

GB6013

VALIDITY & RELIABILITY

DR. FARIZA KHALID

VALIDITY & RELIABILITY



- Reliability and validity are concepts used to evaluate the quality of research.
- They indicate how well a method, technique or test measures something.
- Reliability is about the consistency of a measure, and validity is about the accuracy of a measure.

VALIDITY & RELIABILITY



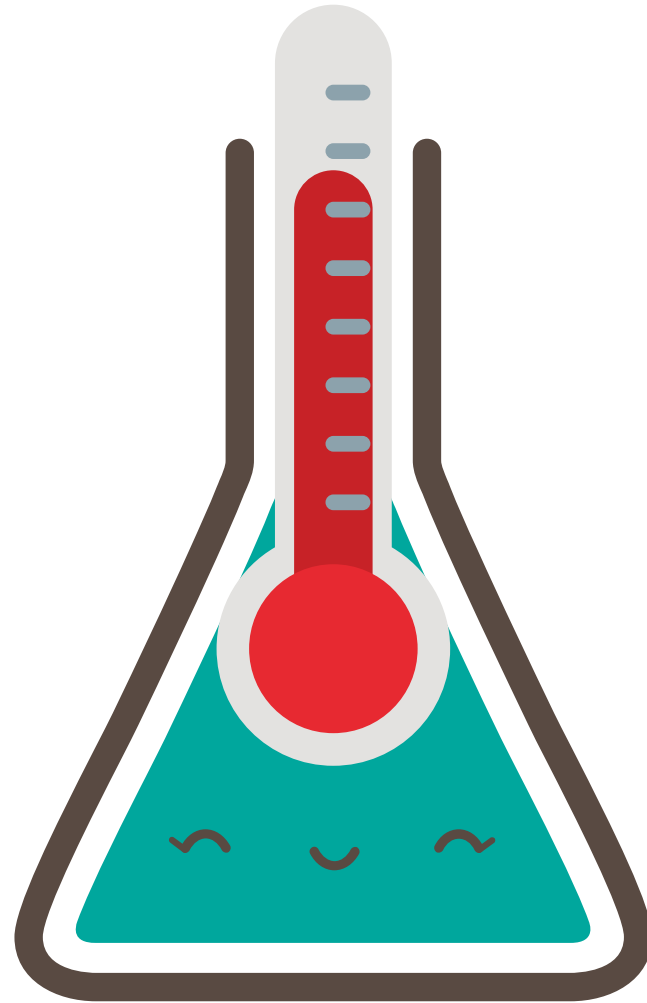
- It's important to consider reliability and validity when you are creating your research design, planning your methods, and writing up your results, especially in quantitative research.

RELIABILITY



- Reliability refers to how consistently a method measures something. If the same result can be consistently achieved by using the same methods under the same circumstances, the measurement is considered reliable.

EXAMPLE



- You measure the temperature of a liquid sample several times under identical conditions. The thermometer displays the same temperature every time, so the results are reliable.

EXAMPLE



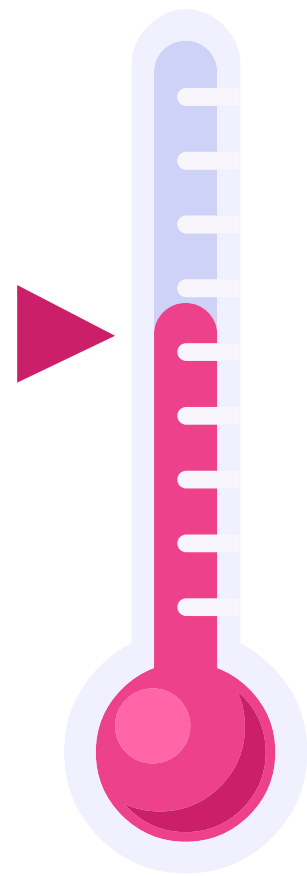
- A doctor uses a symptom questionnaire to diagnose a patient with a long-term medical condition.
- Several different doctors use the same questionnaire with the same patient but give different diagnoses.
- This indicates that the questionnaire has low reliability as a measure of the condition.

