RESEARCH JOURNAL OF FISHERIES AND HYDROBIOLOGY

ISSN:1816-9112
Open Access Journal

A Review on Internal Consistency Measures for Medical Sciences Research

Syed Jawad Ali Shah and Qamruz Zaman

Department of Statistics, University of Peshawar, Pakistan.

Address For Correspondence:

Syed Jawad Ali Shah
Department of Statistics,
University of Peshawar,
Pakistan.

Received: 3 November 2015
Accepted: 20 December 2015
Available online: 30 December 2015

ABSTRACT

In medical sciences research, (specially research based on clinical trials), Likert scale & summated scale are often used in survey instruments. The most important step is to assess the Reliability and Validity of instrument using different approaches. Most of the researchers merely rely on Cronbach's Alpha for calculating internal consistency, and they have the desire to get higher level of Alpha. But higher value of Cronbach's Alpha decreases the validity of the instrument, and also not applicable at every situation due to violation of its assumptions. There are various other internal consistency measures such as angoff's coefficient, Raju's coefficient, Kristoff's coefficient and Guttman's coefficient etc. In this article, a brief description of each of these coefficients is illustrated and at the end, recommendations for proper use at best suitable situation of these coefficients are presented.

KEY WORDS:

INTRODUCTION

Reliability and Validity are two essential concepts in the evaluation of a measurement instrument. In medical sciences research, instruments can be survey questionnaires, conventional knowledge, skill or attitude tests or clinical simulations. The concepts involved in the instrument are sometimes cannot be directly observable (e.g. severity of pain). Most of the times, the attitudes, feeling of pains, opinions about use of a drug etc. involves the use of Likert-type scales. As individuals attempt to quantify latent variables (i.e. the variables which can't be directly measured), researchers generally use multiple-item scales and summated ratings for this purpose. The Likert (1932) approach is often referred to as a "summated rating scale", because the responses of individuals received for each item in a construct are summed (or averaged) to obtain the individual's total score.

Reliability is concerned with the ability of an instrument to measure consistently, while Validity is concerned with the extent to which an instrument is supposed to measure. Cortina (1993) established a principle that use of any reliability approach depends on the source of variance. If the interest of researcher is lying in the phenomenon that error factors associated with passing of time, then test-retest reliability approach or sometimes multiple administration of parallel tests approach should be applied. On the other hand, if the interest of researcher is lying in the phenomenon that error factors associated with different items, then internal consistency approach or single administration of parallel tests should be applied. Vaske, Beaman, & Sponarski (2013) discussed in their article that reliability analysis for summated scale is commonly performed through Internal consistency of the items.

Moreover, in many situations, the Reliability can't be estimated through Test-Retest or Parallel forms approach. So, the only way remained is through Internal Consistency. In most of the medical sciences research, Cronbach's Alpha is merely computed for the purpose of Internal Consistency. Researchers as well as reviewers both are interested in high values of Cronbach's Alpha. Unfortunately, this reliance has so many problems pointed by some researchers (Sijtsma, 2009; Peters, 2014). The problems with Cronbach’s alpha are easily solved by computing readily available alternatives, such as the Greatest Lower Bound or Omega (Peters, 2014).

The objective of current article is to review the available alternatives to Cronbach's Alpha, their limitations/drawbacks and their proper utilization at best suitable situation. The layout of the article is as follows: In the next section, a brief description of Classical Test Theory is presented, which is followed by the most generally used Coefficient, the Cronbach's Alpha formula and its shortcomings are discussed. Then the other alternative Internal Coefficient measures like Guttman's lambda, Greatest Lower Bound, Raju's

Coefficient etc. are explained. All these coefficients are available in R package "Lambda4". In the last, some suggestions are presented for Medical Sciences researchers for the best use of these coefficients.

