



Research Design

Contents

- * Reliability
- * Validity
- * Generalisability
- * Researcher & Participant Biases
- * Reflexivity
- * Transferability & Credibility
- * Sampling Techniques



Reliability

- Reliability is achieved if a study has been set up so that the **IV** can be seen to **affect** the **DV**
 - If the procedure is **replicated**, the study should show **similar results**
 - Replicating a study and finding similar results shows that the measure is **consistent** and not liable to **fluctuate** to any significant degree
- If a study is completed using a **standardised procedure** and obtains the same results, it can be said that it is **reliable**

Reliability across methods

- Lab experiments** are the most **reliable method**, as they:
 - take place in a **neutral space** under **controlled conditions**
 - follow a standardised procedure
 - use **random allocation** of participants to **conditions**
 - tend to use a **control group** as a comparison to the **experimental group**
 - generate **quantitative data** which is easy to **compare** and **analyse**
- Field experiments** generate quantitative data and manipulate an IV but they are exposed to extraneous variables that reduce reliability
- Natural experiments** generate quantitative data but they use a **naturally occurring IV** over which the research has **no control**, making them **less reliable** than lab experiments

Types of reliability

- Internal reliability:** The extent to which a measure is **consistent within itself**
 - E.g., all items in a questionnaire measuring stress should give consistent results

External reliability: The extent to which a measure is **consistent over time**

- E.g., a test should produce similar results if given to the same person at two different times

Measuring reliability

- The **test-retest** method measures **external reliability**:
 - The same participants are given the **same** questionnaire at **separate time intervals** (e.g., with a 6-month gap between testing sessions)
 - If scores are consistent, external reliability is established



- The **split-half** method measures **internal reliability**:
 - The researcher splits the test in half and analyses the responses given to the **first half** of the questionnaire compared to the **second half** of the questionnaire
 - If responses are similar across both halves, internal reliability is established

Inter-observer reliability

- For an **observational study**, the researchers will have previously agreed on which specific behaviours should be **recorded** so that all observers are in agreement before the observation
- Researchers can test the reliability of their observations by **comparing** them with another researcher's recording of their behaviours
 - The level of **consistency** between the two records is then compared
- **Inter-observer reliability** is the level of **consistency** between two or more **trained** observers when they conduct the **same** observation, as follows:
 - Observers agree in advance on behaviour categories and how to record them
 - Each observer records behaviours independently to avoid conformity
 - After the observation, records are compared (often using a **tally chart**)
 - A correlation is calculated between observers' data
 - A strong positive correlation indicates **good inter-observer reliability**
- Establishing good inter-observer reliability means that there is less chance that researcher **bias** has interfered with the observation

An example of a tally chart used to record categories of behaviour in an observation:

Behaviour	Tally
Clenches fist	
Frowns	
Crosses arms	
Raises eyebrows	



Internal validity

- Validity is the extent to which the findings of a study are **true to life** and measure what they are intended to measure
 - Do the findings reflect how people **really think, feel, and behave**?
 - Do they measure what the researcher set out to investigate?
 - Could extraneous factors have influenced the results?
 - Can the findings be **generalised** to wider populations, contexts, and times?
- **Internal validity** measures the extent to which the results are due to the **manipulation** of the **IV** rather than the influence of **extraneous** or **confounding variables**
- This ensures the study demonstrates a clear **cause-and-effect relationship**
 - **Bandura's** (1961) investigation into **social learning theory** used the same toys and the same actions by the aggressive **model**
 - **Asch's** (1951) research on **conformity** used the same line-length stimuli, with the participant seated at the same place at the table, on each of the critical and non-critical **trials**
- High internal validity means that the conclusions drawn from a study are trustworthy and free from outside factors

External validity

- **External validity** measures the extent to which the results can be **generalised** beyond the research setting
- **Ecological validity** is a type of external validity that refers to how realistic the task/environment is
- External validity is high when the task participants are given is more aligned to a real, everyday experience rather than a task that is **artificial** or **contrived** or when their engagement with the task is real, even if the task itself is artificial (lacks mundane realism)
 - **Dickerson** (1992) used **naive participants** in a real setting to investigate the effect of **commitment on prosocial behaviour**
 - **Boyden** (2003) documented the **lived experiences** of children living in war zones
- High external validity means the results are transferable to different populations and settings and scenarios

