## Beth Morling (2017): Research Methods in Psychology: Evaluating a World of Information, 3rd ed., Norton

**p. 131**

Researchers typically will run a correlation-based statistic called **Cronbach’s alpha** (or *coefficient alpha*) to see if their measurement scales have internal reliability.First, they collect data on the scale from a large sample of participants, and thenthey compute all possible correlations among the items. The formula for Cronbach’salpha returns one number, computed from the average of the inter-item correlationsand the number of items in the scale. The closer the Cronbach’s alpha is to 1.0, thebetter the scale’s reliability. (2.) If Cronbach’s alpha is high, there is good internalreliability and researchers can sum all the items together. If Cronbach’s alpha is lessthan .70, then internal reliability is poor and the researchers are not justified in combiningall the items into one scale. They have to go back and revise the items, or theymight select only those items that were found to correlate strongly with one another.

## Dennis Howitt & Duncan Cramer (2020): Research Methods in Psychology, 6th ed., Pearson

**p.241**

However, internal consistency is most often measured using Cronbach’s alpha (Cronbach, 1951), which is a more comprehensive index than split-half reliability. (It calculates the average reliability based on all possible ways of splitting the scale or measure into two halves.) Generally speaking, a reliability of about .70 would be regarded as satisfactory (Nunnally, 1978).

## Hugh Coolican (2018): Research Methods and Statistics in Psychology, 7th ed., Routledge

**p. 236**

Cronbach’s alpha is probably the most commonly used statistic for estimating a test’s reliability. It depends largely on how people vary on individual items. If they tend to vary a lot on the individual items relative to how much they vary overall on the test, then a low value for alpha is achieved and the test is assessed as unreliable.

(If the items in the scale are *dichotomous* [answers are bi-polar, e.g., ‘yes’ or ‘no’ only], then a simpler version is used, known as the Kuder–Richardson measure.) Cronbach’s alpha is equivalent to the average of all possible split-half reliability values that could be calculated on the data set. Good reliability, therefore, is represented with alpha values from around .75 up to 1. However, there is debate about very high levels of alpha as just alluded to in mentioning Cattell’s objections. The basic issue is that, if we get an alpha of .95 or above it suggests that all items are measuring much the same thing and therefore you might as well present just one item.

## Glynis M. Breakwell, Jonathan A. Smith & Daniel B. Wright (2012): Research Methods in Psychology, 4th ed., SAGE

**p. 149**

Coefficient alpha (Cronbach’s alpha) is frequently used in preference to split-half reliability. Cronbach’s alpha is a functional average of all possible split-half reliabilities, and so long as all possible split halves are approximately parallel this works well. Two caveats; first, this is not the case if some of the splits divide meaningful areas of the test specification; for example, all arithmetic items in one half, all geometry items in the other. Second, it is possible, serendipitously but counterproductively, to increase alpha by limiting items to narrower and narrower areas of the test specification, thus excluding important aspects of the desired trait.

What counts as a ‘good’ reliability depends on the circumstances. It is very unusual for an assessment or psychometric test to have a reliability above 0.96; only individual broad spectrum tests of general ability achieve this level. As a rule of thumb, group tests of ability should have reliabilities over 0.8, tests of personality over 0.7, essay marks over 0.6 and interviews over 0.5. Constructs that are more difficult to measure such as creativity, or projective tests such as the Rorscharch inkblot test, rarely have reliabilities greater than 0.3.

## Ronald A. McQueen & Christina Knussen (2013): Introduction to Research Methods and Statistics in Psychology: A practical guide to the undergraduate researcher, 2nd ed., Pearson

**p. 389**

The extent to which all the items in a questionnaire are deemed to be assessing the same things is termed **internal consistency**, and the measure found in much of the literature is known as Cronbach’s alpha.

[…]

Today, almost every journal article describing research involving test items will quote a measure of Cronbach’s alpha (or just alpha) as a measure of reliability. The measure itself is the coefficient of the correlation among the items in a test, and will therefore be expressed as a value between 0 and 1. a good measure of internal consistency will be expressed by an alpha value in the range 0.7 to 0.9 (roughly). Values below 0.7 would suggest that not all items are measuring the same issue, while values greater than 0.9 probably indicate that some items are too similar in what they are measuring and are therefore redundant.