Classical Test Theory:

Suppose z be the observed variable and τ the latent true score (Lord & Novick, 1968), and ϵ be the random measurement error. According to CTT, the classical true score model is:

$$z = \tau + \epsilon \quad ` \tag{1}$$

with the assumption $E(\epsilon) = 0$ and $\rho_{\tau\epsilon} = 0$, i.e. average of random measurement error is zero and there is no correlation between errors and true scores. Therefore, the variance of z can be written as:

$$\sigma_z^2 = \sigma_\tau^2 + \sigma_\epsilon^2 \tag{2}$$

Reliability of z under the model (1) can be defined as the ratio of two variances, that is

$$Reliability = \frac{\sigma_\tau^2}{\sigma_z^2} \tag{3}$$

$$= 1 - \frac{\sigma_\epsilon^2}{\sigma_z^2} \tag{4}$$

$$= \frac{1}{1+\frac{\sigma_\epsilon^2}{\sigma_\tau^2}} \tag{5}$$

The above three equivalent equations give three forms of reliability: The eq (3) says that reliability is the ratio of two variances: the true score variance σ_τ^2 and the total variance σ_z^2. The eq (4) expresses reliability in terms of the measurement error variance σ_ϵ^2 and the total variance σ_z^2. The third equation, eq (5) does not contain σ_z^2 explicitly (Tarkkonen & Vehkalahti, 2005).

Since z is the only observable in model (1), so some further assumptions are required. The classical approach of Reliability is based on the principle of *parallel measurements* (Lord & Novick 1968). According to this principle, two measurements, say z_1 and z_2, given by

$z_1 = \tau + \epsilon_1$ and $z_2 = \tau + \epsilon_2$

are said to be parallel if we assume that $\sigma_{\epsilon_1}^2 = \sigma_{\epsilon_2}^2$ and $\rho_{\epsilon_1\epsilon_2} = 0$. From these assumptions it then can be concluded that $\sigma_{z_1}^2 = \sigma_{z_2}^2 = \sigma_z^2\ (say)$. The correlation between z_1 and z_2 is then becomes:

$$\rho_{z_1z_2} = \frac{\sigma_{z_1z_2}}{\sqrt{\sigma_{z_1}^2\sigma_{z_2}^2}} \tag{6}$$

which after omitting the subscripts from z_1 and z_2, becomes:

$$\rho_{zz} = \frac{\sigma_\tau^2}{\sigma_z^2} = \rho_{z\tau}^2 \tag{7}$$

Hence, the correlation between the two parallel measurements is a way to estimate the reliability, but it requires assumption on the measurement errors as well as the true score is also assumed to be strictly one dimensional (Tarkkonen & Vehkalahti, 2005).

It is important to note that the reliability of a test tells the effect of measurement error on the observed score of a group of objects rather than effect on an individual object. To calculate such individual effect, the Standard Error of Measurement (SEM) must be calculated (Tavakol & Dennick, 2011).

Cronbach's alpha:

Cronbach's Alpha (1951) is used when an instrument/test is split into two half-tests (Gliem & Gliem, 2003), with the inflexible assumptions that each of the items exhibit tau-equivalence (items have equivalent factor loadings) and are uni-dimensional (Novick and Lewis, 1967).

Cronbach's Alpha is computed by the following formula:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^n S^2(Z_i)}{S^2(Z)}\right]$$

where $S^2(Z_i)$ is the variance of scores on part *i* over all *k* persons and $S^2(Z)$ is the variance of total scores over all *k* persons on a single administration of a test (Zimmerman & Burkheimer, 1968).

The α will be 1 if all the items are identical and 0 if none is related to another. Also Zumbo (1999) has proved that Cronbach's alpha give us lower bound of reliability. Zimmerman, Zumbo, and Lalonde (1993) found that when additivity is violated, α underestimate the reliability, while, when the assumption regarding uncorrelated errors is violated, α overestimate the reliability.

Although during the derivation of Cronbach's Alpha, normality assumption is not required, however it is required when we estimate coefficient alpha. As it can be seen from above formula that Alpha is a function of variances, therefore, these variance are estimated through least-squares method, which requires normality assumption (Zumbo, 1999). Also, he mentioned that the exact sampling distribution for Cronbach's Alpha has not yet been determined. However, Feldt (1965) and Kristof (1963) derived a transformation of α, which is based on ANOVA approach and uses F as a test statistic. Hence, normality assumption also applies.

Streiner (2003) recommended a maximum alpha value of 0.90. Note that a high level of Alpha (> 0.90) may also suggest redundancies and is an indication for researcher to remove redundant items from the test. George and Mallery (2003) suggest the following rules of thumb:

" $\alpha \geq 0.90$ ………. Excellent,
$0.80 \leq \alpha < 0.90$ ……….. Good,
$0.70 \leq \alpha < 0.80$ ……….. Acceptable,
$0.60 \leq \alpha < 0.70$ ...………Questionable,
$0.50 \leq \alpha < 0.60$ ……….Poor,
and $\alpha < 0.50$ ………….. Unacceptable".