Temporal validity

- **Temporal validity** measures the extent to which research findings are **relevant** over time



- E.g., **Asch's** research reflect the **post-WWII social climate**; it's unlikely to show the same conformity levels today
- **Bowlby's** maternal-deprivation theory is outdated due to changing family structures
 - E.g., single parents, blended families, stay-at-home fathers, same-sex parents
- High temporal validity means that a study's findings apply across different time periods (past, present, future)

Construct validity

- **Construct validity** refers to how well a study **measures the psychological construct** (theory/idea) it claims to investigate
- It is important for abstract concepts (e.g., intelligence, mood, depression)
- Researchers must show their measurement **aligns with theoretical understanding**.
 - E.g., **Asch** used a task which was **unambiguous** so if the participants conformed it was clear evidence of conformity
- High construct validity means that a study's findings accurately represents the theory in action

Predictive validity

- **Predictive validity** measures how well a study's findings predict a **future** outcome or behaviour
 - E.g., **Ainsworth's (1970) Strange Situation** research categorises people into different **attachment styles** as infants which predict future **relationship** behaviours
- High predictive validity means that a study's findings forecast future behaviour which could be used in educational or business settings for example



Generalisability

- **Generalisability** in research refers to the extent to which a study's findings can be applied to a wider **population**, setting, or time frame
 - E.g., **Kyle et al.'s** (2016) study of the prevalence of **obesity** has good generalisability as they used a sample of 13,483 participants (this ensures secure statistical power)
 - However, as the sample were all nurses in Scotland this limits the extent to which the findings can be generalised to other populations

Types of generalisability

Sample generalisability

- **Sample-population** generalisability involves **inferring** the results derived from a sample and applying it to a population
- To do this, the sample must be:
 - a random and **representative** sample
 - a sufficiently large sample size - the larger the sample the stronger the generalisability
- Small, specific samples cannot easily be generalised, which is why (in **quantitative** research) there is **power in numbers**

Inferential generalisability

- **Inferential** generalisability means that the study's findings can be applied to other settings or populations outside the setting of the study
- This is linked to **external validity**, e.g.,
 - Dickerson (1992) investigated prosocial behaviour in a **real-world context**
 - Boyden (2003) documented the experiences of children in **war zones**

Theoretical generalisability

- **Theoretical** generalisability means that concepts or theories developed from findings can inform **further research** and theory-building
- This type of generalisability is more common with **qualitative research**, where the aim is often to generate new insights rather than to generalise statistically
 - Qualitative research aims for transferability, i.e., whether the **insights** derived from the research can be transferred to **help our understanding of similar contexts**

Link to validity

- Generalisability is a component of **validity**, particularly **external validity**



Your notes

- **External validity** is the extent to which findings can be applied to **real-life contexts, populations, or times**
- **High external validity** is achieved when tasks are realistic or when participants are genuinely engaged, even in artificial settings



Researcher biases

- Researcher biases occur when the **researcher's presence, behaviour, or expectations** interfere with the research process, reducing validity

Investigator effects

- One source of researcher bias is known as the **investigator effect**
 - This occurs when the researcher's characteristics, such as **age, gender and ethnicity, influence** how participants interact with them
 - E.g., a participant may react differently if the researcher reminds them of someone from their past
- The researcher's **tone of voice, non-verbal communication** and what they are **wearing** can impact how participants respond to the research process
 - A **harsh or overly excited** tone of voice adds emotion to a **neutral task**
 - Using **overly dramatic body language** is at odds with **scientific** research and is likely to put participants in the wrong frame of mind
 - Wearing clothing that is overly **vibrant, patterned** or includes **slogans** is too personal and unprofessional

