrief participants on the true extent of the study.

Debriefings are done at the end of the study and it is the researchers responsibility to provide participants with any necessary information needed to complete their understanding of the outcomes of the study.

Debriefing is important to check that participants have not suffered any psychological or physical harm during the process.

If participants have been deceived in any way or full informed consent was not gained at the start, the researchers should fully explain the true purpose of the study before the participant leaves.

Peer Review

Peer review is a process carried out by colleagues or experts in the field of research.

Peer review is the assessment of scientific work by others who are specialists in the same field to ensure that any research intended for publication is of high quality.

Before a piece of research is published it must be subject to the process of peer review/

Peer review is a process to ensure high quality research which accurately represents the field of study.

The main aims of peer review are to allocate research funding, to validate the quality and relevance of research and to suggest amendments or improvements.

Research funding is decided by an independent peer evaluation. Some organisations will have a vested interest in promoting research in that area.

In peer review all elements of research are assessed for quality and accuracy, this includes looking at their hypotheses, methodology and statistical analysis.

In peer review, the reviewers may suggest minor revisions to the work in order to help improve the final published report. In some extreme cases they may deem the work inappropriate and it gets withdrawn from publication.

Peer review contributes to the publication process. All academic research is subject to peer review.

Peer review takes place before a study is published to ensure that the research is valuable and accurately presented.

Peer review is a process of checking the soon to be published research for any errors or miscommunications.

If psychological research was published without the process of peer review, there would be unreliable findings which would damage the integrity of that field of research.

Research often has practical applications for society and if this is not peer reviewed there could be negative consequences.

Peer review promotes and maintains high standards in research. This can have many applications for society, including how funding is allocated.

In Clinical Psychology it is important that any research involving treatments or therapies is peer reviewed. This ensures that reliable and effective treatments are the only ones in the public domain.

Peer review helps to prevent scientific fraud as all work is scrutinised by experts in that field of study.

Peer review helps promote the scientific process and contributes to new knowledge in the field.

Peer review acts as a filter to ensure that only high quality research is published, especially in reputable journals, by determining the validity, significance and originality of the study.

It is important the peer review process is carried out objectively and those involved in the process are kept anonymous. This prevents any conflicts of interest.

Peer review is intended to improve the quality of manuscripts that are deemed suitable for publication.

The benefits of peer review help with the accuracy of the research and ensure the validity of the findings and conclusions.

Peer reviewers have the responsibility to identify strengths and weaknesses in the piece of research. They can provide constructive comments to help the author resolve issues in their work. However, a reviewer should respect the intellectual independence of the author.

Peer reviewers have a responsibility to evaluate the science according to appropriateness of the methodology, completeness and accuracy of the data, and interpretation of the results. However, this can be a subjective process.

A limitation of a peer review is that only statistically significant findings are published which means some key research may be overlooked.

Peer review is usually conducted anonymously; however a minority of reviewers may use this as a way of criticising rival researchers, particularly if they are in competition for research funding.

Publication bias is also a factor in peer review. Many journals prefer to publish significant, positive findings. This means that some research may get overlooked.

The peer review process may suppress opposition to mainstream theories, preferring to publish articles and areas of study that are already in the public domain. This can slow down the rate of change in a particular area of research.

Implications of Psychological Research for the Economy

The implications of psychological research for the economy are concerned with how the knowledge and understanding gained from psychological research may contribute towards our economic prosperity.

The economy is the monetary state in which country holds itself in terms of the production and consumption of goods and services.

The implications of psychological research for the economy can have benefits for society. For example, if more effective treatments can be developed for psychological health problems, this

means that people will be able to return to work and this reduces the burden on the employers, NHS and taxpayer.

There are several ways psychological research can positively impact on the economy, it can contribute to people's well-being, make them more productive, and boost the economy by reducing absence from work.

Schaffer & Emerson (1964) observed that babies form multiple attachments from 10 months onwards and that the father is also a significant attachment figure. This means that both parents can take it in turns to look after the baby, saving money on nursery fees. The mother can also return to work and contribute to society.

Psychological research has shown that both parents are equally capable of providing emotional support necessary for healthy psychological development with their children. This understanding has promoted more flexible working arrangements within the family.

Psychological research can benefit the economy and reduce costs to society, allowing money to be invested into positive interventions.

Absence from work as a result of illness and mental health is extremely costly for the economy. Research into treatments and therapies can offer patients ways to remain in the workplace.

The economic benefit of psychological research into mental health disorders such as anxiety and depression is considerable.

Psychological research can offer suggestions for how governments and institutions spend public money.

Loftus & Palmer (1974) found that leading questions distort people's memories of events and that witnesses are generally unreliable. Therefore, benefits the economy as miscarriages of justice are expensive for society as they involve a trial, the cost of keeping an offender in prison, any appeals process, a re-trial and possible financial compensation for the wrongly accused.

An example of how psychological research has benefited society is in the area of memory research, where funding research into memory loss and dementia has helped reduce costs on the health care system.

There can be economic implications from psychological research for individuals in society. For example, research into government cuts shows that more vulnerable people are at risk of deteriorating mental health. This will put a more costly strain on the health care system.

Reliability

Reliability is the extent to which something is consistent over time. If a research method can be replicated and produce similar results it is believed to have high reliability.

Reliability refers to how consistent a measuring device is, this can include psychological tests or behavioural categories which assess behaviour in an observation.

Reliability is a measure of consistency; if the same outcome is produced time after time, it is described as reliable.

Standardised procedures contribute to high reliability as they are often strictly controlled and can be replicated easily.

Standardised procedures ensure that participants receive the same experience which often result in more reliable findings.

Experimental settings which are strictly controlled, like those in a lab experiment will be highly reliable.

Research methods which have high levels of experimental control particularly over the variables, will have good reliability.

Test-retest reliability is a method of assessing the reliability of a questionnaire or psychological test by assessing the same person on two or more occasions. If the answers are consistent, they are believed to be reliable.

Test-retest reliability involves administering the same test or questionnaire to the same person on different occasions and receiving similar results.

Test-retest reliability is the extent to which a measure or procedure provides the same results each time it is tested.

If a participant scores the same or similar after completing a test, it is thought to have high test-retest reliability.

Test-retest reliability checks the consistency of the results by giving the same test to the same or similar people and comparing the results to the original.

If a patient were to produce similar results after being tested on more than one occasion, they would receive a diagnosis and this would indicate high test-retest reliability.

Inter-rater reliability is when two or more researchers agree on the same results.

In an observation if two or more researchers agree on the same results it suggests there is high inter-observer reliability.

