Estimate Reliability Measurement for Multidimensional Scales

Wahyu Widhiarso
Gadjah Mada University
wahyu_psy@ugm.ac.id

Abstract

The assumptions of multidimensionality measures seem to be violated as researcher apply unidimensional reliability coefficient (i.e. alpha coefficient) to estimate reliability for multidimensional data. There are several reliability coefficients for multidimensional measure is applicable and useful, but very few researchers take them into account. Researches suggest that applying alpha coefficient in multidimensional cases will produce underestimation. This article introduced five reliability coefficients such as alpha stratified coefficient, Mosier's, Raykov's, McDonald's and Hancock-Mueller's reliability coefficient, that applicable for multidimensional measurement. Each of these three measures and its computation is described and examples provided.

Introduction

Development of measurement instruments in social and behavioral science usually uses unidimensional assumption which is conceptually formulated as measurement that consist single factor ability or personality traits. However, many studies have shown that the assumption is difficult to be fulfilled as factor analysis tends produce new emerging factors that contribute to measure. It is hard to produce single factor if we measure broad of range ability or trait with several items or indicators in similar amount of precision (tau equivalent) (Kamata, Turhan, & Darandari, 2003). Actually, single dimension was potentially reached when each item is similar in content, its size is little, focus in certain content, and have similar precision to measure. But, sometimes researchers want to develop measurement on broader range of ability or trait; therefore we must extend the size of item and variation of item content. Ideally, measures was assigned on the basis of a single dimension of individual differences, but in practice researcher often reflect clusters of related yet not identical constructs (Raykov & Shrout, 2002).

The tendency of data from measurement in social and behavioral science to multidimensional was caused by several things. The nature construct in behavioral science measurement tends to have a multidimensional rather than unidimensional measure (Drolet & Morrison, 2001). Different with natural sciences that have obvious construct with single measure (e.g. measure length with ruler), construct in behavioral sciences consist many faces or indicators with multiple measure. For example, to measure person self-esteem, we should look in several items or indicators that reflected the following construct rather that single items. The presence of several items will be more complex if an item have different amount to measure. Different amount of items to measure will also tend to produce new factor to measure as in factor analysis we can see whether each item have different communalities value to measure a factor (Kamata, et al., 2003).

Besides the nature of construct in behavioral sciences, measurement administered in behavioral science also tends to emerge new dimension. Crocker & Algina (1986) give a example in aptitude test that administered under strict time limits. Aptitude test was multidimensional measure because the scale scores reflect a speed factor that different with factor that the test was intended to measure. Drolet and Morrison (2001) also showed that multidimensionality was influenced by the number of items. Too much items can increase new common variance that lead a new dimension rather than a dimensions that originally intended.

Empirical results have shown that assumption of the unidimensionality of measures often seems to be violated (Brunner & Sub, 2005). In practical level, some researcher use Cronbach's alpha formula to estimate reliability measure even with violation of unidimensionality assumption. Cronbach's alpha is not properly used if there is more than one dimension in construct to measure. Cronbach's alpha measures how well a set of items measures a single latent construct in unidimensional manner. When data have a multidimensional structure, Cronbach's alpha will usually be underestimated. When a set of items consisted of multiple dimensional construct to measure, we should separate a reliability estimation based on the data or dimension emerged form factor analysis. We can also apply reliability formula to estimate reliability that in multidimensional manner.

In estimating the reliability, researchers often used the coefficient alpha which already popular. Many of researchers do not understand that alpha coefficients have certain criteria in order to have accurate estimation, such as parallel or tau-equivalent assumption and unidimensionality data. Therefore, for researchers who want to estimate reliability in multidimensional measures, it is recommended to use the reliability coefficient that can accommodate multidimensional model. The following

Reliability Coefficient for Multidimensional Measure

1. *Alpha Stratified Coefficient*

Alpha Stratified Coefficient was introduced by Cronbach, Schoneman, & McKie (1965) is useful for estimating the reliability of the instrument consists of several sub-test. As the original coefficient alpha, stratified coefficient alpha was measure reliability of measurements that involving several components, faces, or dimension.

$$\alpha_s = 1 - \frac{\sum_{i=1}^{k} \sigma_i^2 (1 - \alpha_i)}{\sigma_x^2} \tag{1}$$

σ_i^2 = variance of *i* component

α_i = reliability of *i* component

σ_x^2 = variance of total score

Here is an example of calculating the stratified coefficient alpha. For example a researcher's measure attitude toward political policy that consist three dimensions, including dimensions. Value of this descriptive statistics is presented in Table 1.