Shortcomings:

1) Most of the researchers (Green 2005; Sajtsma, 2009; Tavakol, 2011; Benton, 2013; Hunt, 2013; Peters, 2014) stated that the assumptions, specifically the assumption of uni-dimensionality are often not met in real situation. Hunt (2013) showed that in case of tau-equivalent model as well as in uni-dimensional case, α performs very well. But it demonstrates a large negative bias in multidimensional case. Similar is the situation in case of parallel models. Actually, Cronbach's Alpha is based on the tau-equivalent model, which assumes that each item in the instrument/test measures the same latent characteristic. But, if this assumption is violated then α will underestimate the reliability.

2) Researchers (Cortina, 1993; Nunnally 1994; Streiner 2003; Tavakol, 2011) have also pointed out that high value of Cronbach's Alpha does not necessarily mean a high degree of internal consistency. Because Alpha is also affected by the size of the test.

3) High values of Cronbach's Alpha may also an indication of items with high inter-item correlations which is a sign of narrow coverage and underrepresentation of the instrument, thus lowering the validity of the instrument (Boyle, 1991; Kline, 1979).

4) Cronbach's alpha is a lower-bound estimate of reliability because heterogeneous instrument/test items would violate the assumptions of the tau-equivalent model (Cortina 1993; Tavakol, 2011).

5) Although Cronbach's Alpha is essentially a ratio of two variances (which are squared quantities and always positive), but one can obtain negative values with Cronbach's alpha (Tarkkonen & Vehkalahti, 2005).

Raju's Coefficient:

Cronbach's alpha (1951) severely underestimates test reliability when the several parts have an unequal number of items. Raju (1977) proposed a generalization of alpha, denoted by β_k, called as Raju's Coefficient.

Let a test *S* having *n* items is split into *k* mutually exclusive subsets, $S_1, S_2, \dots, S_k$ and let S_1 contains n_1 items, S_2 contains n_2 items and so on. Let p_i indicates the proportion of items in S_i i.e. $p_i = \frac{n_i}{n}$, while $X, X_1, X_2, \dots, X_k$ denote the total test scores on $n, n_1, n_2, \dots, n_k$ respectively. Then Raju's Beta Coefficient is given by:

$$\beta_k = \frac{\sum_{i \neq j} \sum \sigma_{x_i x_j}}{\sigma_x^2 \sum_{i \neq j} \sum p_i p_j}$$

where $\sigma_{x_i x_j}$ indicates the covariance between X_i and X_j, while σ_x^2 is the variance of *X*. Alternatively,

$$\beta_k = \frac{\sigma_x^2 - \sum_{i=1}^{k} \sigma_{x_i}^2}{\sigma_x^2 (1 - \sum_{i=1}^{k} p_i^2)}$$

Spearman-Brown formula:

Spearman (1910) also introduced a reliability coefficient. Brown (1910) also derived the same formula independently, and therefore, it is known as Spearman–Brown formula.

The Spearman-Brown's reliability coefficient is also denoted by α and is computed as:

$$\alpha = \frac{n\bar{r}}{1 + (n-1)\bar{r}}$$

where "n" indicates the number of tests combined, and $\bar{r}$ indicates the average inter-item correlation. It predicts the reliability of a new instrument/test, which is composed by replicating the current instrument items "n" times (or, saying in other words, by creating an instrument having "n" parallel forms of the current instrument). Thus n = 3 means triplicating the instrument length by adding items with the same characteristics as those in the current instrument.

Kuder-Richardson's KR-20:

The formula developed by Cronbach (1951), known as Cronbach's Alpha, is actually an extension of Kuder-Richardson's (KR)-20 formula and is used to calculate reliability of tests having dichotomously categorized items. It is given by:

$$KR20 = \frac{n}{n-1}\left[1 - \frac{\sum pq}{\sigma^2}\right]$$

where "n" indicates the number of items in the instrument, "p" is the proportion of individuals having value 1, "q" is the proportion of individuals having value 0 and σ^2 is the overall variance of the instrument.

KR-21 formula:

According to this formula , an estimate of reliability from a single administration of a test is provided by

$$KR21 = \frac{n}{n-1}\left[1 - \frac{\bar{Z}(n-\bar{Z})}{n\sigma^2}\right]$$

where $\bar{Z}$ is the mean of total scores over k persons on a single administration of a test. (Zimmerman & Burkheimer, 1968)

Covariance Maximized Lambda 4 ($\lambda_{4(C)}$):

Hunt (2013) introduced an estimator for split-half reliabilities, which is based on maximization method and utilizes a covariance matrix. The arithmetic mean of the vector of split-half reliabilities can then be used to estimate overall test reliability.