Inter-observer reliability is the extent to which there is agreement between two or more observers involved in the observation. This is measured by correlating their observational scores.

Reliability is measured using a correlational analysis. The correlation coefficient should exceed +0.80 for reliability.

Correlational scores of +0.80 or more indicate high levels of reliability.

Reliability can be improved if the researcher has high control in both the setting and over the variables and uses standardised procedures in their method.

The reliability of questionnaires can be measured using the test-retest method. Correlations above +0.80 indicate high reliability.

Many questionnaires lose reliability over time. Questionnaires that produce low test-retest reliability may need to be rewritten or altered.

Using fixed choice questions in a standardised way improves the reliability of interviews.

Interviews that are unstructured are less likely to be reliable.

By operationalising the behavioural categories, an observation can increase its reliability.

Experimental methods can increase their reliability by using standardised procedures in a controlled setting.

Typically quantitative data is more reliable than qualitative data.

Validity

Validity refers to the extent in which something measures what it proposes to measure.

Validity is the extent to which an observed effect is genuine.

Validity refers to the truth of a measure; if it is truly measuring what it is supposed to measure.

Validity refers to the realism in an environmental setting. If the environment is a reflection of real life, it will have high validity.

Validity considers the extent to which the findings can be generalised beyond the research setting.

Validity suggests the results from psychological tests, observations and experiments are authentic and true.

Validity can be both internal and external.

Internal validity refers to the extent that the measure accurately records what it is designed to record.

Internal validity refers to whether the effects observed in an experiment are due to manipulation of the independent variable and not some other factor.

Internal validity is believed to be high when the researchers have strict control over the variables they manipulate and accurately measure what they propose to measure.

If the results of a study are identified as showing clear cause and effect, in that the independent variable is the one factor manipulating the results, then there is high internal validity.

Demand characteristics are one factor that may negatively affect the internal validity of a study.

One way of improving internal validity is to use a control group to compare the findings with that of the experimental group.

If the results show that the independent variable was the only factor affecting the dependent variable in a study, it will have high internal validity. This is often easier to see when comparisons to a control group are made.

Standardised procedures also help to increase the internal validity of a study by minimising the impact of participant reactivity and investigator effects.

Single blind and double-blind procedures can be used to help improve the internal validity of a study.

A single blind procedure is where the participants are unaware of the condition they are allocated. This minimises any demand characteristics or subjectivity.

A double-blind procedure is where both the participants and the researchers are unaware of the conditions they are allocated into. This minimises demand characteristics and investigator effects.

Qualitative data is more likely to be valid. It often gives a true, detailed insight into what people think and how they behave.

Qualitative methods are usually thought to have higher ecological validity than quantitative methods. For example case studies and interviews are believed to be more valid than lab experiments.

Researcher bias and subjectivity will reduce the validity of the findings in a piece of research.

Subjectivity is when research is affected by personal judgements and interpretations. This reduces the validity of a study.

Interpretation bias is a judgement that is made by a researcher or participant in terms of what they think the research outcomes want. This reduces the validity of a study.

Triangulation can be used to enhance the validity within a study. This is the use of a number of different sources or methods through which the data is collected.

In Triangulation, by combining theories, methods or observers in a study, it can help ensure that biases arising from the use of a single method or a single observer are overcome.

Research methods involving questionnaire data can be less valid due to social desirability effect. This is the idea that participants may answer in a way which makes them appear more socially desirable.

Social desirability bias reduces the validity of the research findings.

Many questionnaires and psychological tests incorporate a lie scale within the questions in order to assess the consistency of respondent's answers and control for the effect of social desirability.

In observations, if the behavioural categories are ambiguous or overlap, the validity of those findings will be reduced.

External validity relates to factors outside of the investigation such as generalising to other settings, other populations and times in history.

External validity refers to the extent to which findings can be generalised beyond the research settings in which they were found.

Ecological validity is an example of external validity.

Ecological validity is the extent to which findings from a research study can be generalised to other settings and situations.

Ecological validity relates to the research setting. If this setting represents real life, it will have high ecological validity.

Observational research may produce findings that have higher ecological validity as they are usually conducted in more natural settings and have minimal intervention from a researcher.

Mundane realism is an example of external validity.

Mundane realism relates to the research task. If the task participants are asked to complete represents one similar to their everyday life, it will have high mundane realism.

Temporal validity is an example of external validity.

Temporal validity is the issue of whether findings from a particular study hold true over time.

Temporal validity is the extent to which findings from a research study can be generalised to other historical times.

Many classic studies in Psychology will lack temporal validity as their findings are not generalisable to today's society.

If a study has high temporal validity the findings will be generalisable to other times or eras.

Population validity is the extent to which the findings from a particular sample apply to the wider population.

If a sample is representative of the wider population, it is more likely that their findings will have population validity.

One basic form of validity is face validity. This is whether the test appears ‘on the face of it’ to measure what it intends to measure.

Face validity can be checked by simply looking at the components of the measure or test to see if it outlines/ measures what it is set out to.

Face validity is a basic form of validity in which a measure is scrutinised to determine whether it appears to measure what it's supposed to measure.

If a test has face validity, it should be obvious that it is measuring what it intends to measure. For example if an IQ test on the surface looks like it has questions measuring cognitive ability, it will have face validity.

Construct validity is the degree to which a test measures what it claims or proposes to be measuring. This is based on the theory from which it is constructed.

Construct validity is a type of validity that assesses how well a particular measurement reflects the theoretical construct it is intended to measure.

By looking at individual components within a test, it can be easy to determine whether or not it has construct validity.

Low construct validity implies the test may be measuring something else unintended or that score interpretations are questionable.

Face and construct validity are types of content validity; whether a test or measurement represents all aspects of the intended content.

Concurrent validity is the extent to which a psychological test relates to an existing similar measure.

Concurrent validity is where the findings from one measure agree with the findings from a very similar measure.

If a test produces similar results to that of an identical or similar test, it is believed to have concurrent validity.

The concurrent validity of a particular test is demonstrated when the results obtained are very close to or match another well-established test.

Correlation coefficients can indicate the concurrent validity of two or more tests. If the findings concur a correlation of +0.80 or more would indicate close agreement.

Predictive validity is the extent to which the research gains similar results to other research that has been carried out at a different time, whether the results agree and so predict future outcomes.

Predictive validity assesses how well a test predicts a criterion that will occur in the future.

In diagnosis if a patient presents with symptoms that are true of schizophrenia such as hallucinations or delusions, a psychiatrist can predict the outcome or their treatment. This means there is predictive validity.

