Construct Validity

Martyn Shuttleworth256.7K reads

Construct validity defines how well a test or experiment measures up to its claims. It refers to whether the operational definition of a variable actually reflect the true theoretical meaning of a concept.

The simple way of thinking about it is as a test of generalization, like external validity, but it assesses whether the variable that you are testing for is addressed by the experiment.

Construct validity is a device used almost exclusively in social sciences, psychology and education.

For example, you might design whether an educational program increases artistic ability amongst pre-school children. Construct validity is a measure of whether your research actually measures artistic ability, a slightly abstract label.

What is Construct Validity?

The term 'construct validity' can be a little misleading, because it often makes people think of how an experiment is physically constructed or designed.

A construct refers to a "theorized psychological construct".

Does the theoretical concept match up with a specific measurement [1] / scale used in research?

Construct validity refers to whether a scale or test measures the construct adequately.

An example is a measurement of the human brain, such as intelligence, level of emotion, proficiency or

ability.

Some specific examples could be language proficiency, artistic ability or level of displayed aggression, as with the Bobo Doll Experiment [2]. These concepts are abstract and theoretical, but have been observed in practice.

An example could be a doctor testing the effectiveness of painkillers on chronic back sufferers.

Every day, he asks the test subjects to rate their pain level on a scale of one to ten - pain exists, we all know that, but it has to be measured subjectively.

In this case, construct validity would test whether the doctor actually was measuring pain and not numbness, discomfort, anxiety or any other factor.

Therefore, with the definition of a construct properly defined, we can look at construct ability, a measure of how well the test measures the construct. It is a tool that allows researchers to perform a systematic analysis of how well designed their research is.

Construct validity [3] is valuable in social sciences, where there is a lot of subjectivity to concepts. Often, there is no accepted unit of measurement [1] for constructs and even fairly well known ones, such as IQ, are open to debate.

How to Measure Construct Variability?

For major and extensive research, especially in education and language studies, most researchers test the construct validity before the main research.

These pilot studies [4] establish the strength of their research and allow them to make any adjustments.

Using an educational example, such a pre-test might involve a differential groups study, where researchers obtain test results for two different groups, one with the construct and one without.

The other option is an intervention study, where a group with low scores in the construct is tested, taught the construct, and then re-measured. If there is a significant [5] difference pre and post-test, usually analyzed with simple statistical tests [6], then this proves good construct validity.

There were attempts, after the war, to devise statistical methods to test construct validity, but they were so long and complicated that they proved to be unworkable. Establishing good construct validity is a matter of experience and judgment, building up as much supporting evidence as possible.

A whole battery of statistical tools [7] and coefficients are used to prove strong construct validity, and researchers continue until they feel that they have found the balance between proving validity and practicality.

Threats to Construct Validity

There are a large number of ways in which construct validity is threatened, so here are a few of the main candidates:

Hypothesis Guessing

This threat is when the subject guesses the intent of the test and consciously, or subconsciously, alters their behavior.

For example, many psychology departments expect students to volunteer as research subjects for course credits. The danger is that the students may realize what the aims of the research [8] are, potentially evaluating the result.

It does not matter whether they guess the hypothesis [9] correctly, only that their behavior changes.

Evaluation Apprehension

This particular threat is based upon the tendency of humans to act differently when under pressure. Individual testing is notorious for bringing on an adrenalin rush, and this can improve or hinder performance.

In this respect, evaluation apprehension is related to ecological external validity [10], where it affects the process of generalization [11].

Researcher Expectancies and Bias

Researchers are only human and may give cues that influence the behavior of the subject. Humans give cues through body language, and subconsciously smiling when the subject gives a correct answer, or frowning at an undesirable response, all have an effect.

This effect can lower construct validity by clouding the effect of the actual research variable.

To reduce this effect, interaction should be kept to a minimum, and assistants should be unaware of the overall aims of the project.

See also:
Double Blind Experiment [12]
Research Bias [13]

Poor Construct Definition

Construct validity is all about semantics and labeling. Defining a construct in too broad or too narrow terms can invalidate the entire experiment.