According to Hunt (2013), the procedure for computing ($\lambda_{4(C)}$) involves the following steps:

1. Compute the covariance matrix for all items, and point out the largest covariance between any two items. Place those items (having maximum covariance) on separate splits.
2. Repeat step 1 above for the remaining items (i.e. ignoring those split items), thus creating a vector *t* having (-1, 1) elements.
3. Generate every possible combination of (-1 and 1) elements for *t* that retains the paired-item separations.
4. Compute $\hat{\lambda}_4$ for each t_i using below formula, creating a new vector τ.

$$\lambda_4 = 1 - \left(\frac{\acute{t}\,\Sigma\, t}{\acute{1}\,\Sigma\, 1}\right)$$

5. Take the arithmetic mean or median of τ for $\lambda_{4(C)}$. (Note that if one takes the maximum of τ, then he/she will get Guttman's $\lambda_{4(max)}$.

Hunt (2013) concluded that in case of parallel models, $\lambda_{4(C)}$ shows a low level of bias and has either lower or equivalent consistency as alpha (uni-dimensional case). While, in multidimensional cases, $\lambda_{4(C)}$ has the lowest level of bias and MSE. It performs very well in case of tau-equivalent models and has the lowest level of bias and MSE for multidimensional models. On the other hand, for congenric models, Omega and $\lambda_{4(C)}$ have an equivalent smaller level of bias and MSE.

Guttman's maximal λ_4:

Guttman (1945) introduced his λ_4 as reliability coefficient without any assumption about tau-equivalence model or uni-dimensionality. Its calculation is very simple and involves the split of items into two halves; such that the covariance between scores on the two halves is as high as possible (i.e. it tries to divide the entire test into two halves in such a way that both halves are highly uncorrelated).

The overall test reliability through this method can be computed by the following formula:

$$Reliability = \frac{4\ Cov(first\ half, second\ half)}{var(overall\ test)}$$

It has two drawbacks:

1. In standard statistical packages, like SPSS, its routines are not available. So, researcher has to write his own functions/routines.
2. As Berge and Socan (2004) concluded that there is great chance that it may overestimate the actual reliability in case of small sample or when the test is lengthy. Also, Hunt (2013) revealed that Guttman's maximal λ_4 may overestimate reliability in case of small samples.

McDonald's omega:

McDonald's (1999) introduced a reliability coefficient known as omega (ω_t). Revelle and Zinbarg (2009) preferred McDonald's (1999) ω_t , although they studied that it also has a large positive bias. The formula of McDonald's omega is given as:

$$\omega_h = \frac{\mathbf{1}'\boldsymbol{cc}'\mathbf{1}}{\sigma^2}$$

where $\mathbf{c}$ is a vector of factor loadings, and σ^2 is the overall variance.

Hunt (2013) concluded in his comparative research study that in case of parallel models, it does not have a bias greater than 0.01. While, in case of tau-equivalent model and sample size of 100 or less, it exhibits a bias, low consistency and MSE.

Angoff-Feldt Coefficient:

Angoff (1953) and Feldt (1975) proposed a reliability formula for the situation when a test is composed of two (separately timed) sections. This procedure assumes that the two (timed) sections are parallel in content. Lawrence (1995) pointed out that it will underestimate reliability in the situation when the sections are not strictly congeneric.

The reliability formula of Angoff-Feldt is given by:

$$\text{reliability} = \frac{\text{cov}_{\text{ab}}\text{var}_{\text{T}}}{\text{cov}_{\text{aT}}\text{cov}_{\text{bT}}}$$

where "a" represents first section, "b" represents 2nd section, and variances and covariances are computed as:

$$\text{var}_T = \text{var}_a + \text{var}_b + 2\ \text{cov}_{ab}$$
$$\text{cov}_{aT} = \text{var}_a + \text{cov}_{ab},$$
$$\text{cov}_{bT} = \text{var}_b + \text{cov}_{ab}$$

Note that Angoff-Feldt's procedure does not have a requirement of equal or unequal length of two parts (Sijtsma, 2009). Moreover, in some situations when it is impossible to divide the instrument into two equal parts, then the Angoff-Feldt coefficient, is best alternative to Cronbach's Alpha (Feldt & Charter, 2003).

