SMU Medical Journal

ISSN : 2349 – 1604 (Volume – 4, No. 1, January 2017) Review Article

Indexed in SIS (USA), ASI (Germany), I2OR & i-Scholar (India), SJIF (Morocco) and Cosmos Foundation (Germany) databases. Impact Factor: 3.835 (SJIF)

Reliability and Validity of a Psychological Test

Rajkumari Sanatombi Devi
Assistant Professor (Statistics), Statistics and Epidemiology Cell, Department of Community Medicine, Sikkim Manipal Institute of Medical Sciences, 5th Mile, Tadong, Gangtok. Sikkim

Manuscript received : 29.11.2016
Manuscript accepted: 27.12.2016

Abstract

Reliability and Validity of a test are the two important characteristics of a research tool. These two terms are mainly concerned with test and questionnaire. The most efficient way to understand any test is to examine how the test is constructed and how it is administered to the subject of the study population. A standardized procedure should be followed while administering the test. The evaluation of the test is primarily done by estimating the Reliability and Validity of the test items with the help of Correlation Coefficient which are quantified into numerical scores.

Reliability refers to the consistency of measurements of the test. It is used as the index of the chance of getting the same results when the test is repeated on different occasion with the same subjects. Validity of a test is primarily concerned with the extent to which the test measure

what is intended to measure. It is determined by estimating the Correlation Coefficient between the test scores and some criterion. Validity is an important component of having a reliable estimate of the test scores. It is therefore necessary to decide before administering any test about the purpose, the content and the characteristics of the subjects for whom the test is intended to measure.

Keywords: Psychological test, Validity, Reliability, Coefficient of Correlation

Introduction

According to Cronbach a psychological test is a systematic procedure for comparing the behavior of two or more persons at a particular time or one or more persons at different time periods. A good psychological test must possess the following 8 characteristics, viz,

1. Reliability
2. Validity
3. Objectivity
4. Discriminating power
5. Adequacy
6. Usability
7. Comparability
8. Utility

Here we describe two of the above characteristics, viz. Reliability and Validity which are essential characteristics of a good test. The term Reliability generally refers to the degree of agreement between the scores each time the test is used with the same subjects. Provided the characteristics being measured remain constant, the result should be the same. However, in real situation one will never get the same results if one repeats the measurement of any attribute of physical, biological or psychological phenomena of the subjects. Error due to measurement always exists due to chance factor and it can't be avoided in any test. Therefore, a standardized procedure should be followed while testing the test which may be within the test itself and between the same tests administered at different time periods. The relationship between the test scores is expressed into a numerical score known as Correlation Coefficients where the values should range between 0.00 to +- 1.00. A test is said to be valid if it measures what it is supposed

to measure. It concerns the extent to which a test measures what it is intended to measure. It is determined by finding the Correlation Coefficient value between the test scores and some selected criterion. Validity of a test is required before checking the reliability of the test. A reliable test might or might not be valid. However if a test is not valid then there is no point of discussing about the reliability of the test. Therefore, the content of the selected items should be correct and accurate to the extent that it should support the goals and objectives of the study.

In brief we can say that Reliability means

- the consistency among the test scores.
- the correlation between 2 or more sets of scores on equivalent tests from the same groups of individuals.
- it gives the same results when tried on the same subject under similar condition.
- In its broader sense, "Reliability indicates the extent to which individual differences in test scores are attributed to "true" differences in the characteristics under consideration and the extent to which they are attributable to chance error.

Definition of Reliability

- According to Ghiselli, Campbell and Zedeck **"**Reliability is simply the extent of unsystematic variation in the quantitative description of the amount of some traits an individual possess or manifest when it is measured."
- Nunally defined "Reliability as the precision of measurement regardless of what its measure."

Methods for measuring Reliability

Test –Retest method/Co-Efficient of Stability

1. It means the repetition of the same test on the same individual at different time period.

2. Same test is administered twice to the same individual at different occasion.
3. The test is administered to a group of individual for the first time, known as testing situation
4. After a gap of time say for a week the same test is re-administered to the same individual known as retesting situation.
5. After calculating the co-efficient of Correlation between the two sets of score, an estimate of the consistency of measurement called the reliability coefficient is obtained.
6. If the index of agreement of the relationship comes out to be high enough, we interpret the test is a reliable one.
7. In successive measurement that is in retesting situation if all individual get the same score as they obtained in the first testing or if the entire individual obtained higher or lower scores and their rank remain the same, the test can be called a reliable test.