Confirmation bias

- Another source of researcher bias is **confirmation bias**
 - This is the tendency for researchers to **focus on evidence that supports their hypothesis** while overlooking contradictory findings
 - E.g., Ainsworth (1970) may have overlooked secure behaviours in a child already judged as insecure

Question order bias

- Another source of researcher bias is **question order bias**
 - The researcher could be **biased** in the way that they order the questions on a **questionnaire or interview**
 - Early questions influence how later ones are answered
 - E.g., asking, "Are you against animal testing?" early on may bias subsequent answers
 - A solution is to begin with **neutral questions** and move on to **specific/loaded ones** later

Leading question bias



- Another source of researcher bias is **leading questions** bias
 - This occurs when researchers ask questions which suggest what the answer should be
 - E.g., 'Don't you agree that animal testing is wrong?'
- Bias can also emerge through **instructions** or **task framing**, subtly signalling researcher expectations
 - E.g., if the researcher has **hypothesised** that X condition will result in higher scores, then they may unconsciously communicate this to the participants

Sampling bias

- Another source of researcher bias is sampling bias
 - This occurs when the researcher obtains a sample that does accurately **represent the target population**
 - This is more likely to occur with sampling methods such as opportunity sampling and self-selecting sampling
 - Sampling bias means that the findings cannot be generalised to a wider population, reducing the study's external validity

Participant biases

- Participant biases occur when participants **alter their behaviour** (consciously or unconsciously) because being studied is not a natural experience

Demand characteristics

- Because taking part in research is not an everyday experience, participants may behave in a way which does not **reflect** their usual, everyday behaviour
- Participants pick up **cues** about what is expected
- They may try to confirm (or resist) the researcher's aim
- Influences include:
 - The **research setting** (e.g., artificiality of a lab)
 - Any **verbal or non-verbal communication** from the researcher
 - Participants' **motives** (e.g., to please, rebel, or because of nerves)
- If participants behave according to what they *think* the aim of the research is, it leads to **artificial behaviour** that lacks ecological validity

Social desirability bias

- Another source of participant bias is social desirability bias
 - This is the tendency for people to **under-report negative** and **over-report positive** aspects of themselves when completing a questionnaire or interview



- Social desirability bias is common in self-report methods such as questionnaires and interviews

Dominant respondent bias

- Another source of participant bias is **dominant respondent bias**
 - This occurs when a lone individual (particularly in qualitative research such as **focus groups**) dominates the proceedings, possibly influencing the **opinions** and responses of the other participants
- Unless the dominant participant is handled (carefully and sensitively) by the researcher, they can invalidate research due to the impact of their presence

Acquiescence bias

- Another source of participant bias is **acquiescence bias**
 - This occurs when participants overwhelmingly say 'yes' or give **agreement** to questions asked by the researcher
- It may reflect a desire to **please the researcher**, personality tendencies, or lack of investment in the task
- Acquiescence bias can be reduced by:
 - Using **open-ended questions**
 - Encouraging participants to provide their **own views**



Reflexivity

- **Reflexivity** is a strategy used by **qualitative** researchers to examine how their own subjectivity and choice of methods may influence the findings in order to increase credibility
 - E.g., they must examine their own **biases, prejudices and preconceived ideas** throughout the entire research process
- If a researcher fails to use reflexivity, they could be harming the **validity** of their research

Personal reflexivity

- **Personal reflexivity** can be used in several ways
 - The researcher examining their own values, experiences and beliefs, and how these could affect the research process
 - Contemplating how a particular research topic affects them on a personal level and whether this intrudes into the professional domain
 - Keeping a **journal**/field diary with **reflective comments** so that the researcher's approach and thoughts during the research process are well documented

Epistemological reflexivity

- Epistemological reflexivity can be used alongside personal reflexivity
 - This helps the researcher to examine the ways in which **knowledge** has been generated in a study
 - The researcher should ask questions as to how they have come about this knowledge/insight and if they have made **assumptions** which would **undermine** the **academic integrity** of the research
 - The researcher should reflect on why they chose the specific methods that were used and how such methods led to particular outcomes
 - Strengths and limitations of the methods should form part of this reflection
 - The researcher should aim to be **open, honest and transparent** in the way that they **report** their findings
 - This should lead to the research being viewed as credible and as a product of the participants' behaviour/responses rather than those of the researcher