A prediction could be made on the basis of a new intelligence test that high scorers at age 12 will be more likely to obtain university degrees several years later. If the prediction works out to be true, then the test has predictive validity.

Concurrent and predictive validity are example of criterion related validity. They assess the performance of a test based on its correlation with a known external criterion or outcome.

Criterion validity is the extent to which a measure accurately predicts or correlates with the construct it is supposed to be measuring.

Aetiological validity is the extent to which sufferers of a disorder have the same causal factors, so for a patient to be diagnosed with a particular disorder, they should have the same symptoms as others with the same disorder.

Aetiological or etiological validity can be seen when making a diagnosis of a patient and it means that sufferers of a disease/ disorder have the same causal factors.

Features of science

Psychology is the scientific study of the mind and behaviour. It uses scientific methods to measure observable characteristics.

Science uses an empirical approach. This refers to the belief that knowledge is derived from observable, measurable experiences and evidence, rather than from intuition or speculation.

Psychology is a science because it employs systematic methods of observation, experimentation and data analysis to understand mental processes and predict behaviour.

Psychology is grounded in empirical evidence and subjected to peer review to ensure its credibility.

Empirical methods are scientific approaches which are based on gathering evidence from direct observation and experience.

Empirical evidence aims to be objective and minimises bias or subjective opinion.

Psychology collects empirical evidence which aims to be objective in order to make scientific claims about behaviour.

To be objective, all sources of personal bias are minimised so as not to distort or influence the research process.

Scientific researchers must strive to maintain objectivity as part of their investigations.

Research methods in Psychology associated with the highest level of control, such as lab experiments, tend to be the most objective.

Research methods which collect more quantitative data are more likely to be associated with higher levels of objectivity.

As with other sciences, Psychology follows a scientific method when testing theories or ideas.

The scientific method begins with a theory from which a hypothesis is formed and experimental research is conducted and evaluated.

John Locke (1689) suggested that a theory cannot claim to be scientific unless it has been empirically tested and verified.

Theory construction in Psychology is the process of developing an explanation for behaviour by systematically gathering evidence and organising it into a coherent idea.

A theory is a set of general laws or principles that have the ability to explain particular events or behaviours.

Theory construction occurs through gathering evidence via direct observation or experimentation.

An essential component of a theory is that it can be scientifically tested.

Most theories begin with a number of possible hypotheses which can be tested using objective methods to determine whether or not the idea can be supported or refuted.

Many theories in Psychology are the result of data collection from several research studies, often over time.

Milgram (1974) devised the Agency Theory as a result of data collection through his experiments on obedience to authority in the 1960s.

Bandura (1977) created the idea of mediational processes (ARRM) following his research on the social learning theory in the 1960s, such as the studies using the Bobo doll.

The process of deriving new hypotheses from an existing theory is known as deduction.

In science, deductive reasoning tests will verify hypotheses by deriving specific predictions or expectations that can be tested through experimentation or further observation.

Popper (1959) argued that science would best progress using deductive reasoning as its primary emphasis.

Deduction is drawing conclusions from specific facts but sometimes this can lead to false statements/ conclusions (which may not be entirely true). For example, saying 'flowers are red' and 'tulips are flowers' according to deduction would say 'all tulips are red'.

Older ideas in science focused on inductive reasoning. This is the idea that theories could only be formed following observations which lead to hypotheses.

Inductive reasoning involves drawing general conclusions based on specific observations or patterns, moving from specific instances to broader generalisations.

Induction is making generalisations from facts, usually based on observations following experimental research.

Today science focuses on deductive reasoning, where theories are created and hypotheses are written prior to any observational or experimental testing.

According to Popper (1959), science should attempt to disprove a theory rather than attempt to continually support theoretical hypotheses.

According to Popper (1959), scientific theory should make predictions that can be tested, and the theory should be rejected if these predictions are shown not to be correct.

Some people argue that Psychology is not a true science.

A true science is believed to be one which holds a set of shared beliefs, this is known as a paradigm.

A paradigm is a shared set of assumptions and agreed methods within a scientific discipline.

All traditional sciences, like a paradigm hold a shared set of assumptions and beliefs.

Kuhn (1962) suggested that social sciences like Psychology lack a universal paradigm.

Kuhn (1962) labelled Psychology as a pseudoscience/ fake science an argued that it is probably best seen as a 'pre-science'.

It is argued Psychology with its various approaches and perspectives cannot be seen as a true paradigm.

Some people argue that Psychology is marked by too much internal disagreement and has too many conflicting approaches to qualify as a science.

Psychology has been shown to have had several paradigm shifts over time. In the early 1900s much of Psychology focused on introspection and a mix of methods to study the unconscious mind. Whereas today it is much more focused on objective methods and the field of Biology.

A paradigm shift is the result of a scientific revolution where there is a significant change in the dominant unifying theory within a scientific principle.

According to Kuhn (1962) progress within an established science occurs when there is a scientific revolution.

A scientific revolution is where a handful of researchers begin to question the accepted paradigm. This critique gathers popularity and pace and can lead to a paradigm shift.

Another philosopher working around the same time as Kuhn was called Karl Popper.

Popper (1934) argued that the key criterion for a scientific theory is its falsification.

Falsification is that idea that a theory or idea could be potentially proven wrong. If we cannot falsify it, we must accept it as possibly true.

The idea of falsification suggests that for a theory to be considered scientific, it must be able to be tested and conceivably proven false.

Theories that survive most attempts to falsify them can become the strongest, this is not because they are true but because they have not been proved false.

Freud's work in the early 1900s on psychosexual development cannot be tested experimentally and therefore cannot be falsified. This means his theories must be accepted as plausible.

Most psychologists will avoid using the term 'proves' and instead favour the term 'supports' or 'suggests'.

Like traditional sciences, Psychology states a null hypothesis alongside the research hypothesis. This allows for falsifying, in that one hypothesis will be deemed correct and the other incorrect.

Popper (1959) introduced the idea of falsification. For a theory to be valid it must produce hypotheses that have the potential to be proven incorrect by empirical evidence.

Popper (1959) believed that scientific knowledge is provisional. This is the idea the science is constantly evolving as theories and ideas are proved to be right or wrong.

Genuine scientific theories should hold themselves up for hypothesis testing and for the possibility of being proven incorrect.

A true science is one in which theories are constantly challenged and can potentially be falsified, whereas pseudosciences are harder to be falsified.

Popper (1934) created the hypothetico-deductive method as a scientific process. It is a step-by-step, organised, and rigorous way to find the solution to a problem.

The hypothetico-deductive method begins by formulating a hypothesis in a form that can be falsifiable, using a test on observable data where the outcome is not yet known.

