AQA Psychology A-level

Research Methods

Detailed Notes

This work by PMT Education is licensed under CC BY-NC-ND 4.0

Part 1 — Research Methods

1.1. Experimental method

The **experimental method** involves a researcher **manipulating an independent variable (IV)** to measure its effect on a **dependent variable (DV)**. Other variables are controlled as far as possible to determine whether changes in the DV were **caused by the IV**.

There are four main types of experiment:

Type	Description	Strengths	Limitations	Example
Laboratory	An experiment conducted in a **controlled environment** where the researcher can carefully control extraneous variables.	**+ High control** – makes it easier to identify whether the IV caused changes in the DV. **+ Replicable** – standardised procedures can be repeated to check the findings.	**- Low ecological validity** – the controlled setting may be artificial and unlike real-life situations. - **Demand characteristics** – participants may guess the aim and change their behaviour.	A researcher tests whether **background noise (IV)** affects memory by asking participants to recall words in a controlled laboratory.
Field	An experiment conducted in a **natural environment**, while the researcher still **manipulates the IV** and measures the DV.	**+ High ecological validity** – behaviour is more likely to reflect real-life behaviour.	**- Less control** – extraneous variables are harder to control, making it more difficult to establish cause and effect. **Replication is more difficult** because the setting cannot be controlled as precisely.	A researcher changes the **music played in a supermarket (IV)** and measures how long customers spend shopping.
Natural	An experiment where the **IV occurs naturally** and is not manipulated by the researcher.	+ Allows researchers to study **real-life events** that would be **impossible or unethical to manipulate**, such	**- Low control** – other variables may differ between conditions, making it difficult to establish cause and effect.	A researcher compares people's **stress levels before and after a natural disaster** such as an earthquake.

CC BY NC ND

	The researcher measures its effect on the DV.	as natural disasters.	**- Difficult to replicate** because the event may be unpredictable or rare.	
Quasi	An experiment where the IV is a **pre-existing characteristic or difference between people**, rather than something manipulated by the researcher, such as **age or gender**.	**+** Can investigate variables that **cannot be ethically or practically manipulated**.	- Participants **cannot be randomly allocated** to conditions because the IV already exists. This makes it harder to establish that the IV **caused** differences in the DV.	A researcher compares **memory performance in younger and older adults**, where **age** is the IV.

CC BY NC ND

1.2. Observational Techniques

An **observation** involves systematically **watching and recording behaviour**. Researchers can vary **where the observation takes place**, whether participants **know they are being observed**, and whether the researcher **takes part in the situation**.

Type	Description	Strengths	Limitations	Example
Naturalistic	Behaviour is observed and recorded in the **setting where it would normally occur**.	**+ High ecological validity** – behaviour is more likely to reflect how people behave in everyday life.	**- Less control** over extraneous variables. **- Replication can be difficult** because the setting cannot be standardised.	Observing how children **play in a school playground**.
Controlled	Behaviour is observed and recorded in a **structured or controlled environment**, such as a laboratory.	**+ Greater control** over extraneous variables. **+Easy to replicate** because the setting and procedure can be standardised.	**- Behaviour may be less natural,** reducing ecological validity. **- Demand characteristics** may occur if participants realise they are being observed for a particular purpose.	Observing how participants **interact with a stranger** in a laboratory setting.
Overt	Participants **know they are being observed** and their behaviour is being recorded.	**+ Ethically acceptable** because participants can give **informed consent**.	**- Participants may change** their behaviour because they **know they are being watched**, reducing the validity of the findings.	Telling students that their **classroom behaviour** is being observed during a lesson.

CC BY NC ND

Covert	Participants are **unaware that they are being observed** and their behaviour is recorded without their knowledge.	**+ Natural behaviour as** participants are less likely to change their behaviour	**- Ethical issues** because participants cannot give informed consent. There may also be concerns about **privacy**.	A researcher observes **customer behaviour in a public shop** without revealing that they are being studied.
Participant	The researcher **joins the group** they are observing and takes part in the situation.	**+ Can provide greater insight** into the behaviour and experiences of the group.	**- The researcher may become too involved and lose objectivity**. **- Hawthorne effect,** participants may also change their behaviour if they discover they are being observed	A researcher **joins a youth group** to observe how members interact with each other.
Non-participant	The researcher observes the group **without taking part** in the situation.	**+ The researcher can remain more objective** because they are separate from the group.	- **The researcher may have less insight** into the group's behaviour and experiences. **- Observer bias** may also influence how behaviour is recorded or interpreted.	A researcher watches a **classroom from the back** without taking part in the lesson.

CC BY NC ND

1.3. Self-report techniques

Self-report techniques involve asking participants to provide information about their **thoughts, feelings, experiences or behaviour**. The two main self-report techniques are **questionnaires** and **interviews**.

Questionnaires

A **questionnaire** is a set of written questions used to collect information from participants. There are two main types of question:

Type	Description	Example
Open questions	Participants can answer **in their own words**, with no fixed response options. This produces mainly **qualitative data**.	*"How does school make you feel?"*
Closed questions	Participants choose from a **fixed set of responses**, producing data that can usually be easily **categorised and quantified**.	*"How often do you feel stressed at school?"* **Never / Rarely / Sometimes / Often / Always**

Closed questions can include:

- **Likert scales** – participants indicate their level of **agreement** with a statement.
- **Rating scales** – participants select a value representing the **strength of their feeling or opinion**.
- **Fixed-choice questions** – participants select from a list of **provided options**.

Strengths of questionnaires

- **Quick and cost-effective** – large amounts of data can be collected from many participants relatively cheaply.
- **Easy to analyse** – particularly when using closed questions, as responses can be converted into **quantitative data**.
- **Anonymity** – participants may feel more comfortable giving honest answers, particularly about **sensitive topics**.
- **Standardised** – every participant receives the same questions, making responses easier to **compare**.

CC BY NC ND

Limitations of questionnaires

- **Lack of detail** – closed questions can restrict participants from explaining their thoughts or experiences fully.
- **Response bias** – participants may not answer truthfully or carefully. For example, **social desirability bias** occurs when participants give answers that make them appear more socially acceptable.
- **Acquiescence bias** – participants may have a tendency to **agree with statements**, regardless of their actual opinion.
- **Validity** – if participants misunderstand a question or give socially desirable answers, the questionnaire may not accurately measure what it intends to measure.

Interviews

An **interview** is a self-report technique in which an **interviewer asks questions directly to an interviewee** to obtain information about their thoughts, feelings, experiences or behaviour. There are **two main types of interview**:

Interview type	Description	Strengths	Limitations
Structured	Uses a **set of predetermined questions**, asked in the **same order and usually in the same way** for every participant.	**+ Standardised** – all participants are asked the same questions. **+ Easy to compare** responses between participants. **+ Replicable** because the same interview schedule can be used again.	**- Limited depth** – participants can't develop their answers. **- Interviewer effects** may influence participants' responses. **- Social desirability bias** may cause participants to give socially acceptable answers.
Unstructured	Does not use a **fixed set of questions**. The conversation develops **naturally**, with follow-up questions based on the participant's responses.	+ Produces **rich, detailed information**. + Questions can be **adapted to the individual**, allowing interesting or unexpected responses to be explored.	- **Difficult to compare** responses as participants may be asked different questions. **- Difficult to analyse** due to large amounts of qualitative data. **- Requires a skilled interviewer,** increasing

CC BY NC ND

		+ Provides greater insight into participants' experiences.	the possibility of interviewer bias. **- Social desirability bias** may still affect responses.

1.4. Correlations

A **correlation** is a **non-experimental research method** used to investigate the **relationship between two variables**, known as **co-variables**.

Researchers **measure both co-variables** and analyse the data to see whether there is a relationship between them. **Neither variable is manipulated** by the researcher.

Example: A researcher could investigate the relationship between **hours of sleep** and **memory test scores**. They would measure how much sleep participants get and their memory scores, then analyse whether the two variables are related.

Correlation Coefficients

A **correlation coefficient** is a numerical value that shows the **strength and direction** of the relationship between two co-variables.

Correlation coefficients range from **-1 to +1**.

- The **sign (+ or -)** shows the **direction** of the relationship.
- The **size of the number** shows the **strength** of the relationship.

Coefficient	**Meaning**
+1	Perfect **positive** correlation
+0.80	Strong **positive** correlation
+0.20	Weak **positive** correlation
0	**No correlation**
-0.20	Weak **negative** correlation
-0.80	Strong **negative** correlation
-1	Perfect **negative** correlation

The closer the coefficient is to **+1 or -1**, the **stronger** the relationship. The closer it is to **0**, the **weaker** the relationship.