Kristof Coefficient:

Kristof (1974) developed a formula, which may be used to estimate total test reliability in the situation where test is composed of more than two i.e. three separately timed sections. It will also underestimate reliability when the three sections are not strictly congeneric (Lawrence, 1995).

Kristoff (1974) formula for reliability is given by:

$$\text{reliability} = \frac{[\text{cov}_{\text{ab}}\text{cov}_{\text{ac}} + \text{cov}_{\text{ab}}\text{cov}_{\text{bc}} + \text{cov}_{\text{ac}}\text{cov}_{\text{bc}}]^2}{\text{cov}_{\text{ab}}\text{cov}_{\text{ac}}\text{cov}_{\text{bc}}\text{var}_{\text{T}}}$$

where "a", "b", and "c" represent first, second and third sections respectively, and overall variance is computed by:

$$\text{var}_T = \text{var}_a + \text{var}_b + \text{var}_c + 2\ \text{cov}_{ab} + 2\ \text{cov}_{bc} + 2\ \text{cov}_{ac}$$

This procedure is adopted when three parts are homogeneous in content but may or may not be equal in length. (Sijtsma, 2009)

Conclusion:

Whenever, a researcher wants to compute the reliability of an instrument then he/she must confirm the uni-dimensionality or multidimensionality. Some researchers (Green & Thompson, 2005; Tate, 2003) suggest that a Principal Components Analysis (PCA) should be applied to the set of items. If PCA gives only one factor, then it is an indication of uni-dimensionality. The alternative way, as suggested by Greenberg & Anderson (1988), Shelby (2011), is to conduct a Confirmatory Factor Analysis (CFA).

Peters (2014) suggested that researchers should calculate few diagnostics before using any internal consistency formula. These diagnostics include: (a) Conducting a factor analysis/principal component analysis; (b) inspecting the average values and variation in each item; (c) generating a correlation matrix (if data is ordinal, then polychoric correlation should be computed); (d) examining the histograms of each item's distribution etc.

There are a number of different measures of reliability explained in this paper, each has its own limitations. The most commonly measure is the Cronbach's alpha, but it requires that the each category in the model is essentially tau-equivalent (i.e. uni-dimensional). Through Confirmatory Factor Analysis, if the loadings of each item on the category are equivalent, then it is tau-equivalent. And one should use Cronbach's Alpha. On the other hand, if the model is congeneric (i.e. the factor loadings are not the equal), then Cronbach's Alpha underestimates reliability. In that case, McDonald's omega should be used.

If the instrument is split into two parallel parts, the Spearman-Brown's coefficient may be used to estimate the reliability of instrument. However, if the splitting provides parts which are congeneric (i.e. two parts are unequal in length) then Angoff-Feldt coefficient, is most suitable.

REFERENCES

Angoff, W.H., 1953. Test reliability and effective test length. Psychometrika, 18: 1-14.

Bentler, P.M., 1972. A lower-bound method for the dimension-free measurement of internal consistency. Social Science Research, 1: 343-357.

Benton, T., 2015. An empirical assessment of Guttman's Lambda 4 reliability coefficient. Proceedings of Quantitative Psychology Research, pp: 301-310). Springer International Publishing.

Boyle, G.J., 1991. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? Personality and Individual Differences, 12(3): 291-194.

Brown, W., 1910. Some experimental results in the correlation of mental abilities. British Journal of Psychology, 3: 296-322.

Cortina, J.M., 1993. What Is Coefficient Alpha? An examination of theory and applications. Journal of Applied Psychology, 78(1): 98-104.

Cronbach, L.J., 1951. Coefficient Alpha and the internal structure of tests. Psychometrika, 16(3): 297-334.

Feldt, L.S., 1965. A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. Psychometrika , 34(3): 363-373.

Feldt, L.S., 1975. Estimation of the reliability of a test divided into two parts of unequal length. Psychometrika , 40(4): 557-561.

Feldt, L.S. and R. Charter, 2003. Estimating the reliability of a test split into two parts of equal or unequal length. Psychological Methods, 8(1): 102-109.

George, D. and P. Mallery, 2003. *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gliem, J.A. and R.R. Gliem, 2003. Calculating, interpreting, and reporting Cronbach's Alpha reliability coefficient for Likert-type scales. *Midwest Research to Practice Conference in Adult, Continuing, and Community Education*, (pp. 82-88). The Ohio State University, Columbus.