VALIDITY



- Validity refers to how accurately a method measures what it is intended to measure.
- If research has high validity, that means it produces results that correspond to real properties, characteristics, and variations in the physical or social world.
- High reliability is one indicator that a measurement is valid.
- If a method is not reliable, it probably isn't valid.

EXAMPLE



- If the thermometer shows different temperatures each time, even though you have carefully controlled conditions to ensure the sample's temperature stays the same, the thermometer is probably malfunctioning, and therefore its measurements are not valid.

EXAMPLE



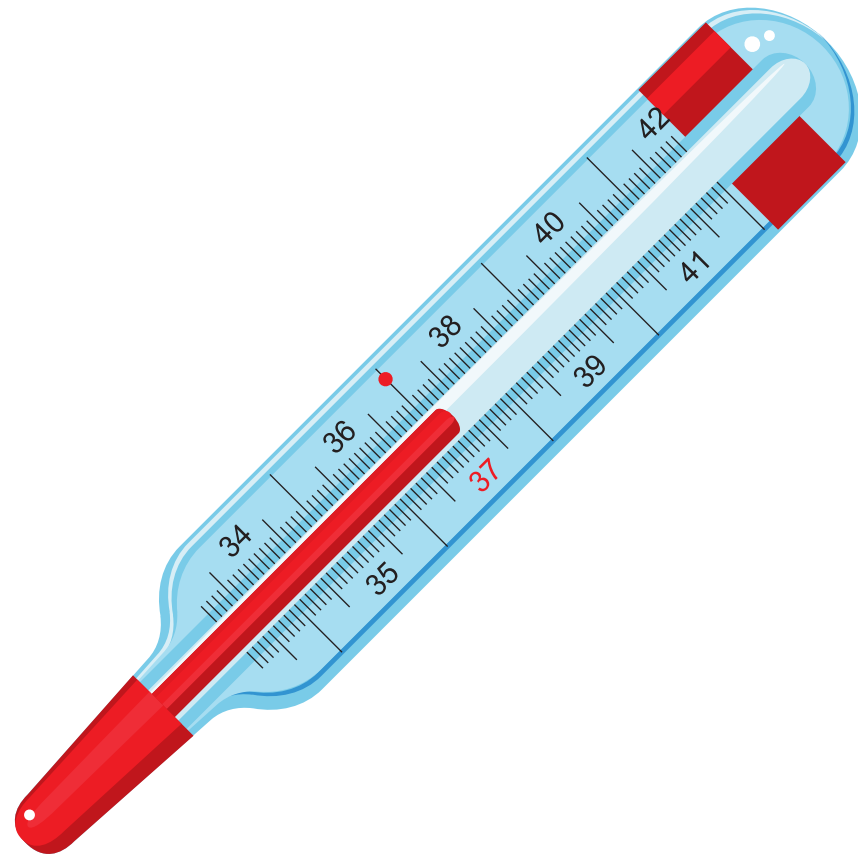
- If a symptom questionnaire results in a reliable diagnosis when answered at different times and with different doctors, this indicates that it has high validity as a measurement of the medical condition.

REMEMBER



- Reliability on its own is not enough to ensure validity. Even if a test is reliable, it may not accurately reflect the real situation.

EXAMPLE



- The thermometer that you used to test the sample gives reliable results.
- However, the thermometer has not been calibrated properly, so the result is 2 degrees lower than the true value. Therefore, the measurement is not valid.

EXAMPLE



- A group of participants take a test designed to measure working memory. The results are reliable, but participants' scores correlate strongly with their level of reading comprehension. This indicates that the method might have low validity: the test may be measuring participants' reading comprehension instead of their working memory.

VALIDITY



- Validity is harder to assess than reliability, but it is even more important. To obtain useful results, the methods you use to collect your data must be valid: the research must be measuring what it claims to measure. This ensures that your discussion of the data and the conclusions you draw are also valid.