The hypothetico-deductive method is one of the most widely used scientific methods for disproving hypotheses and building corroboration for those that remain.

An important part of the hypothetico-deductive method is replicability. For example, a theory can only be trusted if the findings from which it is based on or shown to be replicable.

Replicability is the extent to which scientific procedures and findings can be repeated by other researchers.

Replication is important when determining the reliability of a method. If findings are similar after repeating a study several times, reliability is high.

If a new idea is proposed but cannot be tested and replicated by other scientists it will not be accepted.

Replication can be used to assess the validity of a finding, by repeating the study over a number of different contexts and circumstances to see the extent to which it can be generalised.

Reporting Psychological Investigations

Psychology follows the format of a scientific report when publishing journal articles.

The first section of a scientific report is called the abstract. This includes a short summary of all the major elements within the report.

An abstract is a short summary of about 150 to 200 words and provides key details from the whole of the research report.

An abstract should include a brief summary of the introduction, method, results and discussion.

The introduction section looks at past research (theories and studies) from a similar area of study. It begins quite broadly and gradually becomes more specific.

The introduction section includes the aims and hypotheses of the report.

The method section is divided into several subsections and is a description what the researchers did throughout the study.

The method section includes details about the sample, design, materials/ apparatus, procedure and ethics.

The purpose of a method section is to provide sufficient detail so the other researchers are able to precisely replicate the study if they wish.

The sample includes demographic information about the participants; how many, where they are from and any key characteristics.

The sample section will include the target population and sampling method.

The design section should clearly state the research method and give justification for its use. This can include the experimental design.

The design section will include a description of the variables investigated in the research, namely the independent variable and the dependent variable.

The design section could also include a description of any controls or extraneous variables.

The materials section should give detail of any apparatus or assessment materials used as part of the procedure.

A copy of any materials, such as questionnaires or other paper based tasks can be found in the appendices at the end of the report. This allows for replication by other researchers.

The procedure section should give a verbatim record of everything that was said and done to the participants. This includes the briefing statement, standardised instructions and debriefing.

The procedure section is a step by step outline of everything that happens in the investigation from beginning to end. This is necessary to ensure replication.

The ethics section may outline any explanation of how the issues were addressed within the study.

The results section of a psychological report should summarise any key findings from the research. It is not a place for raw data or calculations (these are found in the appendices).

The results section is likely to describe any descriptive statistics and show any summary illustrations such as graphs or charts.

The results section will also include inferential statistics making reference to the choice of statistical test, calculated and critical values and the level of significance.

The results section will outline whether the research hypothesis was accepted or rejected based on the outcome of the inferential statistics.

If qualitative research methods are used a conclusion of the themes or patterns will be outlined in the results section.

A discussion section of a psychological report takes into consideration what the results show and relates them back to the earlier psychological theories discussed in the introduction.

In the discussion section the researcher will summarise the findings in verbal form and discuss these in context of the evidence presented in the introduction.

The discussion section will consider how this research study relates to the others previously stated in the introduction.

In the discussion section the researcher should outline any limitations of the investigation and perhaps include suggestions of how these might be addressed in the future.

Wider implications and practical applications may be discussed at the end of a discussion section in a psychological report.

The end of a psychological report will discuss what contribution this investigation has made to the existing knowledge base within the field of study.

A full references section will be outlined at the end of a psychological report. This is a list of all of the sources refer to or quoted throughout the article.

The references section should outline every source that was cited in the psychological report. This can include any books, journal articles all websites.

Psychological reports use the Harvard referencing system when quoting the source material cited in the report.

Book referencing follows the Harvard referencing system which has a specific format. For example, author (publication date), title of book, place of publication, publisher.

Referencing journal articles will include the specific journal name, the volume of issue and it's specific page numbers.

The purpose of the reference section is to ensure other researchers can find all of the cited sources stated in the psychological report, should they wish.

At the back of a psychological report there will be an appendices section, which is a place for storing copies of any materials, raw data or calculations conducted during the process of the study.

Data Handling and Analysis

Quantitative and Qualitative Data

Quantitative data is numerical, whereas qualitative data is opinion based.

Quantitative data can be scored or counted and is usually given as numbers.

Quantitative data collection techniques usually gather numerical data in the form of individual scores from participants.

Quantitative data is open to being analysed statistically and can easily be converted into graphs or charts.

Quantitative data is quicker to analyse and comparisons between groups and conclusions can be drawn easily.

Quantitative data is reliable as most of the methods used to collect this type of data are easy to replicate.

Quantitative data tends to be more objective and less open to bias.

Quantitative data does not explain why people behave the way they do, or perform as such on tests, therefore is much narrower in meaning.

Qualitative data is expressed in words and is non-numerical.

Qualitative data may take the form of a written description or personal opinions.

Qualitative data is often gathered from open-ended questions or unstructured interviews.

Qualitative data offers researchers a detailed insight into behaviour.

Qualitative data tends to be more valid than quantitative data.

Qualitative data offers detailed descriptions into why people feel the way they do; this gives us a valid insight into behaviour.

Qualitative data is more time consuming to analyse and may result in an increased risk of interpretation bias.

Qualitative data often relies on self-report which is more subjective.

Psychological research can collect both qualitative and quantitative data offering a valid insight into behaviour based on reliable findings.

Experiments and structured observations tend to collect more quantitative data, whereas open-ended questionnaires and case studies tend to collect more qualitative data.

Primary Data

Primary data is information that has been obtained first hand by a researcher for the purpose of the investigation.

Primary data is gathered directly from participants as part of an experiment, self-report or observation.

Primary data is original data that has been collected for the purpose of an investigation by a researcher themselves.

Primary data is authentic as it is collected directly from the source.

Primary data tends to be more reliable as it is collected first hand.

Primary data can be time consuming and requires effort on the part of the researcher.

Planning, designing and conducting research studies which collect primary data can be very time consuming.

Secondary Data

Secondary data is information that has already been collected by someone else and pre-dates the current research investigation.

Secondary data is information which is obtained from another source or researcher.

Secondary data is collected by someone other than the person conducting the research investigation.

Secondary data can be located in books, journal articles or websites.

Secondary data can produce a valid picture by providing detail about the area of study.

Secondary data may be inexpensive and easily accessed requiring minimal effort from a researcher.

Secondary data is subject to interpretation bias, as the individual researcher will select the information they require.

Secondary data can be time consuming if multiple sources are referred to in the investigation.

There may be variation in the quality in the accuracy of secondary data, and sources which are used for reference must be credible.