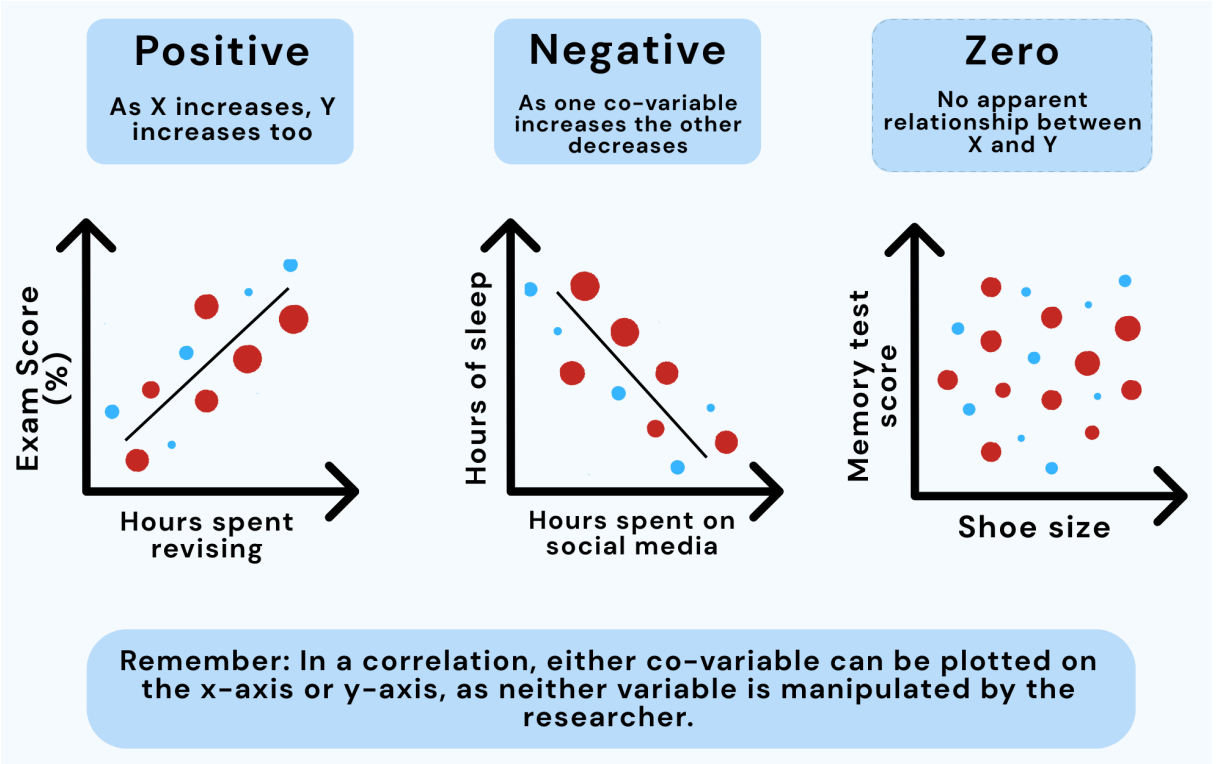
Types of Correlation

- **Positive correlation -** Both **co-variables increase together**. For example, the more hours a student spends **revising**, the higher their **test score** tends to be. The correlation coefficient is **above 0**.
- **Negative correlation -** As one **co-variable increases**, the other **decreases**. For example, **the more hours a student spends on social media, the fewer hours of sleep they get**. The correlation coefficient is **below 0**.
- **Zero correlation -** There is **no relationship** between the two co-variables. For example, **shoe size and test score** would be expected to have no correlation. The correlation coefficient is **0**.

Scattergrams

A **scattergram** is used to display the relationship between two **co-variables**. Each **dot represents a pair of scores** from one participant.

A **line of best fit** can be added to show the **overall direction of the relationship**. The **closer the points are to the line of best fit**, the **stronger** the correlation.



Correlations VS Experiments

The main difference is **what the researcher does with the variables**.

Experiment

- The researcher **manipulates an IV** and measures its effect on a **DV**.
- This allows researchers to investigate **cause and effect**.

Correlation

- The researcher **measures two co-variables** and looks for a **relationship** between them.
- Neither variable is **manipulated**.
- Therefore, a correlation can show an **association**, but **not cause and effect**.

Example: A correlation may find that **stress and poor sleep** are related. However, this does not show that stress causes poor sleep. **Poor sleep could cause stress**, or a **third variable** could cause both.

Experiment → cause and effect
Correlation → association

Strengths	Limitations
Quick and economical – researchers only need to measure the two co-variables, rather than manipulate an IV and control other variables.	**Cannot establish cause and effect** – a correlation only shows an association between variables.
Useful starting point – identifying a relationship can lead to further research into why the relationship exists.	**Third-variable problem** – another variable may be responsible for the relationship between the co-variables.
Useful when experiments are unethical or impractical – researchers can investigate relationships without manipulating potentially harmful variables.	**Can be misinterpreted** – a correlation may be incorrectly reported as showing **causation**, particularly by the media.
Secondary data can be used – researchers can analyse data that has already been collected, saving time and money.	A correlation **does not explain why** two variables are related.

CC BY NC ND

1.5. Content Analysis

A **content analysis** involves studying human behaviour **indirectly** by analysing things people produce, such as **TV programmes, adverts, newspapers, social media posts or films**.

Researchers can use content analysis to identify **patterns, themes, values, beliefs and attitudes** within the material being studied.

How to Conduct a Content Analysis

1. **Identify the aim or hypothesis** – decide what you want to investigate.
2. **Select the material** – choose the sources to be analysed, such as newspaper articles or TV adverts.
3. **Create a coding system** – decide which **categories** will be used to record the features being studied.
4. **Apply the coding system** – go through the material and record each relevant feature
5. **Analyse the data** – calculate frequencies or identify **themes and patterns**, depending on whether the coding produces quantitative or qualitative data.
6. **Write up the findings** – present the research using the format of a **scientific report**.

Coding in Content Analysis

Coding is the process of **turning information from the material being studied into categories** so that it can be analysed.

Researchers:

1. Decide which **categories** they are looking for.
2. **Code each occurrence** into the appropriate category.
3. **Count the frequencies** of each category.

For example, if a researcher is analysing **how aggression is shown in films**, they might create categories and **code each occurrence** into the most appropriate category:

Category	**What is coded**
Physical aggression	Hitting, kicking or pushing
Verbal aggression	Shouting, threatening or insulting
No aggression	No aggressive behaviour

The categories should be **clear, specific and mutually exclusive**, so that each occurrence fits into **only one category**.

CC BY NC ND

The researcher can then **count how often each category occurs**, converting the material into **quantitative data**, such as **frequencies and percentages**.

For example: If physical aggression occurs **20 times** and verbal aggression occurs **15 times**, the researcher can compare the **frequency** of the two types of aggression.

Strengths	Limitations
Real-world material – content analysis often uses naturally occurring material, such as films, adverts or social media posts, giving the research **high mundane realism**.	**Observer bias** – researchers may interpret or code material differently. This can be reduced by using **inter-observer reliability**.
Can produce a **large amount of data** without having to recruit participants.	**Researcher bias** – researchers may choose material or categories that support what they expect to find.
Easy to replicate if the same material and coding system are used.	**Interpretative bias** – qualitative material can be interpreted differently by different researchers.
Usually **fewer ethical issues** when analysing material already available in the **public domain**, as there is less need for informed consent.	**Limited insight** – analysing recorded material does not allow researchers to directly ask people about the reasons behind their behaviour.

CC BY NC ND

1.6. Case studies

A **case study** is a **detailed investigation of an individual, group or event**. It collects information from a range of sources to build a detailed picture of the case, often including **past and present behaviour**.

Case studies usually produce mainly **qualitative data**, although they can also include quantitative data.

Examples from psychology:

- **HM (2000)** – Researchers studied HM after his brain surgery and found that he could learn new **procedural skills**, such as mirror drawing, but had severe difficulties forming new **episodic memories**. This provided evidence that different types of long-term memory are **separate**.
- **Little Hans (1909)** – Freud used information about Little Hans' behaviour and experiences to support the **psychodynamic approach**, particularly the **Oedipus complex**.

Strengths	Limitations
In-depth information – researchers can gain a detailed understanding of a person's behaviour and experiences.	**Low generalisability** – studying one individual or a small group makes it difficult to apply findings to the wider population.
Allows researchers to investigate **rare or unusual cases** that would be difficult to study using other methods.	**Researcher bias** – researchers may selectively interpret information to support their expectations.
Can investigate situations that would be **unethical or impractical** to study experimentally.	**Retrospective data** may rely on people's memories, which can be **inaccurate**.
Can provide a **basis for further research** by generating ideas or hypotheses that can be tested using other methods.	**Time-consuming** and often difficult to **replicate**, particularly when the case is unique.

CC BY NC ND

Part 2 — Scientific Processes

2.1. Aims & Hypotheses

An **aim** is a **general statement** that explains the **purpose of a study**. It outlines **what the researcher intends to investigate**. Aims are often based on **psychological theories** or findings from **previous research**.

Example: *To investigate whether the amount of sleep a person gets affects their memory performance.*

Hypotheses

A **hypothesis** is a **precise, testable prediction** about the relationship between the variables being investigated.

The **alternative hypothesis (H_1)** predicts that there **will be a significant difference or relationship** between the variables being investigated.

There are two types of **alternative hypothesis (H_1)**:

Directional hypothesis (one-tailed)	**Non-directional hypothesis (two-tailed)**
Predicts the **direction** of the expected effect or relationship.	Predicts that there **will be an effect or relationship**, but **does not predict its direction**.
Used when previous research or psychological theory suggests the likely outcome.	**Used when** there is **little previous research** or the evidence is **contradictory**.
Example: *Participants who get **more sleep** will achieve **higher memory test scores** than participants who get less sleep.*	**Example:** *There will be a **difference** in memory test scores between participants who get different amounts of sleep.*

The **null hypothesis (H_0)** is the alternative to the alternative hypothesis. It predicts that there will be **no significant difference or relationship** between the variables.

CC BY NC ND

2.2. Sampling

The **population** is the **whole group of people** that the researcher is interested in studying.

A **sample** is a **smaller group of people selected from the population** to take part in the research.

Example: If a researcher wants to investigate the revision habits of **sixth-form students in the UK**, all sixth-form students in the UK are the **population**. The students actually selected to take part are the **sample**.

