



Internal Consistency Reliability

Martyn Shuttleworth 136.8K reads

Internal consistency reliability defines the consistency of the results delivered in a test, ensuring that the various items measuring the different constructs deliver consistent scores.

For example, an English test is divided into vocabulary, spelling, punctuation and grammar. The internal consistency reliability test provides a measure that each of these particular aptitudes is measured correctly and reliably.

One way of testing this is by using a test-retest method, where the same test is administered some after the initial test and the results compared.

However, this creates some problems and so many researchers prefer to measure internal consistency by including two versions of the same instrument within the same test. Our example of the English test might include two very similar questions about comma use, two about spelling and so on.

The basic principle is that the student should give the same answer to both - if they do not know how to use commas, they will get both questions wrong. A few nifty statistical manipulations will give the internal consistency reliability and allow the researcher to evaluate the reliability of the test.

There are three main techniques for measuring the internal consistency reliability, depending upon the degree, complexity and scope of the test.

They all check that the results and constructs measured by a test are correct, and the exact type used is dictated by subject, size of the data set and resources.

EXPLORABLE Quiz Time!



Quiz:
Psychology 101 Part 2



Quiz:
Psychology 101 Part 2



Quiz:
Flags in Europe

[See all quizzes =>](#)

Split-Halves Test

The split halves test for internal consistency reliability is the easiest type, and involves dividing a test into two halves.

For example, a questionnaire to measure extroversion could be divided into odd and even questions. The results from both halves are statistically analysed, and if there is weak correlation ^[1] between the two, then there is a reliability problem with the test.

The split halves test gives a measurement of in between zero and one, with one meaning a perfect correlation.

The division of the question into two sets must be random. Split halves testing was a popular way to measure reliability, because of its simplicity and speed.

However, in an age where computers can take over the laborious number crunching, scientists tend to use much more powerful tests.

Kuder-Richardson Test

The Kuder-Richardson test for internal consistency reliability is a more advanced, and slightly more complex, version of the split halves test.

In this version, the test works out the average correlation for all the possible split half combinations in a test. The Kuder-Richardson test also generates a correlation of between zero and one, with a more accurate result than the split halves test. The weakness of this approach, as with split-halves, is that the answer for each question must be a simple right or wrong answer, zero or one.

For multi-scale responses, sophisticated techniques are needed to measure internal consistency reliability.

Cronbach's Alpha Test

The Cronbach's Alpha test [2] not only averages the correlation between every possible combination of split halves, but it allows multi-level responses.

For example, a series of questions might ask the subjects to rate their response between one and five. Cronbach's Alpha gives a score of between zero and one, with 0.7 generally accepted as a sign of acceptable reliability.

The test also takes into account both the size of the sample and the number of potential responses. A 40-question test with possible ratings of 1 - 5 is seen as having more accuracy than a ten-question test with three possible levels of response.

Of course, even with Cronbach's clever methodology, which makes calculation much simpler than crunching through every possible permutation, this is still a test best left to computers and statistics spreadsheet programmes.

Summary

Internal consistency reliability is a measure of how well a test addresses different constructs and delivers reliable scores. The test-retest method [3] involves administering the same test, after a period of time, and comparing the results.

By contrast, measuring the internal consistency reliability involves measuring two different versions of the same item within the same test.

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

Links

[1] <https://explorable.com/statistical-correlation>

[2] <https://explorable.com/cronbachs-alpha>

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