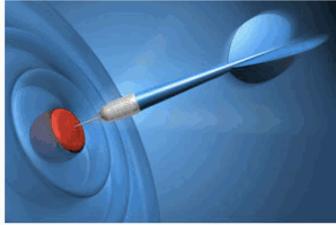


Reliability in Research Design

3 years ago 2 By Tariq



When I think of reliability I imagine always knowing what to expect. If a person is able to produce the same quality work consistently then they are considered reliable. You see it in sports all the time. Certain players have a knack for coming through in key situations no matter how late in the season or how worn down they are. However, I can imagine few jobs that require more reliability than a surgeon. Having an off day for them could prove disastrous. For a measure to be reliable it must demonstrate consistency as well as repeatability. When carrying out research our results should be accurate across a range of measurements. In surveys you would like to think that you would get the same response no matter what mood your respondent is in but that is not always the case. A surgeon must deal with difficult situations while showing the same precision

and reliability. That is a quality to be admired but you can not always expect everyone to act like a surgeon at all times. It's also possible that your respondent won't know what you mean when you ask them a certain question resulting in an answer that is entirely different from what you are attempting to measure.

Test-retest Reliability

If your survey respondent had to take your survey again would they answer the same questions the same way? Test-retest reliability measures reliability over time. A number of factors can affect reliability over time such as a person's mood, time of the day, where the questions are placed in the survey (context effect), circumstantial events, vagueness, etc. A good test will take into account factors that may influence survey results over time and minimize them so that results show little variation. If a test is unreliable then any one of a number of factors can lead to varying results depending on when the question is asked. In general the more time a person takes between retesting the more variation you can expect in the responses.

If you ask Joe Q what he thinks about Candidate X on Tuesday he may view him favorably because X gave a really good speech on Monday. Say Candidate X is indicted later in the week in a corruption scandal. Joe previously indicated that a candidate's integrity is very important. Last week he said that he was leaning toward Candidate X. Now that Candidate X has been exposed you may think he is likely to give you a different response if you asked Joe the same question next Tuesday. The reliability of opinion polls can be doubtful depending on the questions we ask because opinions tend to fluctuate over time. What does Joe Q mean when he rates integrity as very important? Perhaps Joe Q considers anyone that shares his ideology to have integrity. It's possible that Joe Q would vote for Candidate X no matter what he thought of him personally because they share the same ideology. Probing Joe's past voting record would be more indicative of voter preference than asking a subjective question about integrity. Asking him more objective questions that would not fluctuate from week to week would have higher test-retest reliability.

Parallel Forms Reliability

Another challenge reliability faces is in knowing what the best questions to ask are. What does Joe Q mean when he rates integrity as very important? Could we come up with better questions to predict how voters like Joe Q would vote? Another way to improve the reliability of a survey is to ensure that it is representative of the data you are trying to collect. To do this increase the sample size. If you are gathering research to find out whether voters like Joe Q are likely to vote for Candidate X then you need to find more people like Joe and ask them different questions or question sets based on the same construct.

You come up with a large set of questions to ask in your survey. The construct that you are measuring is voter preference. The large question set is split in half and you administer each set to half of the targeted population. You can then take a look at which questions are better indicators of voter preference. This combines what is known as a split test method with parallel form evaluation.

integrity as very important? Could we come up with better questions to predict how voters like Joe Q would vote? Another way to improve the reliability of a survey is to ensure that it is representative of the data you are trying to collect. To do this increase the sample size. If you are gathering research to find out whether voters like Joe Q are likely to vote for Candidate X then you need to find more people like Joe and ask them different questions or question sets based on the same construct.

You come up with a large set of questions to ask in your survey. The construct that you are measuring is voter preference. The large question set is split in half and you administer each set to half of the targeted population. You can then take a look at which questions are better indicators of voter preference. This combines what is known as a split test method with parallel form evaluation.

You can use parallel forms to measure a construct for people that are not like Joe Q. Here you would divide a population that is representative of all likely voters in two. Develop a large question set that measures a particular construct and then administer to each half of your representative population. Now you can learn which questions are better indicators for voter preference for a representative population.

Inter-rater Reliability

This is necessary if you are conducting your survey using an interview process. If multiple people are interviewing Joe Q to ask what his opinion on politics is then inter-rater reliability measures the degree to which the observers agree. This is the best way to measure reliability if you are using observation for your research.

Internal Consistency Reliability

The purpose of asking questions in surveys is to assess a particular construct or idea. Therefore different questions that measure the same construct should yield similar results. Reliability is determined on the basis of whether results are consistent for different items that measure the same construct. For example, you could check for reliability on your survey by asking a respondent two similar questions meant to measure the same thing.

- *Average Inter-Item Correlation* – when we ask a respondent two similar questions to measure the same construct. This compares correlations between this and any other paired questions to measure the same construct by calculating the mean of all paired comparisons.
- *Average Itemtotal Correlation* – where you take the average inter-item correlation and calculate a total score for each item.
- *Split-half Correlation* – you divide items that measure the same construct into two tests, apply them to the same group of people, and calculate the correlation between the two scores.
- *Cronbach's Alpha* – when we calculate the average split half estimates from a sample population.

In order to draw conclusions, formulate theories, or make generalizations about your research you need to ensure the reliability of the data you collect. In general reliability is threatened when assessments are taken over time, rely on different standards of judgment, or assessments are highly subjective. You can improve reliability by ensuring that your surveys are written clearly and without ambiguity. You should construct your response options so that they are appropriate and meaningful.