Researchers use different **sampling methods** to select their sample:

Sampling method	How it works	Strengths	Limitations
Opportunity	The researcher selects people who are **available and willing to take part** at the time the research is being conducted. For example, a researcher studying revision could approach students who are **in the school library** and ask them to participate.	**+ Quick,** easy and inexpensive.	**– Researcher bias,** the researcher selects people who are easiest to access, which may produce an unrepresentative sample.
Random	The researcher first creates a **sampling frame**, which is a list of everyone in the target population. Each person is given a **number**, and participants are then selected **entirely by chance**, using a random number generator, random number table or lottery method.	**+ Reduces researcher bias** because each person in the sampling frame has an **equal chance of being selected**.	**- Time-consuming** because a complete sampling frame is needed. **- Volunteer bias** may occur if selected people refuse to participate.
Systematic	The researcher creates a **sampling frame** and selects participants using a fixed numerical system. For example, they might select **every 10th person** on the list. A **random starting point** is usually chosen first to reduce bias.	**+ Reduces researcher bias** and is simple to carry out.	**- Can produce a biased sample** if there is a pattern in the sampling frame.

CC BY NC ND

Stratified	The researcher divides the **population into subgroups (strata)** based on relevant characteristics, such as **gender or age**. They calculate how many people are needed from each subgroup so the sample has the **same proportions as the population**. Participants are then selected **randomly from each stratum**.	**+ Produces a representative sample** because the proportions of relevant characteristics match those in the population	**- Time-consuming** because the researcher needs information about the population, must calculate the proportions and then select participants from each stratum.
Volunteer	The researcher **advertises the study** and people **choose for themselves** whether they want to participate. For example, a researcher could put an advert on a university noticeboard asking for volunteers for a memory study.	**+ Quick and easy to recruit.** Participants are generally **willing to take part** and may be more cooperative.	**- Volunteer bias** – people who choose to participate may have particular characteristics that make them different from the wider population.

Sampling Bias and Generalisation

Researchers usually cannot study **every member of the population**, so they select a **sample**. Ideally, the sample should be **representative of the population**, meaning the findings are more likely to **generalise to the wider population**.

However, the way participants are selected can introduce **sampling bias**. This can make the sample **unrepresentative**, limiting the **generalisability** of the findings.

Example: If a study on sixth-form students only samples students from **one London school**, the findings may not generalise to **all sixth-form students in the UK** because the sample may not represent the wider population.

2.3 Pilot Studies

A **pilot study** is a **small-scale version of an investigation carried out before the main study**. It is used to identify **problems with the procedure** so that they can be corrected before the full investigation.

For example, a researcher might pilot a questionnaire to check whether the questions are **clear and easy to understand**.

Piloting can help researchers:

- identify and fix **problems with the procedure**
- check that **instructions are clear**
- check that **equipment or materials work properly**
- estimate how long the investigation will take
- improve the **validity and reliability** of the final study
- save **time and money** by identifying problems before the main study

Example: Piloting a questionnaire

1. The researcher creates a questionnaire about **exam stress**.
2. They give it to a **small group of students** as a pilot study.
3. The students **complete the questionnaire**, allowing the researcher to see whether there are any problems.
4. **Problem identified:** Most students answer **"unsure"** to Question 3: *"How often do you feel psychologically overwhelmed by your academic responsibilities?"*
5. **Change made:** The researcher rewrites Question 3 to make it clearer: *"How often do you feel stressed because of your schoolwork?"*
6. The improved questionnaire is then used in the **main study.**

CC BY NC ND

2.4. Experimental Designs

Experimental design refers to **how participants are allocated to the different conditions of an experiment**. The three main designs are **independent groups, repeated measures and matched pairs**.

Design	How it works	Strengths	Limitations	Solution
Independent groups	Participants are **randomly allocated** to one condition of the IV, so each participant takes part in **only one condition**. **Different people in each condition**	**+ No order effects** because participants only complete one condition. **+ Requires fewer trials** per person than repeated measures.	**- Participant variables** may differ between groups and affect the DV. Requires more participants overall.	**Random allocation** gives each participant an equal chance of being placed in each condition, helping to balance participant variables.
Repeated measures	The **same participants take part in all conditions** of the IV. **Same people in every condition**	**+ Participant variables are controlled** because the same people take part in every condition. **+ Fewer participants are needed.**	**- Order effects** may occur, such as **fatigue, boredom or practice**, affecting performance in later conditions.	**Counterbalancing** : half of the participants complete the conditions in one order and the other half complete them in the **opposite order**.
Matched pairs	Participants are **matched in pairs** on important characteristics that could affect the DV. One person from each pair takes part in each condition. **Similar people split between conditions**	**+ Participant variables are controlled** because participants are matched. **+ No order effects** because each participant completes only one condition.	**- Time-consuming** to find and match participants. **- Difficult to match all relevant variables** that could affect the DV.	**Careful matching** on variables known to affect the DV.

CC BY NC ND

2.5. Observational Design

When conducting an observation, researchers need to decide **exactly what behaviour they are looking for and how they will record it**. This helps make observations **systematic and easier to analyse**.

Behavioural Categories

A **behavioural category** is a **specific, observable and measurable behaviour** that the researcher records during an observation. Researchers break broad **target behaviour** into smaller, clearly defined categories.

Example: If the target behaviour is **aggression**, behavioural categories could include:

- **Shouting**
- **Punching**
- **Swearing**

The categories should be **operationalised**, meaning it is clear exactly what counts as each behaviour.

They should also be **mutually exclusive**, meaning each behaviour should fit into **only one category**. For example, **smiling** and **grinning** would be too similar to use as separate categories.

Clear behavioural categories make it easier for different observers to **record the same behaviours consistently**, improving **inter-observer reliability**.

Sampling Methods

Researchers cannot always record **every behaviour that occurs**, especially during a long observation. Sampling methods allow them to decide **when or what behaviour to record**.

Sampling method	How it works	Strengths	Limitations
Time sampling	The researcher records behaviour during **specific time periods** decided before the observation. For example, recording behaviour for **2 minutes every 10 minutes**.	**+ Less time-consuming** as it reduces the amount of behaviour that needs to be recorded.	**- The selected time periods may not be representative** of the person's behaviour throughout the whole observation.

CC BY NC ND

Event sampling	The researcher records **every occurrence of a particular behaviour** during the observation.	**+ Useful for infrequent behaviours,** as the researcher waits for the specific behaviour to occur.	**- Difficult to record every occurrence accurately** especially if the behaviour occurs very frequently. **- Complex behaviours may also be oversimplified.**

2.6. Questionnaire Construction

When designing a **questionnaire**, questions must be carefully written so participants **understand what is being asked** and can give accurate answers.

Principle	What it means	Example
Clear and unambiguous	Questions should have **one clear meaning** and be easy to understand.	*"How often do you feel anxious?"* rather than *"Do you regularly experience anxiety?"*
Avoid jargon	Avoid **technical or unfamiliar language** that participants may not understand.	Avoid *"Do you experience psychomotor agitation?"* with the general public.
Avoid leading questions	Questions should not **encourage a particular answer**.	Avoid *"Don't you agree that school causes stress?"*
Avoid double-barrelled questions	Do not ask **two things in one question**, as participants may have different answers to each part.	Avoid *"Do you enjoy school and get on with your teachers?"*
Avoid double negatives	Double negatives can **confuse participants** and lead to inaccurate answers.	Avoid *"Do you disagree that you don't enjoy school?"*
Use neutral wording	Questions should not suggest that one answer is **better or more acceptable**.	Avoid *"How helpful was the excellent therapy?"*

Response options should also be carefully designed. For closed questions, answer choices should be **clear, appropriate and non-overlapping** so participants can easily select the answer that best applies to them.

Good	Bad
How many hours do you revise each week?	*How many hours do you revise each week?*
0–2 hours 3–5 hours 6–8 hours 9–11 hours 12+ hours	0–5 hours 5–10 hours 10–15 hours 15+ hours
No overlap between options.	**5 and 10 hours overlap**, so participants may be unsure which option to choose.

CC BY NC ND

Design of Interviews

When designing an interview, researchers need to decide **what questions will be asked, how they will be asked and how the interview will be structured**. The design should make sure participants understand the questions and that the information collected is **relevant to the aim of the study**.

Design consideration	Structured interview	Unstructured interview
Preparing questions	A **fixed set of questions** is prepared before the interview.	The researcher prepares **general topics or areas** they want to investigate rather than a fixed list of questions.
Wording	Questions should use **standardised wording** so every participant is asked the same question.	Questions can be **adapted to the participant's responses** and phrased naturally.
Order	Questions are asked in a **pre-determined order**.	There is **no fixed order**; the conversation can develop naturally.
Follow-up questions	Follow-up questions are generally **not added**, as this would make interviews less standardised.	The researcher can ask **follow-up questions** to explore interesting or relevant answers in more detail.
Interviewer	The interviewer should follow the **same procedure** with each participant.	The interviewer needs to be able to **respond to and explore** the participant's answers appropriately.
Recording responses	Responses can be recorded in a **standardised way**, making them easier to compare.	Responses are likely to be **more varied and detailed**, so the researcher needs to decide how these will be recorded and analysed.

Piloting

Both **questionnaires and interviews can be piloted** before the main study. A small group can be used to identify **problems with the questions, wording or procedure**, allowing the researcher to make changes before collecting the main data. *See notes in section 2.3 for examples.*

CC BY NC ND

2.7. Variables

Variables are **anything that can change** in a study. Researchers identify and control variables to make sure the results are **valid**.

Variable	What it means	Sleep and memory example
Independent variable (IV)	The variable **manipulated by the researcher**.	**Hours of sleep:** 4 hours vs 8 hours
Dependent variable (DV)	The variable is measured to see whether it is affected by the IV.	**Memory test score:** marks out of 20
Extraneous variable	Any variable **other than the IV that could affect the DV**.	**Time of day**, test difficulty, caffeine intake

In **natural and quasi-experiments**, the IV is **not manipulated** by the researcher because it already exists naturally.

