**Reliability vs validity: what’s the difference?**

**Date published July 3, 2019 by**[**Fiona Middleton.**](https://www.scribbr.com/author/fionamiddleton/)**Date updated: January 13, 2020**

Reliability and validity are concepts used to evaluate the quality of research. They indicate how well a [method](https://www.scribbr.com/category/methodology/), technique or test measures something. Reliability is about the consistency of a measure, and validity is about the accuracy of a measure.

It’s important to consider reliability and validity when you are creating your [research design](https://www.scribbr.com/research-process/research-design/), planning your methods, and writing up your results, especially in [quantitative research](https://www.scribbr.com/methodology/qualitative-quantitative-research/).

| **Reliability vs validity** | | |
| --- | --- | --- |
|  | **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. |

**Table of contents**

1. [Understanding reliability vs validity](https://www.scribbr.com/methodology/reliability-vs-validity/#understanding-reliability-vs-validity)
2. [How are reliability and validity assessed?](https://www.scribbr.com/methodology/reliability-vs-validity/#how-are-reliability-and-validity-assessed)
3. [How to ensure validity and reliability in your research](https://www.scribbr.com/methodology/reliability-vs-validity/#how-to-ensure-validity-and-reliability-in-your-research)
4. [Where to write about reliability and validity in a thesis](https://www.scribbr.com/methodology/reliability-vs-validity/#where-to-write-about-reliability-and-validity-in-a-thesis)

**Understanding reliability vs validity**

Reliability and validity are closely related, but they mean different things. A measurement can be reliable without being valid. However, if a measurement is valid, it is usually also reliable.

**What is 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.

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.

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.

**What is 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.

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.

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.

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

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.

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 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](https://www.scribbr.com/dissertation/discussion/) of the data and the [conclusions](https://www.scribbr.com/dissertation/write-conclusion/) you draw are also valid.

**How are reliability and validity assessed?**

Reliability can be estimated by comparing different versions of the same measurement. Validity is harder to assess, but it can be estimated by comparing the results to other relevant data or theory. Methods of estimating reliability and validity are usually split up into different types.

**Types of reliability**

Different types of reliability can be estimated through various statistical methods.

| [**Types of reliability**](https://www.scribbr.com/methodology/types-of-reliability/) | | |
| --- | --- | --- |
| **Type of reliability** | **What does it assess?** | **Example** |
| [**Test-retest**](https://www.scribbr.com/methodology/types-of-reliability/#test-retest-reliability) | 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. |
| [**Interrater**](https://www.scribbr.com/methodology/types-of-reliability/#interrater-reliability) | 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). |
| [**Internal consistency**](https://www.scribbr.com/methodology/types-of-reliability/#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](https://www.scribbr.com/methodology/correlational-research/) between the two sets of results. If the two results are very different, this indicates low internal consistency. |

**Types of validity**

The validity of a measurement can be estimated based on three main types of evidence. Each type can be evaluated through expert judgement or statistical methods.

| [**Types of validity**](https://www.scribbr.com/methodology/types-of-validity/) | | |
| --- | --- | --- |
| **Type of validity** | **What does it assess?** | **Example** |
| [**Construct**](https://www.scribbr.com/methodology/types-of-validity/#construct-validity) | 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. |
| [**Content**](https://www.scribbr.com/methodology/types-of-validity/#content-validity) | 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. |
| [**Criterion**](https://www.scribbr.com/methodology/types-of-validity/#criterion-validity) | The extent to which the result of a measure corresponds to **other valid measures** of the same concept. | A [survey](https://www.scribbr.com/methodology/survey-research/) 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. |

To assess the validity of a cause-and-effect relationship, you also need to consider[internal validity](https://www.scribbr.com/methodology/internal-vs-external-validity/) (the [design of the experiment](https://www.scribbr.com/methodology/experimental-design/)) and [external validity](https://www.scribbr.com/methodology/internal-vs-external-validity/) (the generalizability of the results).

How to ensure validity and reliability in your research

The reliability and validity of your results depends on creating a strong [research design](https://www.scribbr.com/research-process/research-design/), choosing appropriate methods and samples, and conducting the research carefully and consistently.

**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.

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](https://www.scribbr.com/methodology/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.

**Where to write about reliability and validity in a thesis**

It’s appropriate to discuss reliability and validity in various sections of your [thesis or dissertation](https://www.scribbr.com/category/dissertation/). Showing that you have taken them into account in planning your research and interpreting the results makes your work more credible and trustworthy.

| **Reliability and validity in a thesis** | |
| --- | --- |
| **Section** | **Discuss** |
| [**Literature review**](https://www.scribbr.com/dissertation/literature-review/) | What have other researchers done to devise and improve methods that are reliable and valid? |
| [**Methodology**](https://www.scribbr.com/dissertation/methodology/) | How did you plan your research to ensure reliability and validity of the measures used? This includes the chosen sample set and size, sample preparation, external conditions and measuring techniques. |
| [**Results**](https://www.scribbr.com/dissertation/results/) | If you calculate reliability and validity, state these values alongside your main results. |
| [**Discussion**](https://www.scribbr.com/dissertation/discussion/) | This is the moment to talk about how reliable and valid your results actually were. Were they consistent, and did they reflect true values? If not, why not? |
| [**Conclusion**](https://www.scribbr.com/dissertation/write-conclusion/) | If reliability and validity were a big problem for your findings, it might be helpful to mention this here. |