

Assignment No: 02

Subject Advance Research Methods.

To: Dr. Farooq Jan Sahib.

From: Pervez Khan

Reg No: 16974

Class: MS Finance

Semester: 1st

Q.1) Validity Types with Explanation? -

Ans. Validity tells you how accurately a method measures something. If a method measure, and the results closely correspond to real world values, then it can be said to have validity. There are four main types of validity.

1) Construct validity:

Does the test measure the concept that it's intended to measure?

2) Content validity:

Is the test fully representative of what it aims to measure?

3) Face validity:

Does the content of the test appear to be suitable to its aim?

P.T.O

4

Criterion validity:

Do the result correspond to a different test of the same thing. Here we determine the accuracy of the actual components of a measure. If you are doing experimental research, you also need to consider internal and external validity, which deal with the experimental design and the generalizability of results.

1

Construct validity:

Construct validity evaluates whether a measurement truly really represents the thing we are interested in measuring. It's central to establishing the overall validity of a method - what is construct validity?

A construct refers to a concept or characteristic that can't be directly observed, but be measured by observing other indicators that are associated with it.

Construct can be characteristics of individuals, such as intelligence, obesity, job satisfaction or depression. They can also be broader concepts applied to organization or social groups, such as gender equality, corporate social responsibility, or freedom of speech. e.g. there is no objective, observable entity called "depression" that we can measure directly -

P.T.O

But base on existing Psychological research and Theory, we can measure depression based on a Collection of symptoms and indicators, such as low self confidence and low energy levels-

→ what is Construct validity?

Construct validity is about ensuring that the method of measurement matches the construct you want to measure. if you develop a questionnaire to diagnose depression you need to know. does the questionnaire really measure the construct of depression, or is it actually measuring the respondent mood self esteem or some other construct- To achieve construct validity you have to ensure that your indicators and measurements developed based on relevant existing knowledge. The questionnaire must include only relevant questions that measure known indicators of depression.

The other types of validity described below can all be considered as forms of evidence for construct validity.

2) Content Validity:

Content validity assesses whether a test is representative of all aspects of the construct.

To produce valid results, the content of a test, survey or measurement method must cover all relevant

P.T.O

Parts of the subject it aims to measure. If some aspects are missing from the measurement (or if irrelevant aspects are included), the validity is threatened -

e.g. A mathematics teacher develops an end of semester algebra test for her class. The test should cover every form of algebra that was taught in the class. If some types of algebra are left out, then the results may not be an accurate indication of students' understanding of the subject. Similarly if she includes questions that are not related to algebra, the results are no longer a valid measure of algebra knowledge.

③ Face validity:

Face validity considers how suitable the content of a test seems to be on the surface. It's similar to content validity, but face validity is a more informal and subjective assessment.

e.g. you create a survey to measure the regularity of people's dietary habit. You review the survey items which ask questions about every meal of the day and snacks eaten in between for every day of the week.

on its surface the survey seems like a good representation of what you want to test, so you consider it to have high face validity.

P.T.O

(5)

As face validity is subjective measure, it is often considered the weakest form of validity. It can be useful in the initial stages of developing a method.

(4) Criterion validity:

Criterion validity evaluates how closely the results of your tests correspond to the result of a different test. What is a criterion?

The criterion is an external measurement of the same thing. It is usually an established or widely used test that is already considered valid.

What is criterion validity?

To evaluate criterion validity, you calculate the correlation between the results of your measurement and the results of the criterion measurement. If there is a high correlation, this gives a good indication that your test is measuring what it intends to measure. e.g. A university professor creates a new test to measure applicants' English writing ability. To assess how well the test really does measure students' writing ability, she finds an existing test that is considered a valid measurement of English writing ability, and compares the results when the same group of students take both tests. If the outcomes are very similar, the new test has a high criterion validity.

Q
Ans

Reliability Types with Explanation?
Reliability tells you how consistently a method measure something - when you apply the same method of measurement to the same sample under the same conditions. you should get the same results. if not the method of measurement may unreliable -

There are four main types of reliability. Each can be estimated by comparing different sets of results produced by the same method -

- ① Test-retest - the same test over time -
- ② Interrater - the same test conducted by different people -
- ③ Parallel forms - different versions of a test which are designed to be equivalent
- ④ Internal consistency - the individual items of a test.

Test retest reliability:

Test retest reliability measure the consistency of results when you repeat the same test on the same sample at a different point in time. you use it when you are measuring something that you expect to stay constant in your sample -

A test of colour blindness for trainee Pilot applicants should have high test retest reliability, because colour blindness is a trait that does not change over time -

P.T.O

Why its important?

Many factors can influence your results at different point in time. e.g respondent might experience different moods, or external conditions might affect their ability to respond accurately.

Test-retest reliability can be used to assess how well a method resists these factors over time. The smaller the difference between the two sets of results the higher the test retest reliability.