Extraneous Variables

An **extraneous variable** is **any variable other than the IV that could affect the DV**. Researchers try to **control extraneous variables** so they can be more confident that any change in the DV was caused by the IV.

Example: In a study investigating the effect of sleep on memory, researchers might control:

Extraneous variable	How it could affect the study
Test difficulty	A harder test could lower memory scores.
Time of day	Participants may perform better at certain times.
Caffeine intake	Caffeine could affect alertness and memory.

Operationalisation of Variables

Operationalisation means **clearly defining exactly how a variable will be measured or manipulated** so that it is **specific, observable and measurable**.

For example: Instead of simply saying *"Sleep affects memory"*, the researcher could operationalise the variables by stating:

- **IV:** Amount of sleep → **8 hours vs 4 hours**
- **DV:** Memory performance → **score out of 20 on a word recall test**

This could be written as a directional hypothesis:

> *Participants who receive **8 hours of sleep** will score **more marks out of 20 on a word recall test** than participants who receive **4 hours of sleep**.*

Why operationalise variables?

- **Replicability** – other researchers can repeat the study using the same procedures.
- **Consistency** – variables are measured or manipulated in the **same way** for all participants.
- **Testability** – the hypothesis becomes **clear and measurable**.

CC BY NC ND

2.8. Control

Researchers can use several methods to **control extraneous variables** and reduce their effects on the results.

Method	What it means	How it helps	Example
Randomisation	The **use of chance** when making decisions during a study to reduce **bias**. It can be used when deciding the **order of conditions**, arranging **materials or stimuli**, or making other decisions within the procedure.	Using chance reduces the likelihood that the **researcher's preferences or expectations** influence the study.	A researcher uses a **random number generator** to decide which condition each participant completes first.
Standardisation	Using the **same procedures and instructions** for every participant.	This controls differences in how the study is carried out, making it less likely that these differences will affect the **DV**.	Every participant receives the **same instructions**, completes the same task and has the **same amount of time** to complete it.
Random allocation	Assigning participants to different conditions **by chance**, giving each participant an **equal chance of being placed in each condition**.	This helps distribute **participant variables** between the groups, making the groups more comparable.	A **random number generator** assigns participants to either the **experimental or control condition**.
Counterbalancing	Used in **repeated measures designs**. Different participants complete the conditions in **different orders**.	Both groups complete **both conditions**, but reversing the order means that practice or fatigue is less likely to affect one condition more than the other.	**Group 1:** Condition A → **Condition B** **Group 2:** Condition B → **Condition A**

CC BY NC ND

Control groups	An **experimental group** is a group that **experiences the IV being investigated**. A **control group** does **not experience the IV** or receives a **standard comparison condition**. A **control group** provides a **baseline for comparison** with the experimental group.	Comparing the groups helps researchers determine whether a change in the **DV** is likely to have been caused by the **IV** rather than other factors.	**Caffeine and memory:** Experimental group receives a **caffeinated drink**; The control group receives a **decaffeinated drink**.

2.9. Demand Characteristics and Investigator Effects

Demand Characteristics

Demand characteristics are **cues from the researcher or research situation that allow participants to guess the aim of the investigation**. This may cause participants to **change their behaviour** rather than behaving naturally.

When participants enter an unfamiliar research situation, they may try to work out **what is expected of them**. This is known as **participant reactivity**.

Example: A participant realises that a study is investigating whether **stress affects performance** and deliberately tries harder because they think this is what the researcher wants.

Participants may respond in different ways:

- **Please-U effect** → the participant behaves in the way they think the **researcher wants** them to.
- **Screw-U effect** → the participant deliberately behaves in a way that **works against the aims of the study**, such as deliberately underperforming.

Effect on research: Demand characteristics can reduce the **validity** of findings because participants may not behave naturally. The results may reflect their **reaction to the research situation** rather than the effect of the **IV**.

Investigator Effects

Investigator effects are **unwanted influences of the researcher on participants or the results of an investigation**, whether the influence is **conscious or unconscious**.

The researcher can influence the study through their:

- **Behaviour**
- **Appearance**
- **Expectations**
- **Interactions with participants**

They can also influence the study through decisions made when **designing or carrying out the research**.

Example: If a researcher expects participants in the experimental condition to perform better, they may unintentionally give them **more encouragement or clearer instructions**.

2.10. Ethics

Ethics are the principles that guide researchers to make sure psychological research is **fair, safe and respectful to participants**.

The **British Psychological Society (BPS) Code of Ethics and Conduct** gives psychologists guidelines for carrying out research responsibly. Researchers should consider **ethical issues when designing and carrying out a study** and take steps to deal with any problems that could affect participants.

Ethical issue	What it means	How it can be dealt with
Informed consent	Participants should agree to take part after receiving enough information to make an **informed decision**. However, researchers may not reveal the true aim because this could cause **demand characteristics**.	**Prior general consent:** Participants agree to take part in a range of studies, some of which may involve deception. **Presumptive consent:** A similar group is asked whether they would consent, and this is used to judge whether the actual participants would agree. **Retrospective consent:** Participants are asked for consent **after** taking part and can decide whether their data can be used.
Deception	Participants are **deliberately misled** or information is withheld. It may be necessary if knowing the true aim would cause demand characteristics, but it should not expose participants to unacceptable harm.	**Debriefing:** After the study, participants are told the **true aim and nature** of the research. They should have the opportunity to ask questions and can request that their data is withdrawn where possible.
Protection from harm	Participants should be protected from **physical and psychological harm**. The risk should not be greater than what they would normally experience in everyday life.	Researchers should assess potential risks before the study and stop the study if participants become distressed. An **ethical review** should consider whether the benefits of the research justify the potential risks.

CC BY NC ND

Privacy	Participants have the right to control information about themselves and how much of their **private life** is observed. This can be difficult in **naturalistic or covert observations**.	Researchers should avoid observing private behaviour without good justification and should consider whether participants could reasonably expect to be observed in that situation.
Confidentiality	Participants have the right to have their **personal information protected** and not disclosed without permission.	Researchers can remove identifying information, use **numbers or initials instead of names**, and store data securely so individuals cannot easily be identified.
Right to withdraw	Participants should be able to **leave the study at any point** without giving a reason.	Participants should be informed of their right to withdraw during the **briefing** and reminded again during the **debriefing**. Where possible, they should also be able to request the removal of their data.

CC BY NC ND

2.11. The Role of Peer Review

Peer review is the **assessment of scientific research by experts in the same field**. It is used to check that research is **good quality, valid and suitable for publication**.

Before research is published in an academic journal, it is usually assessed by **other psychologists or experts who have knowledge of the same area of research**. They examine the research to identify **strengths, weaknesses, errors or problems** with the methodology, analysis and conclusions.

The reviewers then provide **feedback to the researcher** and may recommend that the research is:

- **Accepted** for publication
- **Revised** before it can be published
- **Rejected** if there are serious problems with the research

Peer review therefore acts as a **quality-control process** within psychology. It helps ensure that research is **checked by other experts before it becomes part of the scientific evidence base**.

Main purposes of peer review

- To decide which research is **worthwhile**, helping organisations decide where **research funding** should be allocated.
- To **check the quality and validity of research** before it is published.
- To identify possible **errors, weaknesses or improvements** that could be made to the research.
- To help prevent **fraudulent or poor-quality research** from being published.

AO3 - EVALUATION

1. **It improves the quality of published research.** Research is checked by **experts in the same field** before publication. Reviewers can identify weaknesses in the **methodology, analysis or conclusions** and suggest improvements. This means researchers can correct problems before their findings become part of the scientific literature, increasing the **quality and credibility of psychological research**.

2. **Peer review can be affected by reviewer bias.** Reviewers are usually anonymous, which can protect them from pressure, but this can also allow them to **disadvantage researchers they disagree with or compete against**. This may be particularly problematic when researchers are competing for **research funding**. One solution is **open reviewing**, where the reviewer's identity is known. This may make reviewers more accountable and reduce the opportunity for personal bias.

3. **Publication bias can distort the evidence base.** Journals may be more likely to publish research with **interesting, significant or positive findings** because these are more likely to attract attention. Research that finds no significant effect may be less likely to be published, creating the **file-drawer problem**, where studies with

CC BY NC ND

negative or non-significant findings remain unpublished. This means psychologists may get a **misleading picture of the evidence** because they are seeing mainly the studies that produced interesting results.

4. **Peer review may discourage new ideas.** Reviewers are often experts who are familiar with **established theories and accepted explanations**. They may therefore be more likely to favour research that supports existing ideas and reject research that challenges them. This could **slow scientific progress**, as new or controversial explanations may take longer to become accepted.

5. **Peer review cannot completely prevent fraudulent research.** Even research that has been peer reviewed can contain serious problems. **Wakefield (1998)** published research claiming a link between the **MMR vaccine and autism**, which was later found to have serious ethical and methodological problems and was eventually retracted. This shows that peer review is useful for **quality control**, but it cannot guarantee that all published research is accurate or honest.

CC BY NC ND

2.12. The Implications of Psychological Research for the Economy

Economic implications refers to how psychological research can affect a country's **financial wellbeing**, particularly through **employment, productivity, healthcare costs and the use of public money**.

Psychological research can benefit the economy when it helps people to **stay healthy, return to work and work more effectively**. For example, research can lead to more effective treatments for mental health problems, reducing **sickness absence** and allowing people to remain economically active. Research can also help organisations and governments use **resources more efficiently**, reducing unnecessary costs.

