

- a. Errors due to inadequate sampling of contents indicating that the items for the two tests may have been slightly different.
- b. Errors due to chance response tendencies indicating guessing or random responding to the items.
- c. Errors due to changes in the testing environment indicating the effect of physical conditions such as light, temperature, time of the day, etc.
- d. Errors due to the changes in the person taking the test indicating the influence of health, fatigue, mood, and practice, etc.

Reliability of any selection tool and tests in particular is determined by using any of the following methods.

1. Test-retest Reliability. It refers to giving the same test to the same people in identical conditions after a lapse of time. The time interval has to be such that neither it is too long to bring about the influence of the changes in the individual nor too short for him to use previous exposure to the same test as a facilitator in the second exposure. The latter is called “practice effect”. Some people may remember certain items and locate the correct answer to do well in the subsequent trial.

2. Alternate Form Reliability. The second method is an improvement over the first. It consists of the use of identical tests on the same population at the same time. The two tests are not exactly alike yet they measure the same ability by different items of equal difficulty levels. In this situation, the testing conditions, number of items, time limit, and instructions, etc. remain exactly the same. The only differences is in the items covering the same contents.

3. Split-half Reliability. A single test is administered. For purposes of computing reliability, the test items are divided in two halves usually on the basis of even-odd serial numbers. The two halves constitute two sub-parts of the test.

4. Kuder-Richardson Reliability. It consists of administering a single test to the candidate. After the test is finished, each item is compared with other items of the test to see if responses on these items show inter-item consistency (Anastasi, 1968). The more homogeneous the contents of the test, the higher would be the inter-item consistency.

For the first three methods of determining reliability, Pearson Product-Moment Correlation of Coefficient* is calculated. However, for Kuder-Richardson reliability the method of computation is slightly different. The formula is called “Kuder-Richardson formula 20”.

$$\gamma_{11} = \left(\frac{n}{n-1} \right) \frac{\sigma_t^2 - \Sigma pq}{\sigma_t^2}$$

In this formula γ_{11} = reliability coefficient of the whole test; n =number of items in the test; σ_t = standard deviation of total scores on the test; and Σpq = total of the product of the proportion of persons who pass (p) and proportion who do not (q) each item.

* See Chapter 16 for details and methods of calculating correlation of coefficient.