Meta-analysis

A meta-analysis is the process of combining the findings from a number of studies on a particular topic.

A meta-analysis aims to produce an overall statistical conclusion based on a range of studies.

A meta-analysis compares the data from multiple studies in order to form a statistical conclusion.

A meta-analysis is a research method that uses secondary data.

During a meta-analysis the researcher will process a number of studies by analysing the results and forming a joint conclusion.

The van Ijzendoorn (1988) study investigating cross-cultural variation in attachment is an example of a meta-analysis.

A meta-analysis is different to a review article.

A meta-analysis creates a larger more varied sample from which results can be generalised increasing population validity.

Meta-analysis may be subject to publication bias, as the researchers may only refer to the information that suits their aim.

In a meta-analysis the researcher may not select all relevant studies, choosing to leave out those studies with negative or non-significant results. This ends with a biassed conclusion.

Descriptive statistics

Descriptive statistics are a way of presenting a summary of the results which have been analysed in an investigation.

Descriptive statistics will use graphs and tables to summarise the statistics of an investigation in order to identify trends and analyse patterns.

Descriptive statistics include measures of central tendency. These measures are averages which give us information about the most typical values in a set of data.

Measures of central tendency include the mean the median and the mode.

The mean is an arithmetic average calculated by adding up the all the values in a data set and dividing them by the number of those values.

The mean value is also known as the average.

The mean is calculated by taking the total value of all the scores and dividing them by the number of scores in the data set.

The mean is the most sensitive of the measures of central tendency as it includes all of the values in the data set. This makes it more representative as a whole.

The mean value is easily distorted by extreme values, which can skew the data.

Calculations showing the mean scores are often illustrated using bar charts.

In a perfect normal distribution curve the mean value will be displayed at the peak point of the curve.

The median is a central value in a set of data when all the values are arranged from lowest to highest.

The median is the middle value in a data set when scores are arranged in numerical order.

The median is easy to calculate once the scores are placed in order.

A strength of calculating the median is that it is less affected by extreme values.

The median is less sensitive to variations in the data, so may not present a true picture.

The median is less sensitive than the mean as the actual values of lower and higher numbers are ignored and extreme values may be important.

The mode is the most frequently occurring value in a set of data.

The mode is the most common number which appears in the data set. There may be two modes in the set (bimodal).

The mode is very easy to calculate, however when there are several modes in a data set the analysis is less useful.

The mode is commonly used with category data/ nominal data.

The mode is often used when there is frequency or category data as other measures are not appropriate.

In some cases there may be no modal value which does not provide useful information to form conclusions.

Measures of Dispersion

Measures of dispersion are a measure of the spread or variation in a set of scores.

Measures of dispersion are based on the spread of scores, this is how they vary and differ from one another.

The range is a measure of dispersion.

The range is a simple calculation of the dispersion in a set of scores which is calculated by subtracting the lowest score from the highest score, and often adding 1 as a mathematical correction.

Adding 1 for mathematical correction when calculating the range allows for the fact that often raw scores are rounded up or down when they are recorded.

The range is easy to calculate, however it only takes into account the two most extreme values. This may be unrepresentative of the data set as a whole.

The range can be influenced by outliers and this can skew the data.

The range does not indicate whether most scores are closely group together around the mean or whether they are spread out.

Standard deviation is a more sophisticated measure of dispersion in a set of scores.

The standard deviation tells us by how much, on average each score deviates from the mean.

The standard deviation is a single value that tells us how far the scores are spread away from the mean. The higher the value the more variation there is in the set of scores.

The larger the standard deviation, the greater the dispersion or spread within a set of data.

Below standard deviation value reflects the fact that the data is tightly clustered around the mean. This implies that all participants responded in a very similar way.

Deviation is a more precise measure of dispersion than the range as it includes all values within the final calculation. However like the mean this can be distorted by a single extreme value.

To calculate the standard deviation the difference between each score and the mean is calculated and squared. These are totalled up and divided by the number of scores. This gives the variance on the standard deviation is the square root of this.

Percentages

Percentages are calculated by taking the score, dividing it by the total and multiplying by 100.

Percentages are useful to show overall comparisons between groups or conditions.

To calculate a percentage, divide the numerator (number on the top) by the denominator (number on the bottom) and multiply by 100.

To convert a percentage to a decimal, simply remove the % sign and move the decimal .2 places to the left.

Fractions

A fraction is a part of a whole.

To find a fraction of an amount, divide by the denominator. The denominator is the number of equal parts. E.g., for 1⁄3, the denominator is 3 and multiply by the numerator. The numerator is the number of parts used.

Fractions should be reduced by finding the highest common factor.

Ratios

A ratio is a way of comparing two or more quantities. For example, 2:3 is a ratio, which means for every two parts of one thing, there are three parts of another.

Ratios compare two numbers, usually by dividing them.

A ratio shows how much of one thing there is compared to another. Ratios are usually written in the form: A:B.

Ratios should be reduced by finding the highest common factor.

Presentation and Display of Quantitative Data

There are various ways of representing data either in tables or graphs.

Data can be summarised in the form of a summary table. This does not show any raw data but just the main findings.

Summary tables are often used to display the descriptive statistics. This allows a clear comparison to be seen between the groups or conditions.

It is standard practise to include a verbal summary underneath a table of results explaining what the findings show.

Different graphs are used in Psychology to illustrate the data from an investigation.

A bar chart is a type of graph in which the frequency of each variable/ condition is represented by the height of the bars.

Bar charts visually represent data using individual bars with spaces between.

Bar charts are used when data is divided into categories, known as discrete data.

Mean values are often displayed in bar charts.

In a bar chart the bars are separated to illustrate that there are separate conditions.

In a bar chart categories or conditions are usually plot on the horizontal X axis. Whereas average scores or frequency data is plot on the vertical Y axis.

Histograms are a type of graph which shows frequency, but on like a bar chart the area of the bars, not just the height represents frequency.

The bars in a histogram can be different heights and different widths.

In a histogram the bars touch each other, and the data displayed along the X axis is continuous rather than discrete.

In a histogram the X axis must start at a true 0 as the scale is continuous.

Scattergrams are a type of graph that represents the strength and direction of the relationship between co-variables in a correlational analysis.

Correlations are illustrated using scattergrams.

Scattergrams do not illustrate differences between conditions but instead show associations between co-variables.

In a scattergram one co-variable is plot along the X axis and the other co-variable along the Y axis. Where the two scores meet across is placed, and from this a line of best fit can be drawn.

Scattergrams can indicate the type and strength of a correlation from how the points are plotted.