Examples across the course:

Topic	Psychological research	Economic implication
Clinical Psychology	**Soomro et al. (2009)** reviewed drug treatments for OCD and found that **SSRIs were more effective than placebo** in reducing OCD symptoms.	Effective drug treatment can reduce the severity of OCD, allowing people to **function more effectively and return to employment**. This can reduce the economic cost of **sickness absence** and reduce the need for long-term healthcare support.
Biopsychology	Research into **neurotransmitters** has shown that abnormalities in systems such as **serotonin and dopamine** are associated with psychological disorders. This has contributed to the development of drug treatments that alter neurotransmitter activity.	Effective medication can **reduce symptoms and hospitalisation**, meaning fewer resources need to be spent on long-term care. If patients can return to work, this also **increases productivity and reduces sickness absence**.
Memory	**Geiselman et al. (1984)** found that the **cognitive interview** produced more correct information from eyewitnesses than a standard police interview.	More accurate eyewitness evidence can make criminal investigations **more efficient**, reducing the time and resources spent investigating incorrect leads and potentially reducing the costs associated with **wrongful convictions**.

Social Influence	Research into **minority influence** shows that consistent minorities can influence the majority and contribute to **social change**.	This can be applied to campaigns designed to change behaviour, such as **anti-smoking or environmental campaigns**. Successful behaviour change can reduce costs to services such as the **NHS** and improve productivity.
Clinical Psychology	Research into **CBT** has demonstrated that psychological treatments can reduce symptoms of disorders such as depression.	Treating depression can reduce **sickness absence and unemployment**, allowing people to return to work and contribute to the economy rather than requiring continued healthcare or welfare support.

2.13. Reliability

Reliability refers to the **consistency of findings** from a psychological investigation. If a method is reliable, it should produce **similar results when repeated**.

Reliability is important because researchers need to know that their findings are **consistent** and not simply caused by chance, unclear questions or differences in how behaviour was measured.

Types of reliability and how they are assessed:

Type	What it means	Ways of assessing it
Internal reliability	The **consistency within a measure**. It assesses whether different parts of the same test are consistently measuring the **same construct**.	**Split-half method:** the items in a test are divided into **two halves** (e.g. odd and even questions). Each participant receives a score for each half, and the two scores are **correlated**. A strong positive correlation suggests that both halves are measuring the same thing consistently. A correlation of around **+0.80 or above** is generally considered high reliability.
External reliability	The **consistency of a measure outside the original testing occasion**. A reliable measure should produce similar results when used at a different time or by different observers.	**Test-retest:** the **same test is given to the same participants on two separate occasions** and the two sets of scores are correlated. There should be an appropriate time interval: too short and participants may remember their previous answers; too long and their characteristics or attitudes may genuinely change. **Inter-observer:** two or more observers **independently record the same behaviour** using the same behavioural categories. Their observations are compared to calculate the level of agreement. $\frac{\text{Total number of agreements}}{\text{Total number of observations}} \geq 0.80$ = High Inter-observer Reliability A high correlation/agreement (around **+0.80 or above**) indicates high inter-observer reliability.

Improving reliability

Questionnaires

- Use **clear and unambiguous questions** to reduce differences in interpretation.
- Use appropriate **closed questions** with clear response options.
- If test-retest reliability is low, unclear questions can be **rewritten or removed**.

Interviews

- **Standardise** the interview as much as possible by using the same questions and instructions.
- Interviewers should be **trained** so they conduct the interview consistently and avoid leading or ambiguous questions.
- **Structured interviews** are generally more reliable because the questions and procedure are standardised.

Experiments

- Use **standardised instructions, materials and procedures** for every participant.
- Keep testing conditions as similar as possible.
- This makes the procedure easier for another researcher to **replicate**, allowing the consistency of findings to be assessed.

Observations

- Use **operationalised behavioural categories** that are observable and measurable.
- Categories should be **mutually exclusive** so behaviour cannot fit into two categories.
- Categories should be **exhaustive**, covering all relevant behaviours.
- Train observers to use the categories consistently and calculate **inter-observer reliability** to check their agreement.

CC BY NC ND

2.14. Validity

Validity is about whether the findings of a study are **true and accurate**, whether the research method actually measures what it is supposed to measure and gives a **true representation of what is being investigated**.

Reliability and validity are different. A study can be reliable without being valid. For example, a questionnaire could produce **consistent results** each time but still fail to measure what it claims to measure.

There are two main **umbrella types of validity:**

- **Internal validity** concerns whether the findings are **true within the study itself**.
- **External validity** concerns whether the findings can be **generalised beyond the study**.

Umbrella type	Specific types	Ways of assessing it
Internal validity – the extent to which the results are genuinely caused by the **IV**, rather than another factor. **High internal validity** means we can be confident that changes in the **DV were caused by the IV**, rather than by **extraneous variables, participant variables, demand characteristics or investigator effects**.	**Face validity** – whether a measure **appears** to measure what it is supposed to measure. **Concurrent validity** – whether a new measure produces similar results to an **established measure** of the same construct.	**Face validity:** ask participants, researchers or experts whether the measure appears appropriate. **Concurrent validity:** compare the new measure with an established measure and **correlate the scores**.
External validity – the extent to which findings can be **generalised beyond the original study**. This includes whether findings apply to different **situations, populations and time periods**.	**Ecological validity** – whether findings generalise to **real-life situations**. **Temporal validity** – whether findings remain valid across **different time periods**.	**Ecological validity:** compare the research setting/tasks with real-life situations. **Temporal validity:** see whether findings are replicated at different points in time.

CC BY NC ND

Improving validity

Experiments

- **Standardisation** – using the same instructions, materials and procedure for every participant reduces extraneous variables.
- **Random allocation** – reduces the effect of participant variables by giving each participant an equal chance of being in each condition.
- **Control groups** – provide a baseline for comparison, making it easier to see whether the IV caused changes in the DV.
- **Single/double-blind procedures** – reduce **demand characteristics** and **investigator effects**.

Questionnaires and interviews

- Use **clear, unambiguous questions** to make sure questions measure what they are intended to measure.
- Use **validated measures** that have already been shown to measure the intended construct.
- **Anonymity** can reduce social desirability bias, making answers more likely to reflect participants' true views.

Observations

- Use clear, **operationalised behavioural categories** so observers know exactly what behaviour to record.
- Use **inter-observer reliability** by having two or more observers independently record behaviour and compare their results.
- **Naturalistic observations** can improve ecological validity because behaviour is studied in a real-life setting.

CC BY NC ND

2.15. Features of Science

For research to be considered **scientific**, it must follow certain principles that allow researchers to produce **objective, testable and trustworthy evidence**.

Feature	Explanation	Example / Key point
Objectivity	**Researcher bias is minimised**, so personal beliefs, expectations or opinions do not influence the research process or findings.	Researchers should base conclusions on **evidence**, rather than personal opinions or expectations.
Empirical method	Scientific evidence is collected through **systematic observation and measurement**. Scientific claims should therefore be based on **observable and measurable evidence**.	Research should use **evidence that can be observed and measured**, rather than speculation. The **behaviourist approach** measures observable behaviour, while the **biological approach** uses methods such as **brain scans and drug experiments** to collect measurable evidence.
Replicability	A study should be able to be **repeated by other researchers using the same methods** and produce similar findings.	Similar results increase confidence that the findings are **reliable** and not simply due to **chance, researcher bias or the specific circumstances** of the original study.
Falsifiability	A scientific theory must make predictions that could potentially be **shown to be false**. **Popper (1934)** argued that this is an important feature of science.	A theory can be **tested and potentially rejected** if evidence contradicts its predictions. Theories cannot be proven permanently true; they can only gain greater **support** when they survive attempts to falsify them. The **psychodynamic approach** can be criticised because concepts such as the **unconscious mind** are difficult to test or disprove, making them **unfalsifiable**.

CC BY NC ND

Theory construction & hypothesis testing	A **theory** is a set of general principles used to **explain observations and predict behaviour or events**. **Theory construction** uses research evidence to develop explanations, while **deduction** uses an existing theory to develop a **specific, testable hypothesis**.	If findings consistently support a theory's predictions, confidence in the theory increases. If findings contradict them, the theory may be **revised or rejected**.
Paradigms	A **paradigm** is a set of **shared ideas, assumptions and ways of working** accepted by scientists within a discipline.	**Kuhn (1962)** argued that mature sciences have a **dominant paradigm**. Psychology has 5 competing approaches, so Kuhn described it as **pre-paradigmatic**.
Paradigm shifts	A **paradigm shift** occurs when there is a **major change in accepted scientific assumptions**, usually because problems with the existing paradigm lead to a **scientific revolution**.	Psychology's competing approaches can be used to argue that it does not yet have one dominant paradigm and is therefore **not a fully established science**.

CC BY NC ND

2.16. Reporting Psychological Investigations

Psychologists use a particular format to write up research for publication, known as a **scientific report**. A scientific report follows a standard structure so that research can be **clearly communicated, evaluated and replicated**.