VALIDITY & RELIABILITY

	Reliability	Validity
What does it tell you?	The extent to which the results can be reproduced when the research is repeated under the same conditions.	The extent to which the results really measure what they are supposed to measure.
How is it assessed?	By checking the consistency of results across time, across different observers, and across parts of the test itself.	By checking how well the results correspond to established theories and other measures of the same concept.
How do they relate?	A reliable measurement is not always valid: the results might be reproducible, but they're not necessarily correct.	A valid measurement is generally reliable: if a test produces accurate results, they should be reproducible.

TYPES OF RELIABILITY

Type of reliability	What does it assess?	Example
Test-retest	The consistency of a measure across time : do you get the same results when you repeat the measurement?	A group of participants complete a questionnaire designed to measure personality traits. If they repeat the questionnaire days, weeks or months apart and give the same answers, this indicates high test-retest reliability.

TYPES OF RELIABILITY

Type of reliability	What does it assess?	Example
Interrater	The consistency of a measure across raters or observers : do you get the same results when different people conduct the same measurement?	Based on an assessment criteria checklist, five examiners submit substantially different results for the same student project. This indicates that the assessment checklist has low inter-rater reliability (for example, because the criteria are too subjective).

TYPES OF RELIABILITY

Type of reliability	What does it assess?	Example
Internal consistency	The consistency of the measurement itself : do you get the same results from different parts of a test that are designed to measure the same thing?	You design a questionnaire to measure self-esteem. If you randomly split the results into two halves, there should be a strong correlation between the two sets of results. If the two results are very different, this indicates low internal consistency.

TYPES OF VALIDITY

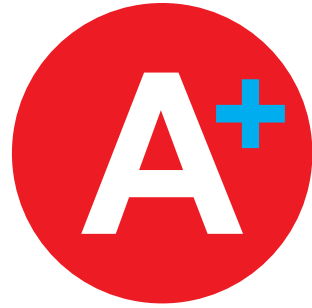
Type of validity	What does it assess?	Example
Construct	The adherence of a measure to existing theory and knowledge of the concept being measured.	A self-esteem questionnaire could be assessed by measuring other traits known or assumed to be related to the concept of self-esteem (such as social skills and optimism). Strong correlation between the scores for self-esteem and associated traits would indicate high construct validity.

TYPES OF VALIDITY

Type of validity	What does it assess?	Example
Content	The extent to which the measurement covers all aspects of the concept being measured.	A test that aims to measure a class of students' level of Spanish contains reading, writing and speaking components, but no listening component. Experts agree that listening comprehension is an essential aspect of language ability, so the test lacks content validity for measuring the overall level of ability in Spanish.

TYPES OF VALIDITY

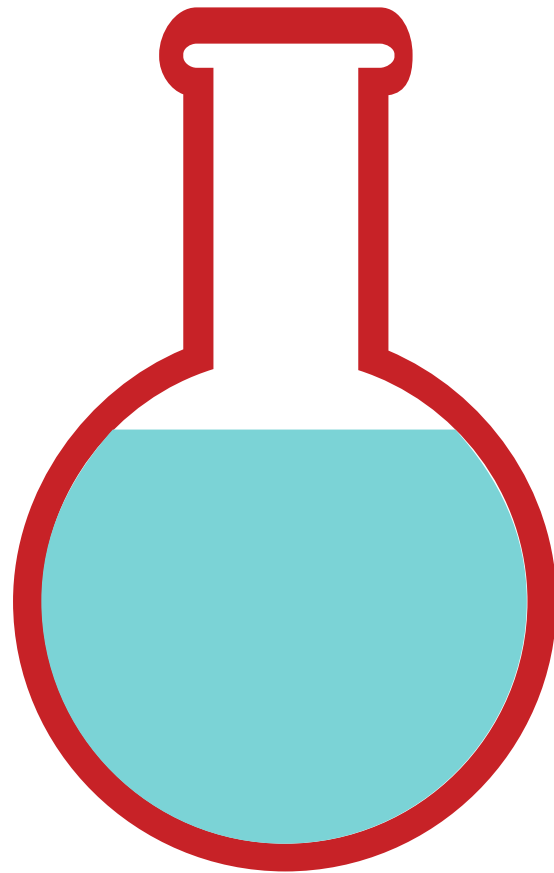
Type of validity	What does it assess?	Example
Criterion	The extent to which the result of a measure corresponds to other valid measures of the same concept.	A survey is conducted to measure the political opinions of voters in a region. If the results accurately predict the later outcome of an election in that region, this indicates that the survey has high criterion validity.