For example, a researcher might try to use job satisfaction to define overall happiness. This is too narrow, as somebody may love their job but have an unhappy life outside the workplace. Equally, using general happiness to measure happiness at work is too broad. Many people enjoy life but still hate their work!

Mislabeling is another common definition error [14]: stating that you intend to measure depression, when you actually measure anxiety, compromises the research.

The best way to avoid this particular threat is with good planning and seeking advice before you start your research program.

Construct Confounding

This threat to construct validity occurs when other constructs mask the effects of the measured construct.

For example, self-esteem is affected by self-confidence and self-worth. The effect of these constructs needs to be incorporated into the research.

Interaction of Different Treatments

This particular threat is where more than one treatment influences the final outcome.

For example, a researcher tests an intensive counseling program as a way of helping smokers give up cigarettes. At the end of the study, the results show that 64% of the subjects successfully gave up.

Sadly, the researcher then finds that some of the subjects also used nicotine patches and gum, or electronic cigarettes. The construct validity is now too low for the results to have any meaning. Only good planning and monitoring of the subjects can prevent this.

Unreliable Scores

Variance [15] in scores is a very easy trap to fall into.

For example, an educational researcher devises an intelligence test that provides excellent results in the UK, and shows high construct validity.

However, when the test is used upon immigrant children, with English as a second language, the scores are lower.

The test measures their language ability rather than intelligence.

Mono-Operation Bias

This threat involves the independent variable [16], and is a situation where a single manipulation is used to influence a construct.

For example, a researcher may want to find out whether an anti-depression drug works. They divide patients into two groups, one given the drug and a control [17] given a placebo [18].

The problem with this is that it is limited (e.g. random sampling error [19]), and a solid design would use multi-groups given different doses.

The other option is to conduct a pre-study that calculates the optimum dose, an equally acceptable way to preserve construct validity.

Mono-Method Bias

This threat to construct validity involves the dependent variable [20], and occurs when only a single method of measurement [1] is used.

For example, in an experiment to measure self-esteem, the researcher uses a single method to determine the level of that construct, but then discovers that it actually measures self-confidence.

Using a variety of methods [21], such as questionnaires [22], self-rating, physiological tests, and observation [23] minimizes the chances of this particular threat affecting construct validity.

Don't Panic

These are just a few of the threats to construct validity, and most experts agree that there are at least 24 different types. These are the main ones, and good experimental design [24], as well as seeking feedback from experts during the planning stage, will see you avoid them.

For the 'hard' scientists, who think that social and behavioral science students have an easy time, you could not be more wrong!

Source URL: https://verify.explorable.com/construct-validity

Links
[1] https://verify.explorable.com/scientific-measurements
[2] https://verify.explorable.com/bobo-doll-experiment
[3] http://www.socialresearchmethods.net/kb/constval.php
[4] https://verify.explorable.com/pilot-study
[5] https://verify.explorable.com/significance-test
[6] https://verify.explorable.com/significance-test-2
[7] https://verify.explorable.com/statistics-tutorial
[8] https://verify.explorable.com/aims-of-research
[9] https://verify.explorable.com/research-hypothesis
[10] https://verify.explorable.com/ecological-validity
[11] https://verify.explorable.com/what-is-generalization
[12] https://verify.explorable.com/double-blind-experiment
[13] https://verify.explorable.com/research-bias
[14] https://verify.explorable.com/defining-a-research-problem
[15] https://verify.explorable.com/statistical-variance
[16] https://verify.explorable.com/independent-variable
[17] https://verify.explorable.com/scientific-control-group
[18] https://verify.explorable.com/placebo-effect
[19] https://verify.explorable.com/random-sampling-error
[20] https://verify.explorable.com/dependent-variable
[21] https://verify.explorable.com/research-designs
[22] https://verify.explorable.com/survey-research-design
[23] https://verify.explorable.com/observational-study
[24] https://verify.explorable.com/true-experimental-design