Sections of a Scientific report

- **Abstract** – A brief summary of the **key details of the investigation**, including the **aim, hypothesis, method, main results and conclusion**. It allows readers to quickly understand what the study investigated and decide whether they want to read the full report. **It may contain a brief sentence summarising each section of the report.**
- **Introduction** – Provides the **background to the research**. It discusses relevant **theories, concepts and previous research**. The introduction usually starts broadly before becoming more specific and ends by stating the **aim and hypothesis/hypotheses** of the investigation.
- **Method** – Provides a detailed description of **exactly how the investigation was carried out**. It includes the **research design, sample, materials, procedure and ethical considerations**. There should be enough detail for another researcher to **replicate the investigation**.
- **Results** – Presents the **findings of the investigation** without explaining what they mean. This can include **descriptive and inferential statistics**, tables, graphs and other appropriate forms of data presentation. For qualitative research, this may include the results of **thematic analysis**.
- **Discussion** – Explains and interprets **what the findings mean**. The researcher may compare the findings with previous research and theories discussed in the introduction. The discussion can also consider **conclusions, limitations, improvements and implications** of the findings.
- **Referencing** – Provides details of the **sources referred to in the report**, allowing readers to locate the original material. Sources may include **journal articles, books and websites**. References should be presented using an appropriate referencing system and in a consistent format.

CC BY NC ND

Referencing books and journal articles

Researchers must **reference their sources** to give credit to the original authors and show where their information and ideas came from. This allows readers to **check the original source** and helps prevent **plagiarism**, which is presenting someone else's work or ideas as your own.

Different types of sources are **referenced in different ways**. For example, the information included in a **book** is referenced differently from a **journal article** or a **website**.

- **Book: Author(s) → date → title of book → place of publication → publisher**

Example: Flanagan, C. and Berry, D. (2016). *A Level Psychology Year 2*. Cheltenham: Illuminate Publishing.

- **Journal article: Author(s) → date → title of article → journal name → volume → issue → page range**

Example: Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*(4), 371–378.

CC BY NC ND

3. Data Handling and Analysis

3.1. Quantitative and Qualitative data

Type and description	Data collection techniques	Strengths	Limitations
Qualitative data – data presented in **words or other non-numerical forms**, such as participants' opinions, experiences and descriptions.	**Open questions**, **unstructured interviews** and observations where detailed descriptions of behaviour are recorded.	**+ Rich and detailed data** – provides a deeper understanding of participants' **thoughts, feelings and experiences**. **+ Validity** – participants can explain their views in their own words, giving a more accurate representation of what they think or feel. **+ Unexpected findings can emerge** – researchers may discover new themes or ideas that were not anticipated before the study.	**– Time-consuming to analyse** – responses must be organised and interpreted, making analysis slower than quantitative data. **– Difficult to compare responses** – participants may give different answers, making it harder to identify patterns or make direct comparisons. **– Researcher bias** – analysing and interpreting qualitative data can be **subjective**, so researchers may interpret the same responses differently.
Quantitative data – data presented in **numerical form**, such as scores, frequencies, measurements and percentages.	**Closed questions**, **structured observations** using behavioural categories, and numerical measurements in **experiments**.	**+ Can be analysed using statistical tests** and presented using **graphs and tables**. **+ Easy to compare and identify patterns** **+ More objective** – numerical data is generally **less open to interpretation**,	**– Less depth and detail** – numerical data may not provide a detailed understanding of participants' **thoughts, feelings and experiences**. **– Does not explain why** – numerical data shows **what** participants responded to, but may

CC BY NC ND

		reducing researcher bias.	not explain the reasons behind their responses. **– Can oversimplify complex behaviour** – reducing thoughts, feelings or behaviours to **numbers** may leave out important detail and context.

CC BY NC ND

3.2. Primary and Secondary data

Type and description	Strengths	Limitations
Primary data – information **collected first-hand by the researcher** specifically for their investigation.	**+ Can be tailored to the research aims** – researchers can collect the **exact information needed** for the investigation. **+ Greater control over data collection** – the researcher can choose **how, when and under what conditions** the data is collected.	**– Time-consuming** – collecting primary data can take a significant amount of **time and effort**. **– Expensive** – collecting large amounts of data or **recruiting participants** can be costly.
Secondary data – information that has been **collected by someone else** but is used by a researcher for their own investigation. It is sometimes referred to as **desk research**.	**+ Quick and inexpensive** – the researcher does not have to **collect the data themselves**, saving time and money. **+ Access to large amounts of existing data** – researchers can use data that has already been collected, allowing them to study **large samples or datasets**.	**– May be outdated or incomplete** – existing data may no longer accurately reflect the population being studied. **– May not meet the research aims** – the data was collected for a different purpose, so it may not contain **exactly the information the researcher needs**. **– Less control over data quality and collection**
Meta-analysis – a method of analysing **secondary data** by statistically combining findings from **multiple studies investigating the same topic** to produce an **overall conclusion**.	**+ Combines evidence from many studies** – provides a **larger overall evidence base** than an individual study. **+ Increases confidence in findings** when results are **consistent across studies**, as similar findings are less likely to be due to chance. **+ Identifies overall patterns** – combining results can reveal **patterns or trends** that may not be clear in individual studies.	**– Publication bias** – studies with **non-significant or negative findings** are less likely to be published (**file-drawer problem**). This can make the overall evidence appear **more positive than it really is**. **– Depends on the quality and comparability of included studies** – if the studies have **poor methodology** or use very different methods/measures, the conclusions of the meta-analysis may be unreliable.

CC BY NC ND

3.3. Descriptive Statistics

Descriptive statistics are numerical methods used to **summarise and describe data**. They include measures of **central tendency**, measures of **dispersion**, percentages and correlations. Data can also be presented visually using **tables and graphs**.

Measures of central tendency

Measures of central tendency describe the **typical or central value** within a set of data.

Measure	Calculation method	Strengths	Limitations
Mean – the arithmetic average of a set of scores. - Useful for **interval-level data.**	**Add all the scores together and divide by the number of scores.** **For example:** **Scores -4, 6, 7, 8, 10** **Mean = (4 + 6 + 7 + 8 + 10) ÷ 5 = 7**	**+ Uses all values** in the dataset. **+ Precise measure** of the average.	**- Affected by extreme scores (outliers)**, which can make the mean unrepresentative of the typical score.
Median – the **middle score** when data is arranged in order from lowest to highest. - Useful for **ordinal data**.	Arrange scores in order and identify the **middle value**. If there are two middle values, calculate their mean.	**+ Not affected by extreme scores.** **+ Easy to calculate and interpret.**	**- Does not take the exact value** of every score into account **- Can be less sensitive** than the mean.
Mode – the score that occurs **most frequently**. - Useful for **nominal data**.	Count how often each score occurs and identify the **most frequent score**.	**+ Easy to identify.**	**– There may be more than one mode.** **– Some datasets may have no mode.** **– Does not use all the data.**

CC BY NC ND

Measures of dispersion

Measures of dispersion describe **how spread out or varied the scores are** within a dataset.

Measure	Calculation method	Strengths	Limitations
Range – the difference between the **highest and lowest scores**.	**Highest score − lowest score**	**+ Quick and easy to calculate.** **+ Gives an indication of how spread out the data is**, showing the difference between the **highest and lowest scores**	**- Only uses the highest and lowest scores.** **- Strongly affected by extreme values.**
Standard deviation (SD) – a measure of how much scores **vary around the mean**.	Measures how far each score is from the **mean**. The differences are squared, added together, divided by the number of scores, and the **square root** is taken.	**+ Uses all scores** in the dataset. **+ Useful for comparing the spread of different datasets**, as a smaller standard deviation indicates scores are more closely clustered around the mean. **+ Provides a more precise measure of dispersion** because it considers how **all scores vary around the mean**.	**– More difficult to calculate and interpret than the range.** **– Can be affected by extreme scores**

i. Calculating Standard Deviation

$$s = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$$

Where:

- s = standard deviation
- **x** = each individual score
- **x̄** = mean
- **n** = number of scores

CC BY NC ND

Steps:

1. **Calculate the mean ($\bar{x}$)** of all the scores.
2. **Subtract the mean from each score** to find how far each score is from the mean.
3. **Square each difference** to remove negative values.
4. **Add all the squared differences together**.
5. **Divide by the number of scores minus 1 (n-1)**.
6. **Take the square root** of the answer.

Example

Scores: **2, 4, 6, 8, 10**

1. **Mean:** (2 + 4 + 6 + 8 + 10) ÷ 5 = **6**
2. **Subtract mean from each score:** 2 − 6 = **−4**, 4 − 6 = **−2**, 6 − 6 = **0**, 8 − 6 = **2**, 10 − 6 = **4**
3. **Square each answer:** 16, 4, 0, 4, 16
4. **Add:** 16 + 4 + 0 + 4 + 16 = **40**
5. **Divide by (n-1):** 40 ÷ (5-1) = **10**
6. **Square root:** √10 = **3.16**

Standard deviation = 3.16

ii. Calculating percentages

A **percentage** expresses a value as a **proportion out of 100**.

To calculate a percentage:

- **Percentage = (number in the category ÷ total number) × 100**
- For example, if **30 out of 50 students** preferred online learning:
- **(30 ÷ 50) × 100 = 60%**

CC BY NC ND

3.4. Presentation and Display of Quantitative data

Quantitative data can be presented in different ways to make the **patterns, differences and relationships** in the data easier to understand.

Tables

Tables are used to **organise and summarise numerical data** so that results can be easily compared.

Tables can be used to present **nominal, ordinal or interval data**, depending on what is being measured. They are particularly useful for presenting **summary statistics**, such as the mean, median, mode, frequency and percentages, so that results can be easily compared.

A brief description may be provided alongside or below the table to **highlight the main pattern in the data**.

Example:

Descriptive statistic	4 hours of sleep (Condition A)	8 hours of sleep (Condition B)
Mean	12	17
Median	12	17
Mode	11	18

The table shows that participants who slept for **8 hours generally had higher memory scores** than those who slept for **4 hours**.