Correlations

Correlation is a mathematical technique in which a researcher investigates an association between two variables, called co-variables.

A positive correlation states that as one co-variable increases so does the other.

A negative correlation states as one co-variable increases the other decreases.

Zero correlation is when there is no relationship between the co-variables.

Correlations are numerical values depicted by a correlation coefficient.

A correlation coefficient is a number between -1 and +1 that represents the direction and strength of a relationship between co-variables.

The correlation coefficient for perfect positive correlation is +1.

The correlation coefficient for a perfect negative correlation is -1.

The correlation coefficient no/ zero correlation is 0.

A correlation coefficient closer to a score of 1 indicates a stronger correlation.

Correlation coefficient closer to 0 indicates a weaker correlation.

A correlation coefficient of 0.95 indicates a strong positive correlation, whereas a correlation coefficient of -0.01 indicates a weak negative correlation.

Correlation coefficients that appear to indicate weak correlations can still be statistically significant depending on the size of the data set.

Distributions

Normal Distribution

Normal distribution is a symmetrical spread of frequency data that forms a bell shaped pattern.

In a normal distribution the mean, median and mode are all located at the highest peak.

In a perfect normal distribution the mean, median and mode are identical values.

Within a normal distribution most scores are located in the middle area of the curve, with very few scores at the extreme ends.

In a normal distribution curve the tails of the curve never touch the horizontal X axis (never reach 0) as more extreme scores are always possible.

Skewed Distributions

Not all distributions form a balanced symmetrical pattern. Some data sets derive from psychological scales or measurements may produce skewed distributions.

Skewed distribution is a spread of frequency data that is not symmetrical, where the data clusters to one end.

A skewed distribution is one that appears to lean to one side or the other. There could be a positive or a negative skew in the data set.

A positive skew is a type of frequency distribution in which the long tail is on the positive (right) side of the peak and most of the distribution is concentrated on the left.

A positive skew indicates the average score is being dragged up, shifting the mean score towards the long tail on the right.

In a positive skew there could be one extreme value shifting the average score up towards the right.

In a positive skew the majority of scores remain in the middle of the distribution curve, with only a few scores pulling the average score up.

A negative skew is a type of frequency distribution in which the long tail is on the negative (left) side of the peak and most of the distribution is concentrated on the right.

A negative skew indicates the average score is brilliant dragged down, shifting the mean score towards the long tail on the left.

In a negative skew that could be one extreme value shifting the average score down towards the left.

In a negative skew the majority of scores remain in the middle of the distribution curve, with only a few scores pulling the average score down.

Statistical calculations such as the standard deviation can indicate the spread of scores in a normal distribution. The wider the distribution the more variation in the scores.

Levels of Measurement

Quantitative data can be classified into types or levels of measurement.

There are three levels of measurement, nominal, ordinal and interval.

Levels of measurement influence the choice of statistical test used when analysing results from investigations.

Levels of measurement are quantitative, numerical values which are used in statistical analysis.

Nominal data is represented in the form of categories.

Nominal data is sometimes referred to as categorical data.

Nominal data is discrete in that one item can only appear in one of the categories.

Nominal data is often collected in observations through tally charts.

Nominal data can be gathered through questionnaires asking simple questions like yes or no.

For nominal data the most appropriate measure of central tendency is the mode.

Ordinal data is data which can be placed in order.

Ordinal data is usually ranked giving positions of first to last place.

Ordinal data is gathered from individual scores, this is usually per participant.

Ordinal data does not have equal intervals between each unit.

Ordinal data lacks precision because it is based on subjective opinion rather than objective measures.

Ordinal data can be gathered from questionnaires (total scores per participant) or correlations (two sets of numerical values).

Raw data is converted to rank scores to become ordinal data. This is then used in statistical tests for analysis.

For ordinal data the most appropriate measure of central tendency is the median and the most appropriate measure of dispersion is the range.

Interval data is based on numerical scales all measurements the include units of equal precisely defined size.

Interval data comes from specific units of measurement, for example, centimetres on a ruler or time in minutes from a stopwatch.

Public scales of measurement that produce data based on accepted units of measurement make up interval data.

Interval data is the most precise and sophisticated form of data in psychology.

Interval data is a necessary criterion for the use of parametric tests, such as related or unrelated T-tests or a Pearson's r correlation test.

For interval data the most appropriate measure of central tendency is the mean and the most appropriate measure of dispersion is the standard deviation.

Content Analysis and Coding

A content analysis is a research technique the enables the indirect study of behaviour by examining communications that people produce, for example in texts, TV and film or other forms of media.

The aim of a content analysis is to summarise and describe communication in a systematic way so overall conclusions can be drawn.

A content analysis will transfer qualitative data into quantitative data for comparison and analysis.

Coding is the stage of content analysis in which the communication is analysed by identifying each instance of the chosen categories.

Coding is the initial stage of content analysis.

The qualitative data gathered in a content analysis will need to be categorised into meaningful units. This is the process of coding.

Coding may involve counting the number of times a particular word or phrase appears in the text. This will produce quantitative data that can be analysed.

An example of coding in a content analysis, would be to count the number of times the word paranoid appears in a transcript of an interview with a schizophrenic patient.

Content analysis is useful in that it usually avoids ethical issues in research.

Content analysis can be a very time consuming process.

Content analysis may be subjective as there is a danger the researchers may attribute their own opinions to the research.

Content analyses are seen as less objective than other research methods, especially when more descriptive forms of thematic analysis are used.

Thematic Analysis

A thematic analysis is a qualitative approach which involves identifying ideas or concepts within the data. Themes will often emerge once the data has been coded.

Thematic analysis is a form of content analysis but the outcome is qualitative rather than quantitative.

In a thematic analysis the main process involves the identification of key themes or patterns.

A theme in content analysis refers to any idea that is recurrent in the communication being studied.

Thematic analyses are a valid way of analysing data as they typically report qualitative themes.

Thematic analysis is likely to be more descriptive and may struggle to code information.

As with content analysis, thematic analysis can be a very time consuming process.

Inferential Testing

Inferential testing is used in Psychology to determine whether a significant difference or correlation exists within the results.

In a study, a difference may be seen in average (mean) scores but this does not indicate whether or not the difference is significant, for this a statistical test must be calculated.

Statistical tests are used to determine the significance of the findings in an investigation.

All statistical tests end with a calculated value which is numerical. This number will determine whether the researcher has found a statistically significantly result or not.

Calculated values from statistical tests are compared against critical values in order to determine their significance.

A critical value is a numerical boundary determined by mathematicians, outlined in critical value tables which tell us whether a result is significant or not.

