

8.7: Scale Reliability

Learning Objectives

At the end of this section you should be able to answer the following questions:

- How would you explain *internal consistency* reliability?
- What is a good range for alpha?

When researchers have a final set of items that form a functional scale, reliability calculations for all the scale items must be conducted by analysing the total number of items as a set.

This is most often done by using a type of reliability called *internal consistency* reliability, which is based on formulas that give an index of how much variability is shared and accounted for by the set of items, thus reflecting their degree of interrelationship. High internal consistency reliability reflects that items are consistent with other items in the set, and that the items are measuring the same construct.

Cronbach's alpha is a common statistic used to measure internal consistency, and it measures the correlation between multiple items in a factor. When using Cronbach's alpha, it is important to make sure that all items are related and measured in a similar way – but not with exact similarity in wording or in regard to aspects of the construct measured.

If dealing with multiple constructs or factors that cluster within a higher order factor, Cronbach's alpha should be run for both the total scale, and the items in each factor.

When conducting EFA procedures, scale reliability should be tested for each factor following the last EFA and the finalisation of the factor structure.

PowerPoint: Alpha Output

Please have a look at the link below for SPSS Output on Internal Consistency calculations:

- [Chapter Eight – Alpha Output](#)

As can be seen here, the alpha for the total scale used is 'good' as seen by a .85 value. You can also check and see if some items were deleted, would the alpha improve. If removing an item improves the score by .01-.02 it might be worth removing the item.

In general, "good" alpha estimates range from .7 – .9 (George & Mallery, 2003), with the following interpretations:

<.50 = Unacceptable

.51-.60 = Poor

.61-.70 = Questionable

.71-.80 = Acceptable

.81-.90 = Good

.91-.95 = Excellent

If the alpha is greater than .95, it is likely that there are a number of items that ask very similar, or the same question. For example, "I often feel down" and "I am often down".

This page titled [8.7: Scale Reliability](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Erich C Fein](#), [John Gilmour](#), [Tayna Machin](#), and [Liam Hendry](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.