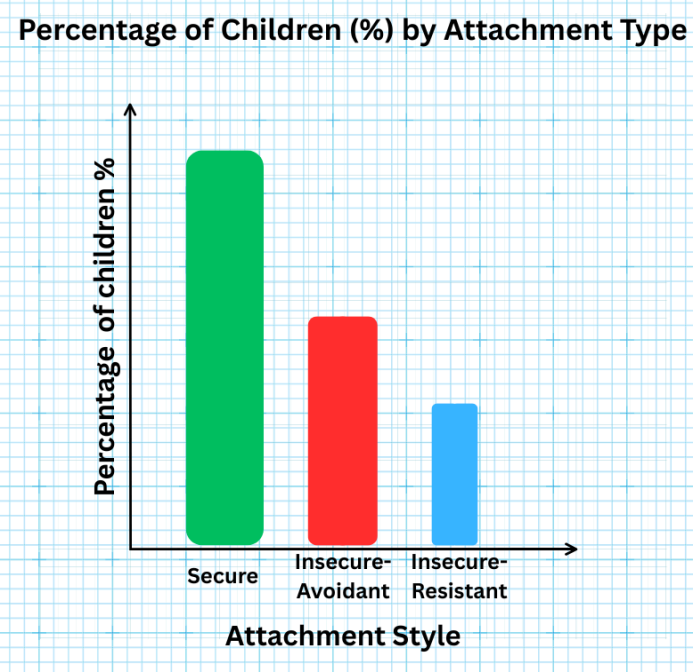
CC BY NC ND

Bar charts

Bar charts are used to display **discrete data**, where scores belong to **separate categories**. This includes **nominal data**, where categories have no meaningful order, such as **type of attachment**. They can also be used to display **ordinal categories**.

The **x-axis** shows the **categories or conditions**, while the **y-axis** shows a **numerical measure**, such as the **number of participants, frequency, percentage or mean score**. The **height of each bar** represents the value for that category, making it easy to **compare different groups or conditions**.

The bars **do not touch** because the categories are **separate from one another**. This distinguishes a bar chart from a **histogram**, where the bars touch because the data represents continuous measurements.



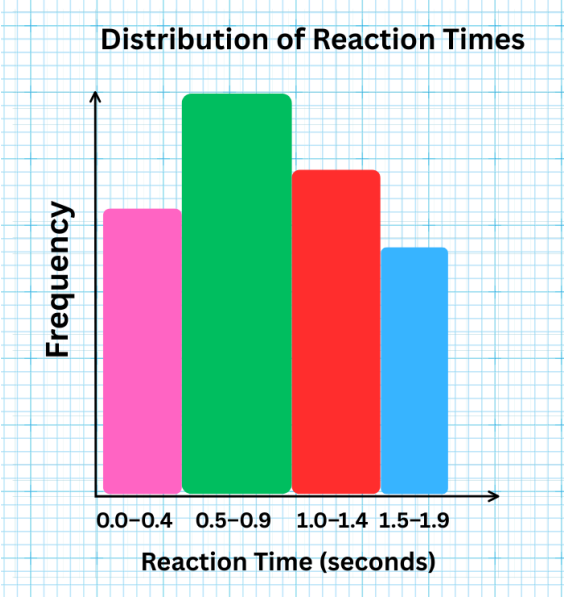
Histograms

Histograms are used to display **continuous data**, where there can be **many possible values between two points**. For example, **height** can be 160 cm, 160.5 cm, 160.51 cm, and so on.

This is usually **interval data**, where the **differences between scores are equal and meaningful**.

The **x-axis** shows the continuous data divided into **intervals (class intervals)**, while the **y-axis** shows the **frequency**, which is the **number of scores that fall within each interval**.

The bars **touch each other** because the intervals are **continuous**, meaning there are **no gaps between the values**.



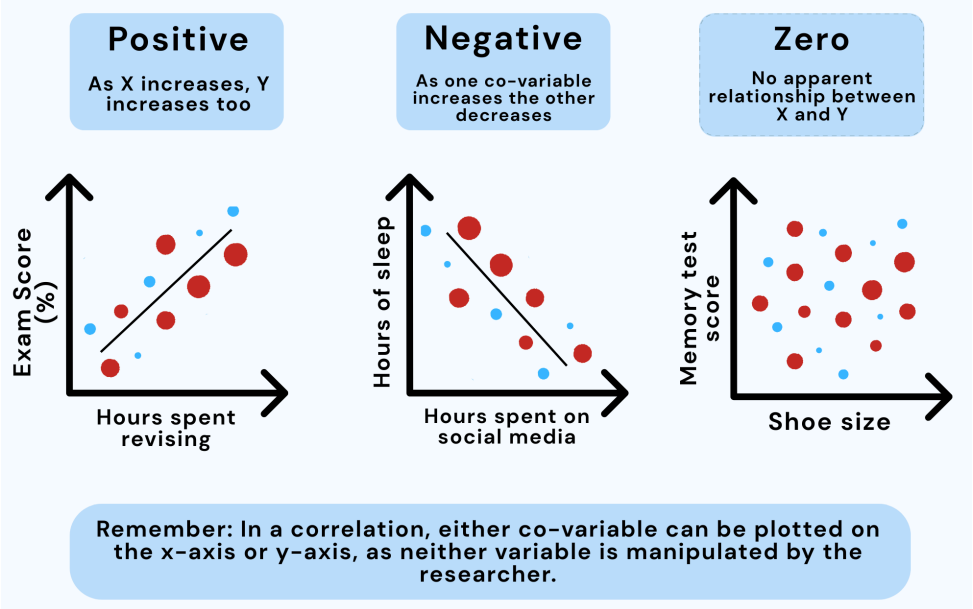
CC BY NC ND

Scattergrams

Scattergrams are used to show the **relationship between two co-variables**, rather than differences between conditions. The co-variables can be measured using **ordinal or interval data**.

Each **point** represents the **paired scores of one participant** on the two co-variables. For example, one participant's **hours of sleep** and **memory test score** would be represented by one point. Either co-variable can be placed on either axis because **neither variable is being manipulated as an IV**. The **x-axis** shows one co-variable and the **y-axis** shows the other.

A scattergram can show a **positive, negative or zero correlation**, depending on the pattern of the points.



Revision Tip: Exam questions may ask you to **draw an appropriate diagram** to display quantitative data. These can be worth **6 marks**, so make sure you have **correctly labelled axes, operationalised variables, an appropriate scale and an accurate diagram**.

CC BY NC ND

3.5. Distributions

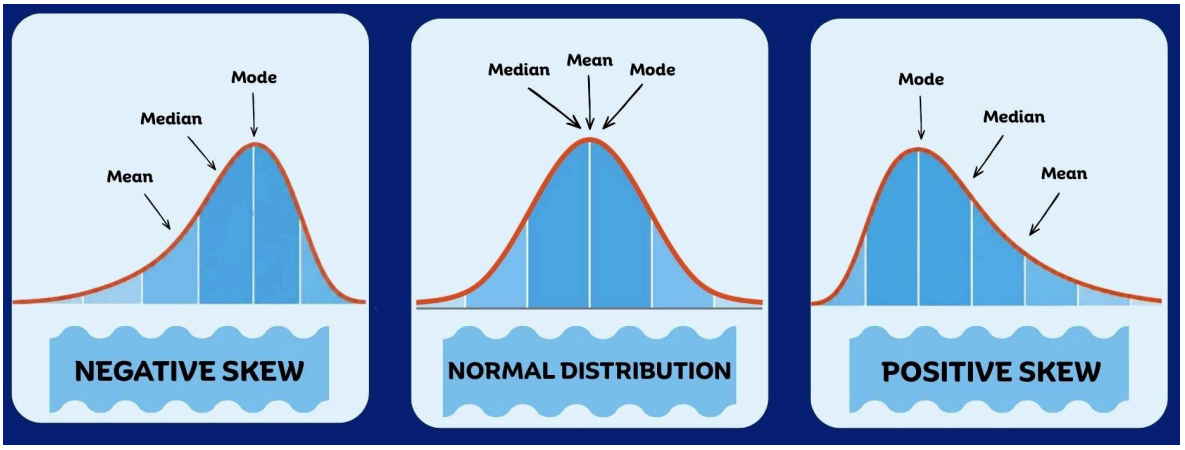
A **distribution** shows how scores are **spread across a set of data**. The shape of a distribution can be **normal** or **skewed**.

Normal distribution is a **symmetrical, bell-shaped distribution** where scores are evenly distributed around the centre. In a normal distribution, the **mean, median and mode are all approximately equal**.

A **skewed distribution** is **not symmetrical**, with scores concentrated more towards one end and a **long tail** extending towards the other. There are two types:

- **Positive skew (right skew)** – most scores are concentrated towards the **left/lower end**, while a small number of higher scores create a **long tail to the right**. The **mean is greater than the median**, which is greater than the mode.

- **Negative skew (left skew)** – most scores are concentrated towards the **right/higher end**, while a small number of lower scores create a **long tail to the left**. The **mean is less than the median**, which is less than the mode.



3.6. Levels of Measurement

Data can be measured at different **levels of measurement**: **nominal, ordinal and interval**. The level of measurement determines which **statistical tests and descriptive statistics** can appropriately be used.

Level	What it means	Psychology example
Nominal	Data is placed into **separate categories with no order or ranking**. Each person or item will **usually belong to one category only**.	**Attachment type:** secure, insecure-avoidant or insecure-resistant. The categories have **no numerical value** and cannot be ranked from highest to lowest.
Ordinal	Data can be **ordered or ranked**, but the **gaps between scores are not necessarily equal**.	**Rating enjoyment of psychology:** 1 = strongly dislike to 10 = strongly like. We know that **8 is higher than 6**, but cannot assume the difference between 8 and 6 is the same as between 4 and 2.
Interval	Numerical data where the **difference between each score is equal and meaningful**.	**Height in centimetres:** 160 cm, 161 cm, 162 cm. The difference between each centimetre is equal. **Temperature in °C** is another example.