Significance is a statistical term that tell us how certain we are a difference or correlation exists.

Researchers in Psychology begin their investigations by writing a hypothesis. This may be directional or non-directional depending on how confident the researchers are with the outcome.

A directional hypothesis is usually written if there is prior research in this area of study.

A research hypothesis will predict the outcome of the results, whereas a null hypothesis will state there is no difference or correlation.

A null hypothesis will state there is no significant difference and any difference found will be due to chance.

A statistical test determines which hypothesis should be accepted and which should be rejected.

If a statistical test concludes a result is significant the research hypothesis will be accepted.

If a result is significant the research hypothesis will be accepted and the null hypothesis rejected.

Probability and Significance

Probability is a measure of the likelihood that a particular event will occur where 0 indicates statistical impossibility, and one indicates statistical certainty.

Statistical tests work on the basis of probability rather than certainty and all tests employ a significance level.

A significance level is the point at which the researcher can claim to have discovered a large enough difference or correlation within the data to state an effect has been found.

The usual level of significance in Psychology is 0.05 (5%).

In Psychology the probability is stated as $p \leq 0.05$. This means the probability that the observed effect (outcome of the results) occurred when there is no effect in the population is equal to or less than 5%.

In Psychology even when a researcher claims to have found a significant difference or correlation, there is still up to a 5% chance that this is not true.

Psychologist can never be 100% certain about a particular result, as they have not tested all members of the population under all possible circumstances. They are also working with people, who unlike chemicals in Chemistry may react unpredictably.

In psychological research, the 5% significance level ensures that there is equal to or less than a 5% chance the results are down to external factors, not manipulated by the researchers.

In Psychology researchers accept a 5% margin of error, which could be the result of chance, and instead suggest the outcome of the result is 95% due to the manipulation of the researchers.

Typical traditional sciences like Biology will often use a more stringent level of probability, such as $p \leq 0.01$.

The more stringent the level of probability, the more certain the researchers are that the results are not due to chance.

Use of Statistical Tables

Once the statistical test has been calculated the final numerical value is called the calculated value.

To check for statistical significance, the calculated value from a statistical test must be compared with the critical value.

The critical value tells us whether or not we can accept the research hypothesis and reject the null hypothesis.

Each statistical test has its own table of critical values which were developed by statisticians.

Each statistical test will have a rule underneath the critical values, explaining what determines a statistical outcome.

For some statistical tests, the calculated value must be equal to or greater than the critical value to be significant. Whereas in other tests the calculated value must be equal to or less than the critical value.

All the statistical tests with a letter 'r' in their name have a rule which states; the calculated value must be equal to or more than the critical value to be significant.

When using critical value tables there are three criteria; identifying the direction of the hypothesis, noting the number of participants and deciding on the level of significance.

In order to identify the critical value in the table the researcher must check to see if the hypothesis is one-tailed or two-tailed.

In order to identify the critical value in the table the researcher must check the number of participants in the study, this is often noted as 'N'. Some statistical tests will have used category data, in which case, degrees of freedom are calculated (e.g. Chi Squared test).

To identify the critical value in the table the researcher must check the level of significance or P value. In most research the level of significance is the standard 0.05 level.

Type I and Type II Errors

As researchers can never be 100% certain that they have found statistical significance, it is possible that the wrong hypothesis may be accepted.

When an incorrect hypothesis is accepted it is known as an error.

A type I error is made when the research hypothesis is accepted and the null hypothesis rejected, when in fact it should be the other way round.

A type I error is a mistake of accepting the research hypothesis when in fact it is incorrect.

A type I error is also known as a false positive.

A type I error is the incorrect rejection of a true null hypothesis.

Researchers are more likely to make a type I error if the significance level is too lenient/ too high, e.g. p≤0.10 (10%).

A type II error is made when the research hypothesis is rejected and the null hypothesis is accepted, when in fact it should be the other way round.

A type II error is a mistake of accepting the null hypothesis when in fact it is incorrect.

A type II error is also known as a false negative.

A type II error is the failure to reject a false null hypothesis.

Researchers are more likely to make a type II error if the significance level is too stringent/ too low, e.g. p≤0.01 (1%), as potentially significant values may be missed.

Psychologists prefer to use the 5% level of significance as it best balances the risk of making a type I or type II error.

Introduction to Statistical Testing

When choosing a statistical test researchers will have decided if the investigation is a test of difference or a correlation.

Statistical tests can be parametric or non-parametric.

Parametric tests are used on normally distributed data, and non-parametric tests are on data that is not normally distributed.

Parametric tests are those that make assumptions about the parameters of the population distribution from which the sample is drawn.

Parametric tests are often based on the assumption that the population data is normally distributed.

In a parametric test the data can be assumed to be normally distributed because the population from which it is drawn would seem to be normal for the variable under investigation.

Non-parametric tests are 'distribution-free' and may include data which can be skewed by outliers or extreme scores.

In a non-parametric test the data might not be normally distributed due to extreme scores.

Statistical tests vary depending on whether they are a test of difference or a correlation.

In most experimental designs researchers are looking for a difference between groups or conditions.

Experimental designs can be independent groups, repeated measures or matched pairs.

An independent groups design is also known as an unrelated design. The participants in one condition are unrelated/ separate to the participants in another condition.

Repeated measures and matched pairs designs are also known as related designs.

In a repeated measures design, the same group of participants are used in all conditions of the experiment.

In a matched pairs design, participants in each condition are separate but have been 'matched' on key characteristics, which makes them related.

Levels of measurement are considered when choosing which statistical test to use. Some tests will use the basic level of nominal data, whereas others will use ranked ordinal data.

Nominal data is represented in the form of categories. This is discrete data as it can only appear in one category.

Non-parametric tests such as the Sign test and Chi Squared test will collect nominal data.

Ordinal data is based on participant scores which can be rated from first to last position. Most typical experiments in Psychology collect ordinal data.

Non-parametric tests such as Mann Whitney, Wilcoxon and Spearman's Rho will collect ordinal data.

Interval data is based on specific units of measurement or numerical scales, where each unit is equal and precisely defined. Experiments using specific apparatus will often collect interval data. E.g. a stopwatch will record interval data, in specific minutes or seconds.

Parametric tests such as related and unrelated t-tests and Peason's r correlation, use interval data.

Parametric tests are not calculated in A-level Psychology.

Sign Test

The sign test is a non-parametric statistical test of difference that allows a researcher to determine the significance of their investigation.

The sign test is a statistical test for a difference in scores between related items, where data should be nominal or better.