ENSURING VALIDITY

- If you use scores or ratings to measure variations in something (such as psychological traits, levels of ability or physical properties), it's important that your results reflect the real variations as accurately as possible.
- Validity should be considered in the very earliest stages of your research, when you decide how you will collect your data.

Choose appropriate methods of measurement



- Ensure that your method and measurement technique are high quality and targeted to measure exactly what you want to know. They should be thoroughly researched and based on existing knowledge.

Choose appropriate methods of measurement



- For example, to collect data on a personality trait, you could use a standardized questionnaire that is considered reliable and valid. If you develop your own questionnaire, it should be based on established theory or findings of previous studies, and the questions should be carefully and precisely worded.

Use appropriate sampling methods to select your subjects



- To produce valid generalizable results, clearly define the population you are researching (e.g. people from a specific age range, geographical location, or profession). Ensure that you have enough participants and that they are representative of the population.

ENSURING RELIABILITY



- Reliability should be considered throughout the data collection process. When you use a tool or technique to collect data, it's important that the results are precise, stable and reproducible.

Apply your methods consistently

- Plan your method carefully to make sure you carry out the same steps in the same way for each measurement. This is especially important if multiple researchers are involved.
- For example, if you are conducting interviews or observations, clearly define how specific behaviours or responses will be counted, and make sure questions are phrased the same way each time.

Standardize the conditions of your research

- When you collect your data, keep the circumstances as consistent as possible to reduce the influence of external factors that might create variation in the results.
- For example, in an experimental setup, make sure all participants are given the same information and tested under the same conditions.

Standardize the conditions of your research

- When you collect your data, keep the circumstances as consistent as possible to reduce the influence of external factors that might create variation in the results.
- For example, in an experimental setup, make sure all participants are given the same information and tested under the same conditions.

VALIDITY & RELIABILITY



- Are used in quantitative research

CONFORMABILITY, CREDIBILITY, DEPENDABILITY & TRANSFERABILITY



- Are used in qualitative research

QUALITATIVE RESEARCH



- The concepts of objectivity, reliability, validity and generalisability provide a basic framework for conducting and evaluating traditional quantitative research.
- However, qualitative researchers contest and reject these positivist concepts (Linda, 2006).

QUALITATIVE RESEARCH



- Guba and Lincoln (1994) propose four criteria for qualitative research: confirmability, credibility, dependability and transferability, to replace the concepts of reliability, validity and generalisability in positivist approaches.

CONFORMABILITY



- In a qualitative study, which involves interpretation of the data by the researcher, the issue of the involvement of the researcher's 'self' in the interpretation of the data arises, as well as the possibility of keeping an open mind and being willing to consider alternative and competing explanations of the data (Denscombe, 2010).

CONFORMABILITY



- Confirmability is the 'qualitative researcher's comparative to the quantitative (or positivist) researcher's notion of objectivity' (Shenton, 2004, p. 72).

CONFORMABILITY

- However, as truth is something subjective to the context, it is impossible to achieve total confirmability (Slack, 2009); thus, an important factor for confirmability is for researchers to acknowledge their positionality as well as the reasons behind their choice of research approach as opposed to other approaches (Miles and Huberman, 1994 cited in Shenton, 2004), and to offer a self-critically reflexive analysis of the methodology used in the research (Linda, 2006).

CREDIBILITY



- Credibility replaces quantitative or positivist researchers' notion of validity in order to ensure that any research measures what it is supposed to measure (McNiff and Whitehead, 2009).

CREDIBILITY



- To achieve credibility in a qualitative study, it is important that the researchers establish confidence in the 'truth' of their findings (Linda, 2006).

CREDIBILITY



- However, no research is absolutely valid, be it quantitative or qualitative, and therefore researchers should attempt to 'minimise invalidity' and 'maximise validity' (Cohen, Manion and Morrison, 2001, p. 105).