Transferability & credibility

- **Transferability** is a type of **generalisability** used specifically with **qualitative data**
 - It refers to the extent to which a study's findings can be applied/generalised to other populations/contexts outside of the study's parameters
- The aim is not statistical generalisation but to ensure that findings **resonate with readers** by providing **rich insight and detail**
- Good transferability evokes a sense of **familiarity and shared understanding** of the experience described
 - This may happen easily for topics which are more common among a wider population (e.g., school days)
 - The findings are **less likely** to be transferable if the topic/experience is more 'niche' or **idiosyncratic** (e.g., becoming a refugee)
 - Research which is **outdated** may not transfer easily to modern contexts, which is something for researchers to consider as well
- To ensure transferability, researchers should fully describe
 - the **methods** (e.g., interview type, structured/unstructured)
 - **sampling** (how participants were obtained, why selected)
 - **procedure** (how the study was conducted, materials used, recording methods)
 - **analysis** (how data were analysed, why that method was chosen)
 - **findings** (how they relate to and support the theory being investigated)

Credibility

- Credibility is the **qualitative equivalent of reliability**
- Since qualitative research often uses **small samples** and **non-replicable methods**, reliability is reframed as **trustworthiness and believability**
- A credible study:
 - transparently explains **methods and materials** used
 - demonstrates the **competence and integrity** of the researcher
 - is published after scrutiny (e.g., **peer review**)

Ways to enhance credibility

- **Triangulation:**



Your notes

- One dataset checks another for consistency (e.g., interview transcript vs. questionnaire responses).
- One researcher checks interpretations with another (agreement on design, coding, or findings).
- **Researcher reflexivity:**
 - Honest reflection in the **discussion section** on both strengths and weaknesses
 - Transparency about limitations increases trustworthiness



Opportunity sampling

- Opportunity sampling is also known as **convenience** sampling
- This is where the researcher selects participants who are **available and willing** at the time
 - E.g., supermarket shoppers at 11am; students in a lecture; parents at a baby yoga group
- The sample used in psychological research is taken from a **target population**
 - A researcher who wishes to investigate the effects of being a single teen parent will require their sample to be drawn from that **specific population**
 - This is an example of a **distinct** target population
 - A researcher is running an experiment on the **duration of short-term memory (STM)**
 - Generally, anyone from the age of **18 to 60 years old** would suffice, i.e., a distinct target population is **not** required
- The researcher draws the sample from the population to take part in the experiment and then **generalises the findings** across the target population

Evaluation of opportunity sampling

Strengths

- The 'convenience' aspect of opportunity sampling is a strength, as it is a **quick and easy** way of obtaining participants for a study
- As people have been approached and agreed to take part, this is more likely to mean that the research process **runs smoothly**
 - Having **unwilling or resistant** participants could damage the **validity** of the findings

Limitations

- It is **not** possible to **generalise** from an opportunity sample, as the sample only **represents** those who were available and willing to participate at the time
- The researcher may be prone to **unconscious bias** when they approach people to take part in the study, e.g.,
 - they may tend to approach people they feel **comfortable** with
 - they may select only those participants who they think will be **interested**
 - they may **avoid** some **social groups** that they are wary of, e.g., males aged 18–25

Self-selecting sampling



- **Self-selecting sampling** (also known as **volunteer sampling**) involves people **actively selecting themselves** to participate in a study, i.e., they volunteer to take part
- A researcher finds a specific way or ways in which to find their sample, e.g.,
 - putting up **posters and flyers** around a **university campus** asking for volunteers to participate in a study
 - posting an advert on various **social media platforms**
 - running an advert in a **print newspaper**
- The research will **advertise when and where** the study is taking place and how they can participate
- The advert may specifically ask for people with certain characteristics, e.g.,
 - first-time **parents**
 - people with **ADHD**
 - bilingual people

