



EXPLORABLE
Think Outside The Box

Published on *Explorable.com* (<https://explorable.com>)

[Home](#) > Test-Retest Reliability

Test-Retest Reliability

Martyn Shuttleworth 182.4K reads

The test-retest reliability method is one of the simplest ways of testing the stability and reliability of an instrument over time.

For example, if a group of students takes a test, you would expect them to show very similar results if they take the same test a few months later. This definition relies upon there being no confounding factor during the intervening time interval.

Instruments such as IQ tests and surveys are prime candidates for test-retest methodology, because there is little chance of people experiencing a sudden jump in IQ or suddenly changing their opinions.

On the other hand, educational tests are often not suitable, because students will learn much more information over the intervening period and show better results in the second test.

The banner features the Explorable logo and the text "Quiz Time!". Below this are three quiz cards:

- Quiz: Psychology 101 Part 2 (Image: Red roller skates on a wooden deck)
- Quiz: Psychology 101 Part 2 (Image: A fan of colorful pencils)
- Quiz: Flags in Europe (Image: A Ferris wheel at sunset)

[See all quizzes =>](#)

Test-Retest Reliability and the Ravages of Time

For example, if a group of students take a geography test just before the end of semester and one when they return to school at the beginning of the next, the tests should produce broadly the same results.

If, on the other hand, the test and retest are taken at the beginning and at the end of the

semester, it can be assumed that the intervening lessons will have improved the ability of the students. Thus, test-retest reliability will be compromised and other methods, such as split testing, are better.

Even if a test-retest reliability process is applied with no sign of intervening factors, there will always be some degree of error. There is a strong chance that subjects will remember some of the questions from the previous test and perform better.

Some subjects might just have had a bad day the first time around or they may not have taken the test seriously. For these reasons, students facing retakes of exams can expect to face different questions and a slightly tougher standard of marking to compensate.

Even in surveys [1], it is quite conceivable that there may be a big change in opinion. People may have been asked about their favourite type of bread. In the intervening period, if a bread company mounts a long and expansive advertising campaign, this is likely to influence opinion in favour of that brand. This will jeopardise the test-retest reliability and so the analysis that must be handled with caution.

Test-Retest Reliability and Confounding Factors

To give an element of quantification to the test-retest reliability, statistical tests [2] factor this into the analysis and generate a number between zero and one, with 1 being a perfect correlation between the test and the retest.

Perfection is impossible and most researchers accept a lower level, either 0.7, 0.8 or 0.9, depending upon the particular field of research.

However, this cannot remove confounding factors [3] completely, and a researcher must anticipate and address these during the research design to maintain test-retest reliability.

To dampen down the chances of a few subjects skewing the results, for whatever reason, the test for correlation [4] is much more accurate with large subject [5] groups, drowning out the extremes and providing a more accurate result.

Source URL: <https://explorable.com/test-retest-reliability?gid=1579>

Links

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

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

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

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

[5] <https://explorable.com/social-science-subjects>