Parallel form method or Equivalence form method

1. Here two forms A and B are given to the same groups of individuals. Here the second form B should not be given immediately after the first form. The two forms must not be alike yet the test must match the number of items, uniformity in contents, range and levels of difficulty and format.
2. It is administered twice..

Split-half method

1. Sometimes it may not be possible to administer the test twice or there may not be equivalent form available, in such cases we use the Split-half method.
2. In this method, the score is divided into two equivalent halves. A common method of

splitting the scores into halves is to split the scores into odd number items and even number items separately.

3. After that the correlation between odd and even items is calculated.
4. If the degree of correlation is high, the test may be regarded as reliable.
5. The test is more reliable when the test is long and the items are placed in order of difficulty from the least to the most difficult.
6. It assumes that the test is homogeneous. It involves the characteristics of stability and equivalence.
7. Spearman Brown Formula for reliability of used in Split half test: rt = 2r/1+r, Where rt = reliability co-efficient of the whole test, r = Correlation between the two series of scores

Kuder Richardson reliability test

The formula is given as rt = n/n-1 x (σ^2– pq/ σ^2), rt= reliability coefficient, P= proportion of right responses, q= proportion of wrong responses, σ = variance of the test score

Crobanch's alpha test

It is a measure of internal consistency that is how closely a set of items are related as a group. It is written as a function of the number of test items and the avarage intercorrelation among the items.

α = N.C/ γ + (N -1).C, where N =no. of items, c = average inter item Covariance among the items, γ is the average variance.

Characteristics of a Reliable test

- Accuracy
- Consistency
- Truthfulness

- Precision
- Stability
- Dependability

Validity of a test

Reliability is a condition for validity. It is the most critical criterion for reliability of the test. Validity indicates the degree to which an item measures what it is supposed to measure. Validity is the extent to which difference found with a measuring instrument reflects "True" differences among those being tested.

Definition of Validity of a test

According to Annastasi, "Validity of a test is concern what the test measure and how well it does so." Cronbach defined "A test is valid, to the degree that we know what it measure or predict."

For determining the Validity of a test, the test must be compared with some ideal independent measure. Validity is reported as a Correlation Co-efficient between the test and some external criterion. The Correlation of Co-efficient computed between the test and an ideal measure of Criteria is known as Validity Coefficient.

Types of Validity

Face Validity

It is not scientific. By seeing the appearance of the test we determine the validity of the test. A test is said to have face validity when by appearance it looks like it measures what it is meant and intended to measure. Here the judgmental process is used to determine and search whether the test appears to correspond to that of the question. Face validity is evaluated by a group of judges, some experts who read or look at a measuring technique and decide whether in their opinion it measures what its name suggests. Face validity is probably the easiest validation procedure to explain but the most difficult to carry out in the course of actual field work.

Content Validity

Content Validity is concerned with the scores or range of items used to measure the variable. It is the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. It can be determined by using a panel of person who shall judge how well the measuring instrument meets the standard of the test items but there is no numerical way to express it, it is a logical method.

Criterion Related Validity

Criterion Related Validity considers the degree to which an instrument correlates with some criterion measures of the variable of interest. It is the ability of the test to predict some outcome or to estimate the existence of current condition. It is of two types.

1. Predictive Validity: It is the ability of the test to identify future differences.

2. Concurrent Validity: Concurrent validity is the degree to which scores on a test are related to the scores on another , already established test administered at the same time or to some other valid criterion available at the same time. It is employed for diagnosis of existing status rather than prediction of future success. This procedure is useful only if the test used as criterion is an accepted one. Criterion Related Validity is expressed as the Co-efficient of Correlation between test scores or some measures of future performance or between test scores and scores on another measure of known validity.