Evaluation of self-selecting sampling

Strengths

- This method is **quick, easy and cost-effective**
 - It is one of the most used (probably *the* most popular) sampling methods by psychologists
- Participants are likely to be **willing and enthusiastic** about the research
 - They have made a **conscious decision** to take part in the research, which means that they are less likely to jeopardise the study and its results

Limitations

- This method is prone to **volunteer bias**
 - The results are not easy to **generalise**, as volunteer participants often have **personality traits** in common, e.g. sociable, outgoing, etc.
- Volunteers may be prone to acquiescence bias
 - This tendency to please the researcher may give rise to **demand characteristics**, which in turn affect the **validity** of the findings

Random sampling

- **Random sampling** involves obtaining a sample taken from a population in way which has the **least bias** out of all of the sampling methods
- With random sampling, **every member** of the population has an **equal chance** of being selected
- How a random sample could be achieved includes:



- putting all the names of the population in a **container** and then drawing out one name at a time until the **optimum sample size** has been reached
 - e.g., a sample size of 50
- employing a computer **name-generator software**
 - this is more useful if a large sample is required, e.g., 2,000 participants

Evaluation of random sampling

Strengths

- This sampling method eliminates **researcher bias**, as the researcher has no control over who is selected to be in the sample
- Using a random sample means that the study results should be fairly **representative**
 - This means that the findings can be **generalised** to the target population

Limitations

- Random sampling can be **time-consuming** and **impractical**
 - It is not always possible to get access to all the information on a target population
 - Additionally, not everyone selected for the sample may want to take part in the study
- Random sampling can result in a **non-representative sample**
 - Due to the nature of the method, the sample could be **unbalanced**
 - An **all-male** sample could be selected randomly, which may **not be a true representation** of the target population

Stratified sampling

- **Stratified sampling** generates a **small-scale** reproduction of the target population
 - The target population is divided and categorised according to **key characteristics** required by the research, e.g.,
 - **age**
 - **gender**
 - **education level**
 - **ethnicity**
 - **profession**
 - The population is sampled within each category **proportional** to the overall total
 - E.g., If the whole population has a total of 18% of males aged 30–40, then the **representative** sample will have 18% of males aged 30–40

Evaluation of stratified sampling



Strengths

- The sample is **representative** of the target population as it is based on exact proportions of the target population
 - This means that it is easy to **generalise** data from the sample to the wider population
- Stratified sampling means that the researcher has **control** over the chosen **categories**, which can be selected according to how relevant they are in terms of the research **aim**

Limitations

- Stratified sampling can be difficult when researchers **cannot confidently classify** every member of the population into a subgroup
- Gathering the sample population can be **time-consuming**
 - It is not always possible to get access to all the information on a target population

Snowball sampling

- **Snowball sampling** is used when the researcher wishes to gain access to a population that may be difficult to find for various reasons such as:
 - they belong to an **exclusive** group or one which is tricky to access, e.g., ex-members of a cult, gang members
 - they would be **unlikely to respond** to the usual type of advertisement asking for participants, e.g., people who are heroin users and people who have spent time in prison
 - they may feel **nervous or compromised** if the researcher were to approach them directly, e.g., people who regularly shoplift and people who suffer from extreme post-natal **depression**
- To get around the above problems, the researcher finds their first one (or a few) participants and asks them if they can **recruit** other, similar participants for the study
 - Once the existing participant(s) start recruiting others, then these new participants may in turn recruit more participants

Evaluation of snowball sampling

Strengths

- This type of sampling means that hard-to-reach populations can be represented in research
 - Without this sampling method it would be difficult to understand the experiences of people who have **unconventional** lives
- Being recommended to a researcher by someone in a similar situation (e.g., a fellow heroin user or ex-cult member) can instil feelings of **trust** in the participant
 - Establish good rapport between researcher and participant is particularly important in qualitative research

Limitations

- The very specific choice of participants – who are likely to report similar experiences – means that the **scope** of the research is limited somewhat
- The researcher has little **control** over who joins the sample, having to rely on **recommendations and referrals** from other people
 - This may threaten the **credibility** of the findings



Your notes