How to measure it:

to measure test retest reliability you conduct the same test on the same group of people at two different point in time. Then you calculate the correlation between the two sets of results.

Test retest reliability example.

You devise questionnaire to measure the IQ of a group of participants. (a property that is unlikely to change significantly over time). you administer the test two months apart to the same group of people - but the results are significantly different so the test retest reliability of the IQ questionnaire is low.

=> Improving test retest reliability:

when designing tests or questionnaires try to formulate questions, statements and tables in a way that won't be influenced by the mood or concentration of participants. when planning your methods of data collection.

P.T.O

try to minimize the influence of external factors and make sure all samples are tested under the same conditions -

Remember that changes can be expected to occur in the participants over time and take these into account.

⇒ Interrater reliability:

Interrater reliability (also called interobserver reliability) measures the degree of agreement between different people observing or assessing the same things. You use it when data is collected by researchers assigning ratings, scores or categories to one or more variables -

In an observational study where a team of researchers should agree on how to categorize or rate different types of behavior.

⇒ Why it's important:

People are subjective so different observers, perceptions of situations and phenomena naturally differ. A valid research aims to minimize subjectivity as much as possible so that a different researcher could replicate the same results.

When designing the scale and criteria for data collection, it's important to make sure that different people will rate the same variable consistently with the minimal bias. This is especially important when there are multiple researchers involved in data collection or analysis -

⇒ How to measure it:

To measure interrater reliability different researchers conduct the same measurement or observation on the same sample. Then you calculate the correlation between their different sets of results. If all the researchers give similar ratings, the test has high interrater reliability.

⇒ Interrater reliability Example:

A team of researchers observe the progress of wound healing in patients. To record the stages of healing rating scales are used with a set of criteria to assess various aspects of wounds. The results of different research assessing the same set of patient are compared and there is a strong correlation between all of results. So the tests has high interrater reliability.

⇒ Improving interrater reliability:

Clearly define your variables and the methods that will be used to measure them. Develop detailed, objective criteria for how the variable will be rated, counted or categorized. If the multiple researchers are involved ensure that they all have exactly the same information and training.

⇒ Parallel forms reliability:

Parallel forms reliability measure correlation between two

D.T.O

equivalent versions of a test. You use it when you have two different assessment tools or sets of questions designed to measure the same thing.

⇒ Why it's important:

If you want to use multiple different versions of a test (e.g. to avoid respondents repeating the same answers from memory) you first need to make sure that all the sets of questions or measurement give reliable results.

In educational assessment, it is often necessary to create different versions of test to ensure that students don't have access to the question in advance. Parallel forms reliability means that if the same students take two different versions of a reading comprehension test, they should get similar results in both tests.

⇒ How to measure it:

The most common way to parallel forms reliability is to produce a large set of questions to evaluate the same thing. Then divide these randomly into two question sets.

The same group of respondents answer both sets, and you calculate the correlation between the two. A high correlation indicates high parallel forms reliability.

P.T.O

=> Parallel forms reliability example:

A set of questions is formulated to measure financial risk aversion in a group of respondents. The questions are randomly divided into two sets, and the respondents are randomly divided into two groups. On both tests, group A takes test A first and group B takes test B first. The results of the two tests are compared and the results are almost identical, indicating high parallel forms reliability.

=> Improving parallel forms reliability:

Ensure that all questions or test items are based on the same theory and formulated to measure the same things.

=> Internal consistency:

Internal consistency assesses the correlation between multiple items in a test that are intended to measure the same construct.

You can calculate internal consistency without repeating the test or involving other researchers, so it's a good way of assessing reliability when you only have one data set.

=> why its important:

When you devise a set of questions or ratings that will be combined into an overall score you have to make sure

P.T.O

That all of the items really do reflect the same thing - if response to different items contradict one another, the test might be unreliable.

To measure customer satisfaction with an online store, you could create a questionnaire with a set of statements that respondents must agree or disagree with, internal consistency tells you whether the statements are all reliable indicators of customer satisfactions.

⇒ How to measure it:

Two common methods are used to measure internal consistency. you calculate the correlation between the results of all possible pairs of items and then calculate the average.

⇒ Split-half reliability:

you randomly split a set of measure into two sets - after testing the entire set on the respondent you calculate the correlation between the two sets of response.

⇒ Internal Consistency Example:

A group of respondents are presented with a set of statements designed to measure optimistic and pessimistic mind sets. They must rate their agreement with each statement on a scale from 1 to 5. if test is internally consistent an optimistic respondent should generally give high rating to optimism indicators and low

ratings to pessimism indicators. The correlation is calculated between all the responses to the optimistic statements, but the correlation is very weak. This suggests that the tests has low internal consistency.

⇒ Improving internal consistency:

Take care when dividing questions or measures. Those intended to reflect the same concept should be based on the same theory and carefully formulated.

∩ The End ∩