CC BY NC ND

4. Inferential Testing

4.1. Introduction to Statistical Testing: The Sign test

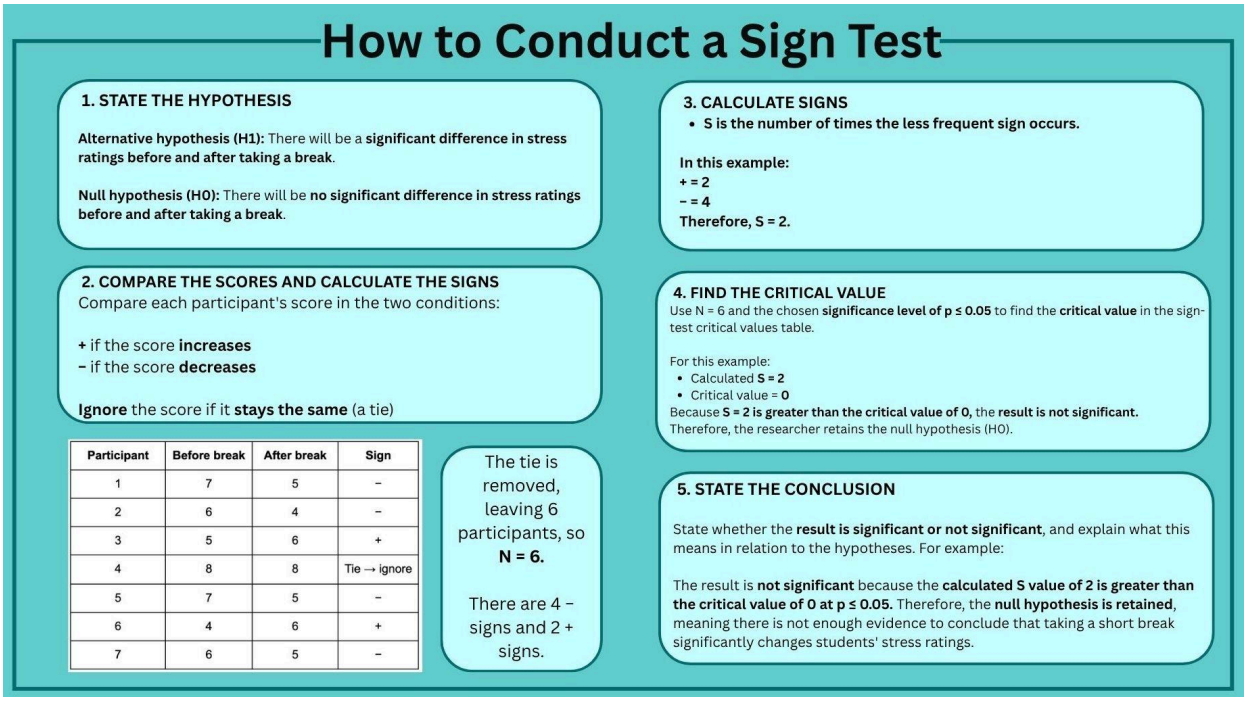
Statistical testing allows psychologists to decide whether a difference or relationship found in their data is **statistically significant** or whether it could have occurred **by chance**.

A result is **statistically significant** when the probability of getting the result by chance is sufficiently low. In A-level Psychology, the usual significance level is **$p \leq 0.05$**, meaning there is a **5% or less probability that the results occurred by chance**, assuming the null hypothesis is true.

Sign Test

The **sign test** is used to test whether there is a **significant difference between two related conditions** when the data is **ordinal**.

Example: A researcher wants to investigate whether taking a short break affects students' **stress levels**. The same students rate their stress **before and after** taking a break.



Revision tip: The **sign test is the only statistical test you may be asked to actually complete in the exam**. You need to **know the steps involved in conducting the sign test**, but you will be given **all the relevant information needed to carry it out**, including the **critical value table**. Therefore, **you do not need to memorise the critical values**.

4.2. Probability and Significance

Probability refers to **how likely something is to happen by chance**. It ranges from **0 to 1**, where 0 means impossible and 1 means certain.

Significance refers to whether a statistical test shows that a finding is **unlikely to have occurred by chance**.

The usual **significance level** in psychology is **$p \leq 0.05$**. This means there is a **5% or less probability that the result occurred by chance**. Therefore, we can be **95% confident that the result is not simply due to chance** and is likely to reflect an effect of the **IV on the DV**.

Why 0.05?

The **0.05 significance level** is commonly used in psychology because it provides a **balance**:

- It requires reasonably strong evidence that a finding is **unlikely to be due to chance**.
- It is not so strict that psychological research would require an **extremely high level of evidence** every time.

In some areas of research, the **consequences of making a wrong decision are much more serious**, so a stricter significance level may be used, such as **$p \leq 0.01$**. For example, in **medical or drug research**, incorrectly concluding that a treatment is safe or effective could have **serious consequences for people's health**.

The more serious the consequences of an incorrect conclusion, the stronger the evidence researchers may require.

Deciding Whether a Result Is Significant

After carrying out a statistical test, the researcher uses the **p-value** to decide whether the result is statistically significant. The **p-value is compared with the significance level** (**$p \leq 0.05$**).

If the result is significant:

- **$p \leq 0.05$**
- There is a **5% or less probability of obtaining a result this extreme by chance**, assuming the null hypothesis is true.
- The result is considered **unlikely to have occurred by chance**.
- **Reject the null hypothesis (H_0)**
- **Support the alternative hypothesis (H_1)**

If the result is not significant:

- **$p > 0.05$**
- There is a **greater than 5% probability of obtaining a result this extreme by chance**, assuming the null hypothesis is true.

CC BY NC ND

- There is **not enough evidence to conclude that the finding is statistically significant**.
- **Retain the null hypothesis (H_0)**
- There is **not enough evidence to support the alternative hypothesis (H_1)**

Using Statistical Tables and Critical Values

Statistical tests are used to determine whether a **significant difference or relationship** exists. This is decided by comparing the **calculated value** with the **critical value**.

The **calculated value** is the result obtained from carrying out the statistical test.

The **critical value** is the **cut-off point** that the calculated value must reach for the result to be **statistically significant**.

The critical value is found using the appropriate **statistical table**. It depends on factors such as:

- whether the test is **one-tailed or two-tailed**
- the chosen **significance level (p)**, usually **$p \leq 0.05$**
- the **number of participants (N)** or **degrees of freedom (df)**, depending on the test

The researcher:

1. Calculates the **calculated value** from the data.
2. Uses the appropriate statistical table to find the **critical value**.
3. **Compares the calculated value with the critical value.**
4. Decides whether the result is **significant**.

Different statistical tests use different rules for deciding whether a result is significant. For some tests, the calculated value must be greater than or equal to the critical value. For other tests, the calculated value must be lower than or equal to the critical value.

Rule for significance	Statistical tests
Calculated value ≥ critical value	**Pearson's r, Spearman's rho, Chi-squared, related t-test, unrelated t-test**
Calculated value ≤ critical value	**Sign test, Wilcoxon, Mann–Whitney U**

Revision Tip: Rule of R, if the test has an 'R' in its name, the calculated value must be greateR than or equal to the critical value for the result to be significant.

CC BY NC ND

Type I and Type II errors

Even when researchers use statistical testing, they can make errors when deciding whether to **reject or retain the null hypothesis**.

Error	What happens	What the researcher concludes	When is it more likely?
Type I error	The null hypothesis is **rejected when it is actually true**.	The researcher concludes there is a significant effect/difference when there isn't one (**false positive**).	More likely when the **significance level is set too high**, such as **$p \leq 0.10$**, because it is easier to reject the null hypothesis.
Type II error	The null hypothesis is **retained when it is actually false**.	The researcher concludes there is no significant effect/difference when there actually is one (**false negative**).	More likely when the **significance level is set too low**, such as **$p \leq 0.01$**, because it is harder to reject the null hypothesis.

Revision Tip: Remember the **Boy Who Cried Wolf**:

- **Type I:** At **first,** he says there's a wolf, but there **isn't → false positive**
- **Type II:** The **second** time there **is** a wolf, but the villagers **don't believe him → false negative**

CC BY NC ND

4.3. Factors Affecting the Choice of Statistical Test

When choosing an **inferential statistical test**, there are **three things** you need to consider:

1. **What are you testing?** Are you looking for a **difference** between conditions or a **relationship/association** between variables?
2. **What level of measurement is the data?** Is it **nominal, ordinal or interval**?
3. **What is the experimental design?** If you are testing a difference, is the design **related** or **unrelated**?

Related design = the same or matched participants take part in both conditions.
Unrelated design = different participants take part in each condition.

Choosing the test

	Test of Difference		**Association**
	Related Groups	**Unrelated Groups**	**Correlations**
Nominal	Sign Test (S)	Chi Squared (χ^2)	
Ordinal	Wilcoxon (T)	Mann Whitney (U)	Spearman's Rho (r_s)
Interval	Related T-Test (t)	Unrelated T-Test (t)	Pearsons (r)

To remember the order of the tests in the table, use the mnemonic:

"Silly Cats Wear Massive Silver Rings Under Pressure."

Revision tip: In an exam, you may be given a research scenario and asked to **select the most appropriate statistical test and explain why**. This question can be worth **7 marks**: **1 mark** for correctly identifying the statistical test and **6 marks** for explaining **three reasons why the test is appropriate**, **2 marks for each reason**. You therefore need to identify the **type of test, level of measurement and experimental design**

CC BY NC ND