Green, S.B. and M. Thompson, 2005. *Structural equation modeling in clinical psychology research.* (M. Roberts, & S. Ilardi, Eds.) Oxford: Wiley-Blackwell.

Greenberg, D.W. and J.C. Anderson, 1988. An updated paradigm for scale development incorporating unidimensionality and its assessment. Journal of Marketing Research, 25: 186-192.

Guttman, L., 1945. A basis for analyzing test-retest reliability. Psychometrika, 10: 255-282.

Hunt, T., 2013. Covariance Maximized Lambda 4: An introduction of a low-biased reliability, pp: 1-22.

Hunt, T., 2013. *Lambda4: Collection of Internal Consistency Reliability Coefficients.. R package version 3.0*. Retrieved June 18, 2015, from http://CRAN.R-project.org/package=Lambda4.

Kline, P., 1979. *Psychometrics and psychology.* London, United Kingdom: Academic Press.

Kristof, W., 1974. Estimation of reliability and true score variance from a split of a test into three arbitrary parts. Psychometrika , 39: 491-499.

Kristof, W., 1963. The statistical theory of stepped-up reliability coefficients when a test has been divided into several equivalent parts. Psychometrika, 28(3): 221-238.

Lawrence, I.M., 1995. *Estimating reliability for tests composed of item sets.* Educational Testing- service, Princeton, New Jersey.

Likert, R., 1932. A technique for the measurement of Attitudes. Archives of Psychology, 140: 1-55.

Lord, F.M. and M.R. Novick, 1968. *Statistical Theories of Mental Test Scores.* London: Addison-Wesley.

McDonald, R.P., 1999. *Test theory: A unified treatment.* New York: Psychology Press.

Novick, M.R. and C. Lewis, 1967. Coefficient Alpha and the reliability of composite measurements. Psychometrika, 32(1): 1-13.

Nunnally, J.C. and I.H. Bernstein, 1994. *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill, Inc.

Peters, G.Y., 2014. The Alpha and the Omega of scale reliability and validity. The European Health Psychologist, 16(2): 56-69.

Raju, N.S., 1977. A generalization of Coefficient Alpha. Psychometrika, 42(4): 549-565.

Revelle, W., and R.E. Zinbarg, 2009. Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. Psychometrika, 74(1): 145-154.

Shelby, L.B., 2011. Beyond Cronbach‟s alpha: Considering confirmatory factor analysis and segmentation. Human Dimensions of Wildlife, 16: 142-148.

Sijtsma, K., 2009. Reliability beyond theory and into practice. Psychometrika, 74(1): 169-173.

Spearman, C., 1910. Correlation calculated from faulty data. British Journal of Psychology, 3: 271-295.

Streiner, D.L., 2003. Starting at the beginning: An introduction to coefficient alpha and internal consistency. Journal of Personality Assessment, 80(1): 99-103.

Tarkkonen, L. and K. Vehkalahti, 2005. Measurement errors in multivariate measurement scales. Journal of Multivariate Analysis, 96: 172-189.

Tate, R., 2003. A comparison of selected empirical methods for assessing the structure of responses to test items. Applied Psychological Measurement, 27: 159-203.

Tavakol, M., and R. Dennick, 2011. Making sense of Cronbach’s alpha. International Journal of Medical Education, 2: 53-55.

Vaske, J.J., J. Beaman and C.C. Sponarski, 2013. Rethinking Internal Consistency in Cronbach's Alpha. 19th International Symposium on Society and Resource Management 2013 Conference. Estes Park, Colorado.

Zimmerman, D.W. and G.J. Burkheimer, 1968. Coefficient Alpha, test reliability, and heterogeneity of score distributions. The Journal Of Experimental Education, 37(2): 90-95.

Zimmerman, D., B. Zumbo and C. Lalonde, 1993. Coefficient Alpha as an estimate of test reliability under violation of two assumptions. Educational and Psychological Measurement, (53): 33-49.

Zumbo, B.D., 1999. A Glance at Coefficient Alpha with an eye towards robustness studies: Some mathematical notes and a simulation model (Paper No. ESQBS-99-1). Prince George, B.C.: University of Northern British Columbia. Edgeworth Laboratory for Quantitative Behavioural Science, pp: 277.