CREDIBILITY



- Lincoln and Guba (1985) recommend prolonged engagement in the field, persistent observation and triangulation of data to enhance research credibility.

CREDIBILITY

- Denscombe (2010) also recommends triangulation, where the researcher can use contrasting data sources to bolster confidence and persuade readers.
- Example, data were generated from multiple sources: online questionnaires, focus group interviews, one-to-one interviews and online discussions

DEPENDABILITY

- According to Cohen et al. (2001), reliability relates to the notion of the same test being able to be repeated using the same conditions as previous tests.
- As a result, the question of reliability translates 'Would the research instrument produce the same results when used by different researchers (all other things being equal)?' into 'If someone else did the research, would he or she have got the same results and arrive at the same conclusions?' (Denscombe, 2010, p. 300).

DEPENDABILITY



- In an absolute sense, there is probably no way of knowing this for certain. However, there are ways of dealing with this issue in qualitative research, through the idea of dependability (Lincoln and Guba, 1985).

DEPENDABILITY



- To achieve dependability in qualitative research, it is vital to provide details of the research through documentation of data, methods and decisions about the research, i.e., an audit trail.

DEPENDABILITY



- In order to evidence dependability, the methods used within a study must be reported in detail to allow other researchers doing similar research, using the same methods, to be able to compare the findings between the studies (Shenton, 2004, p. 9).

DEPENDABILITY



- Example – To achieve dependability in a study, a researcher provides a high level of transparency by presenting details about the strategy of the research design and how it was implemented and writes about the way data were generated and the entire process of the study.

TRANSFERABILITY



- The concept of transferability replaces the concept of external validity for quantitative research (Lincoln and Guba, 1985).
- Transferability relates to the extent to which 'the findings within one study can be applied to other studies' (Shenton, 2004, p. 7).

TRANSFERABILITY

- As for transferability, qualitative researchers do not seek to statistically infer findings from a sample to the whole population.
- Instead, an imaginative process is used in which the reader of the research uses information about the particular instance that has been studied to arrive at a judgement about how far it would apply to other comparable instances (Denscombe, 2010).

TRANSFERABILITY

- The main focus in qualitative research is to show that the findings can be transferred and may have meaning or relevance if applied to other individuals, contexts and situations (Linda, 2006).
- Therefore, the rich, in-depth data derived from multiple sources in the triangulation process may enhance the transferability of the findings.

TRANSFERABILITY

- Example – in a study, individual stories have been heard and, whether single or multiple voices, all have been considered.
- However, all are subject to interpretation and confined within the study, which makes it 'problematic to demonstrate whether the findings and conclusions are applicable to other situations and populations' (Denscombe, 1998, p. 46).

TRANSFERABILITY

- However, Bryman (2004) reminds us of the fact that the findings in case study research are impossible to statistically generalise beyond the specific research context.
- This happens due the sampling of case study research, in which the selection of participants is not done using probability sampling of smaller sample size to represent the population (Bryman, 2004), resulting in the findings being impossible to generalise to the whole population.

TRANSFERABILITY

- Responding to this, Yin (2009) argues that case studies can actually be analytically generalised. In other words, this means that the findings of case studies based on the theoretical propositions of a study can be projected onto a new situation.

TRANSFERABILITY

- In addition, fuzzy generalisation (Bassegy, 2001) is also possible in case studies, where the generalisation is done based on prediction rather than on calculation; for example, the findings of case studies can be useful for other researchers to predict what may happen in their context of study based on the previous case study's findings, so what has occurred in classrooms in others' case studies can help others to predict what may happen in their classrooms (Bassegy, 2001).

TRANSFERABILITY

- In addition, fuzzy generalisation (Bassegy, 2001) is also possible in case studies, where the generalisation is done based on prediction rather than on calculation; for example, the findings of case studies can be useful for other researchers to predict what may happen in their context of study based on the previous case study's findings, so what has occurred in classrooms in others' case studies can help others to predict what may happen in their classrooms (Bassegy, 2001).