Relationship between Reliability and Validity of a test

Reliability and Validity is the two dimension of the same test. Both measure the efficiency of the test. A test which has poor reliability is not expected to have low validity. Reliability is a precondition to Validity but the reverse is not so. Reliability is not a matter of degree but validity is a matter of degree. A highly reliable test may not be a valid test but a highly valid test can't be unreliable. High reliability requires items of equal difficulty and high

inter-correlation between the items whereas high validity requires items of different difficulty and low inter-correlation among the items. A test constructor should not always aim at having high reliability and validity in the same test because sometimes the goals of reliability and validity are incompatible.

Factors influencing the Reliability and Validity of a test

These factors may be categorized into 2 heads

1. **Extrinsic or External factors:** Extrinsic factors are those factors which lie outside the test itself and tend to make the test reliable or unreliable. Some of them are as follow:

   - **Guessing:** It is an important source of reliability. In two alternative response systems, there is 50% chance of answering an item correctly on the basis of guessing. In multiple choice tests, it is reduced.

   - **Environmental condition**: Environmental condition such as light, sound etc are equal and uniform to all the respondents. Otherwise it will reduce the reliability of the test scores. Outside disturbance also reduce the reliability.

   - Cheating, illness, worry, duration of time, inadequacy, and faulty direction lowers the reliability.

2. **Intrinsic or internal factors:** Faulty items (tricky and confusing style of asking a question), poorly constructed items, excessively too difficult, easy items should be eliminated. Used of inappropriate items, length of the test (time interval between the test and re-test), homogeneity (in the subject matter) of the test, discriminating item/ ambiguous and ill structured item, interdependent items, irregularity of respondent/no clear cut in instruction, irregularity / inconsistency in test administration, method of obtaining data, faulty scoring procedure are factors affecting the Reliability and Validity of a test.

Increased Reliability of a test

Use technically correct, unambiguous items, standardized the administrative procedure and standardized the scoring procedure. Be alert for respondent irregularities. Make the test long enough to include a good sample of items. Be certain each item on the test measure the same outcome or set of outcomes. Construct items of an appropriate level of difficulty.

Reliability level

Interpretation of r (Coefficient of Correlation):

- r+- =0.25= low coefficient of reliability
- r+-= 0.50= Medium
- r+- = 0.75 = Highly coefficient
- r+- = 1.0 = perfect positive/ perfect negative

When properly applied, the use of Valid and Reliable assessment instruments will help us to make better decisions. Additionally, by using a variety of assessment tools as part of an assessment program, we can more fully assess the skills and capabilities of people, while reducing the effects of errors associated with any one tool on our decision making.

References

[1] Garrett, H. E (1 976). Statistics in psychology and education. Bombay: Vakils. Feffer and Simons Ltd.

[2] Best, J.W., & Kahn, J.V. (1999). Research in Education (7th Edition). NewDelhi: Prentice-Hall of India Private Limited.

[3] Guilford, J. P. & Benjamin Fruchter. (1981).Fundamental statistics in psychology and education (9th Edition). New York, NY, US: McGraw-Hill.

[4] Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.). Educational Measurement (2nd edition). Washington, D. C.: American Council on Education.

[5] Nunnally, Jum, & Bernstein, Ira H. (1994). Psychometric theory (3rd Edition.). New York: McGraw-Hil.

[6] Anastasi, Anne. (1982). Psychological testing (5th Edition.). New York: Macmillan.

[7] Ghiselli, Edwin E., Campbell, John P., & Zedeck, Sheldon. (1981). Measurement theory for the behavioral sciences. San Francisco: Freeman.

[8] Fernald / Fernald. (2007). Munn's Introduction to Psychology (5th Edition). A.I.T.B.S Publisher, Delhi., India.

Authors Column

Dr. Rajkumari Sanatombi Devi, Ph.D, Msc.(Statistics), M.Ed, M.P.S, PGDC is presently acting as Assistant Professor (Statistics) in the Department of Community Medicine, Sikkim Manipal Institute of Medical Sciences under the Sikkim Manipal University. Her present field of research is Epidemiology and Public Health. She attended few conferences and published nine papers in National & International Research Journals of repute.

SMU Medical Journal, Volume – 4, No. – 1, January, 2017, PP. 185 - 194
© SMU Medical Journal