A sign test is used in studies that have a repeated measures or matched pairs design, where the data collected is nominal.

A sign test looks for a difference, using a repeated measures design collecting nominal data.

When conducting a sign test the nominal data must be converted by identifying which condition participants performed better or worse in.

If a one-tailed hypothesis is used during a sign test, this experimental condition becomes the positive sign. All participants with a higher score in this condition will receive a plus (+) sign.

In a sign test, participants who score higher in the condition which is not predicted to influence the results will receive a minus (-) sign.

In a sign test, if a participant gives identical scores in both conditions, their results can be discarded. This reduces the overall sample size.

A sign test results in a calculated value known as S.

In a sign test the total number of pluses and minuses are compared in order to reach the calculated value. The less frequent sign becomes the calculated value of S.

The calculated value in a sign test (S) is compared against the critical value in the table to determine whether the result is significant or not.

The calculated value from a sign test (S) must be equal to or less than the critical value to be significant.

In a sign test, if the calculated value is equal to or lower than the critical value then the results are significant at the 5% level of probability.

Mann-Whitney Test

The mann whitney test is a non-parametric statistical test of difference that allows a researcher to determine the significance of their investigation.

The mann whitney test is a statistical test of difference between unrelated items, where data should be ordinal or better.

A mann whitney test is used in studies that have an independent measures design, where the data collected is ordinal.

A mann whitney test looks for a difference, using an independent measures design collecting ordinal data.

The calculated value in a mann whitney test is also known as U.

A mann whitney test uses a formula to calculate the value of U.

In order to calculate the formula in a mann whitney test, the data from each group must be ranked and given a total.

When ranking data in a mann whitney test, the data must be viewed as one whole group, where the lowest number receives a rank of 1.

In a mann whitney test the final U value is the smaller of Ua and Ub.

Most experiments in Psychology using an independent measures design will use a mann whitney test.

The calculated value from a mann whitney (U) test must be equal to or less than the critical value to be significant.

In a mann whitney test if the calculated value is equal to or lower than the critical value then the results are significant at the 5% level of probability.

Wilcoxon Test

The wilcoxon test is a non-parametric statistical test of difference that allows a researcher to determine the significance of their investigation.

The wilcoxon test is a statistical test of difference between related items, where data should be ordinal or better.

A wilcoxon test is used in studies that have a repeated measures design, where the data collected is ordinal.

A wilcoxon test looks for a difference, using a repeated measures design collecting ordinal data.

Most experiments in Psychology using a repeated measures design will use a wilcoxon test.

The calculated value in a wilcoxon test is also known as T.

To identify the T value in a wilcoxon test, the difference between each participant's scores in both conditions must be calculated, then the difference must be ranked.

In a wilcoxon test we ignore any scores with values of 0 or differences with a 0 score. Also any signs in front of values are ignored when calculating the difference.

The calculated value from a wilcoxon (T) test must be equal to or less than the critical value to be significant.

In a wilcoxon test if the calculated value is equal to or lower than the critical value then the results are significant at the 5% level of probability.

Chi-Squared Test

The chi-squared test is a non-parametric statistical test of difference or association that allows a researcher to determine the significance of their investigation.

The chi-squared test is a statistical test of difference or association between unrelated items, where data should be nominal.

A chi-squared test is used in studies that have an independent measures design, where the data collected is nominal.

A chi-squared test looks for a difference or association, using an independent measures design collecting nominal data.

Most observations in Psychology collecting nominal data will use the chi-squared test.

The calculated value in a chi-squared test is also known as χ^2.

To identify the calculated value in a chi-squared test, expected frequencies must be calculated.

Expected frequencies in a chi-square test use a formula based on the data for each cell in the contingency table.

Chi-squared tests use nominal data which means categories are counted and there are no participant numbers. This means degrees of freedom are used to determine the significance.

Degrees of freedom (df) are calculated by multiplying (rows – 1) x (columns – 1) In the contingency table.

The calculated value from a chi-squared (χ^2) test must be equal to or more than the critical value to be significant.

In a chi-squared test if the calculated value is equal to or more than the critical value then the results are significant at the 5% level of probability.

Spearman's Rho Test

The spearman's rho test is a non-parametric statistical test of association that allows a researcher to determine the significance of their investigation.

The spearman's rho test is used as a test of correlation between two sets of data.

The spearman's rho test is a statistical test of association between related items, where data should be ordinal or better.

A spearman's rho test is used in studies that have a typical repeated design, however, as correlations are not experiments, there is no such experimental design necessary.

A spearman's rho test must collect data which is at least ordinal in level.

Most correlations in Psychology collecting ordinal data will use the spearman's rho test.

The calculated value in a spearman's rho test is also known as rho.

The calculated value in a spearman's rho test is the correlation coefficient. This value must be between -1 and +1.

A spearman's rho test uses a formula to calculate the value of rho.

In order to calculate the formula in a spearman's rho test, the data from each group must be ranked and given a total.

When ranking data in a spearman's rho test, the data must be viewed as two separate groups, where the lowest number receives a rank of 1.

In a spearman's rho test, when identifying the critical value from the table, the positive or negative correlation sign can be ignored.

The calculated value from a spearman's rho test must be equal to or more than the critical value to be significant.

In a spearman's rho test if the calculated value is equal to or more than the critical value then the results are significant at the 5% level of probability.

Pearson's r Test

Pearson's r is a test of correlation between two sets of values.

Pearson's r test uses interval data.

Pearson's r test is a parametric test used to investigate correlational research.

Correlational research analyses the results from two separate variables. These are used in a formula to calculate the result of the Pearson's r test.

In a Pearson's r test if the calculated value is more than the critical value, the result is significant.

T-Tests

A t-test is used to test for a difference between two data sets/ conditions.

A t-test uses data that is interval in level.

A t-test is a parametric test which is used instead of the wilcoxon test or the mann-whitney u test.

In a t-test it is assumed that participants are drawn from a normally distributed population and there is homogeneity a variance between the two data sets.

There are two types of t-test; the related t-test and the unrelated t-test.

Related T-Test

The related t-test is a parametric test used to replace the non-parametric wilcoxon test.

The related t-test uses a repeated measures design or matched pairs design.

The related t-test uses interval data.

In a related t-test if the calculated value of T is greater than the critical value, the result is significant.

Unrelated T-Test

The unrelated t-test is a test of difference between two sets of data.

The unrelated t-test is a parametric test used to replace the non-parametric mann whitney u test.

The unrelated t-test uses an independent measures design.

The unrelated t-test uses interval data.

In an unrelated t-test if the calculated value of T is less than the critical value, the result is significant.