

Test Validity

Martyn Shuttleworth61.2K reads

Test validity is an indicator of how much meaning can be placed upon a set of test results. In psychological and educational testing, where the importance and accuracy of tests is paramount, test validity is crucial.

Test validity incorporates a number of different validity types, including criterion validity, content validity and construct validity. If a research project scores highly in these areas, then the overall test validity is high.

Criterion Validity

Criterion validity establishes whether the test matches a certain set of abilities.

- Concurrent validity measures the test against a benchmark test, and high correlation indicates that the test has strong criterion validity.
- Predictive validity is a measure of how well a test predicts abilities, such as measuring whether a good grade point average at high school leads to good results at university.

Content Validity

Content validity establishes how well a test compares to the real world. For example, a school test of ability should reflect what is actually taught in the classroom.

Construct Validity

Construct validity is a measure of how well a test measures up to its claims. A test designed to measure depression must only measure that particular construct, not closely related ideals such as anxiety or stress.

EXPLORABLE *Quiz Time!*



Quiz:
Psychology 101 Part 2



Quiz:
Psychology 101 Part 2



Quiz:
Flags in Europe

[See all quizzes ⇒](#)

Tradition and Test Validity

This tripartite approach has been the standard for many years, but modern critics are starting to question whether this approach is accurate.

In many cases, researchers do not subdivide test validity, and see it as a single construct that requires an accumulation of evidence to support it.

Messick, in 1975, proposed that proving the validity of a test is futile, especially when it is impossible to prove that a test measures a specific construct. Constructs are so abstract that they are impossible to define, and so proving test validity by the traditional means is ultimately flawed.

Messick believed that a researcher should gather enough evidence to defend his work, and proposed six aspects that would permit this. He argued that this evidence could not justify the validity of a test, but only the validity of the test in a specific situation. He stated that this defense of a test's validity should be an ongoing process, and that any test needed to be constantly probed and questioned.

Finally, he was the first psychometrical researcher to propose that social and ethical implications of a test were an inherent part of the process, a huge paradigm shift from the accepted practices. Considering that educational tests can have a long-lasting effect on an individual, then this is a very important implication, whatever your view on the competing theories behind test validity.

This new approach does have some basis; for many years, IQ tests were regarded as practically infallible.

However, they have been used in situations vastly different from the original intention, and they are not a great indicator of intelligence, only of problem solving ability and logic.

Messick's methods certainly appear to predict these problems more satisfactorily than the traditional approach.

Which Measure of Test Validity Should I Use?

Academics are generally very resistant to change, and a huge number of educationalists and social scientists stick with the traditional methods.

Both methods have their own strengths and weaknesses, so it comes down to personal choice and what your supervisor prefers. As long as you have a strong and well-planned test design, then the test validity will follow.

Works Cited

Wainer, H. Braun, H.I. (1988) Test Validity. New Jersey: Lawrence Erlbaum Associates.

Source URL: <https://explorable.com/test-validity>