

Instrument Reliability

Martyn Shuttleworth 71.8K reads

Instrument reliability is a way of ensuring that any instrument used for measuring experimental variables gives the same results every time.

In the physical sciences, the term is self-explanatory, and it is a matter of making sure that every piece of hardware, from a mass spectrometer to a set of weighing scales, is properly calibrated.

 **EXPLORABLE**
Quiz Time!



Quiz:
Psychology 101 Part 2



Quiz:
Psychology 101 Part 2



Quiz:
Flags in Europe

[See all quizzes =>](#)

Instruments in Research

As an example, a researcher will always test the instrument reliability of weighing scales with a set of calibration weights, ensuring that the results given are within an acceptable margin of error [1].

Some of the highly accurate balances can give false results if they are not placed upon a completely level surface, so this calibration process is the best way to avoid this.

In the non-physical sciences, the definition of an instrument is much broader, encompassing everything from a set of survey questions to an intelligence test. A survey [2] to measure reading ability in children must produce reliable and consistent results if it is to be taken seriously.

Political opinion polls, on the other hand, are notorious for producing inaccurate results and delivering a near unworkable margin of error.

In the physical sciences, it is possible to isolate a measuring instrument from external factors [3], such as environmental conditions and temporal factors. In the social sciences, this is much more difficult, so any instrument must be tested with a reasonable range of reliability.

Test of Stability

Any test of instrument reliability [4] must test how stable the test is over time, ensuring that the same test performed upon the same individual gives exactly the same results.

The test-retest method [5] is one way of ensuring that any instrument is stable over time.

Of course, there is no such thing as perfection and there will be always be some disparity and potential for regression [6], so statistical methods [7] are used to determine whether the stability of the instrument is within acceptable limits.

Test of Equivalence

Testing equivalence involves ensuring that a test administered to two people, or similar tests administered at the same time give similar results.

Split-testing is one way of ensuring this, especially in tests or observations where the results are expected to change over time. In a school exam, for example, the same test upon the same subjects will generally result in better results the second time around, so testing stability is not practical.

Checking that two researchers observe similar results [8] also falls within the remit of the test of equivalence.

Test of Internal Consistency

The test of internal consistency [9] involves ensuring that each part of the test generates similar results, and that each part of a test measures the correct construct.

For example, a test of IQ should measure IQ only, and every single question must also contribute. One way of doing this is with the variations upon the split-half tests, where the test is divided into two sections, which are checked against each other. The odd-even reliability [10] is a similar method used to check internal consistency [9].

Physical sciences often use tests of internal consistency, and this is why sports drugs testers take two samples, each measured independently by different laboratories, to ensure that experimental [11] or human error did not skew or influence the results [12].

Source URL: <https://explorable.com/instrument-reliability>

Links

[1] <https://explorable.com/statistics-margin-of-error>

[2] <https://explorable.com/survey-research-design>

[3] <https://explorable.com/confounding-variables>

[4]

<http://www.gifted.uconn.edu/Siegle/research/Instrument%20Reliability%20and%20Validity/Reliability.htm>

[5] <https://explorable.com/test-retest-reliability>

[6] <https://explorable.com/linear-regression-analysis>

[7] <https://explorable.com/statistics-tutorial>

[8] <https://explorable.com/reproducibility>

[9] <https://explorable.com/internal-consistency-reliability>

[10] <http://allpsych.com/stats/unit2/46>

[11] <https://explorable.com/type-I-error>

[12] <https://explorable.com/statistically-significant-results>