Table 1. Statistik Deskriptif Hasil Pengukuran Konsep Diri

Dimensi	variance	reliability in each dimension
Dimension A	3	0,80
Dimension B	3	0,80
Dimension C	2	0,70
Total Score	6	

Based on the following information, if we use the coefficient alpha to estimate reliability measurements, we will get reliability results in low estimation (underestimated).

$$\alpha = \frac{2}{3} * 1 - \frac{(3+3+2)}{6} = -0{,}50$$

Reliability coefficient will produce results more satisfactory estimation using stratified alpha coefficient was applied.

$$\alpha_s = 1 - \frac{3(1-0.8)+3(1-0{,}8)+2(1-0.7)}{6} = 0{,}70$$

2. Mosier's Reliability of Composite Score

In 1943 Mosier (1943) develop a reliability coefficient that imposes on the structure of multidimensional measurements. This coefficient can be applied in measurements that have independent multidimensional structure that reflected in several components. Mosier note that this coefficient is a general formula for the reliability of a weighted composite that can be estimated from a knowledge of the weights whatever their source, reliabilities, dispersions, and intercorrelations of the components. For example, reliability estimation of aptitude tests that consists several sub-test. Different with alpha stratified coefficient, reliability of composite score can accommodate differences in the weighting of each sub-test score.

$$r_{xx'} = 1 - \frac{(\Sigma w_j^2 s_j^2) - (\Sigma w_j^2 s_j^2 r_{jj'})}{(\Sigma w_j^2 s_j^2) + 2(\Sigma w_j w_k s_j s_k r_{jk})} \qquad (2)$$

w_j^2 = wighted for j dimension
$r_{jj'}$ = reliability for j dimension
r_{jk} = correlation between j and k dimension
s_j^2 = variance of j dimension

To calculate the composite score reliability we need information about the reliability of each dimension, weighting value, variance score of each dimension and correlation between dimension scores. For example, a problem attribution scale consists of two dimensions, namely attribution to cause of problems and troubleshooting. You can find each statistical value in Table 2.

Table 2. Statictical value of each dimension score

dimension	reliability	variance	weighted	correlation between dimension
Problem Caused	0,85	25	2	0,4
Troubleshooting	0,75	36	3	

$$r_{xx'} = \frac{[(2^2.25)+(3^2.36)]-[(2^2.5.0{,}85)+(3^2.6.0{,}75)]}{[(2^2.25)+(3^2.36)]+[2.(2.3.5.6.0{,}4)]} = 0{,}8310$$

The characteristics of this composite reliability coefficient are: (a) reliability estimation will achieve of 1.00 if all reliability of each dimension are also 1.00; (b) the greater correlation value between each dimensions, the greater reliability estimation produce (c) reliability estimation tend to be larger than the average reliability of each dimension, except we find situation that component reliability, variance value and dimension weighted was similar and also correlation between the dimension is near zero. If this is satisfied, analysis result will produce composite reliability as average of reliability each dimension.

3. Raykov's Reliability of Composite Score

Raykov (1997) develop formula to estimate reliability for composite score based on structural equation modeling that applicable in exploring the factorial structure of an item set that called composite reliability for congeneric measurement models. This coefficient is based on fitting a correspondingly constrained structural equation model and covariance structure analysis methods for scale reliability estimation with congeneric tests (Raykov & Shrout, 2002). Approximate standard error and confidence interval for the reliability coefficient in this approach was permitted, whether unidimensional or not.

Composite reliability coefficient is obtained by ratio of construct measure variance and composite score variance. This equation is based on classical test theory score that defined the reliability as ration division between true variance and observed score variance. As approach at the level of latent construct, confirmatory factor analysis take into account.

$$r_{xx'} = \frac{\mathrm{var}\,(\sum_{i=1}^{p} \sum_{j=1}^{k} \lambda_{ij}\eta_i)}{\mathrm{var}\,(\sum_{i=1}^{p} \sum_{j=1}^{k} \lambda_{ij}\eta_i + \sum_{i=1}^{p} E_1)} \qquad (3)$$

λ_{ij} = *unstandardize loading factor* of Y_i indicator in η_i factor

η_i = i *factor*.

E_i = Error measurement of Y_i indicator

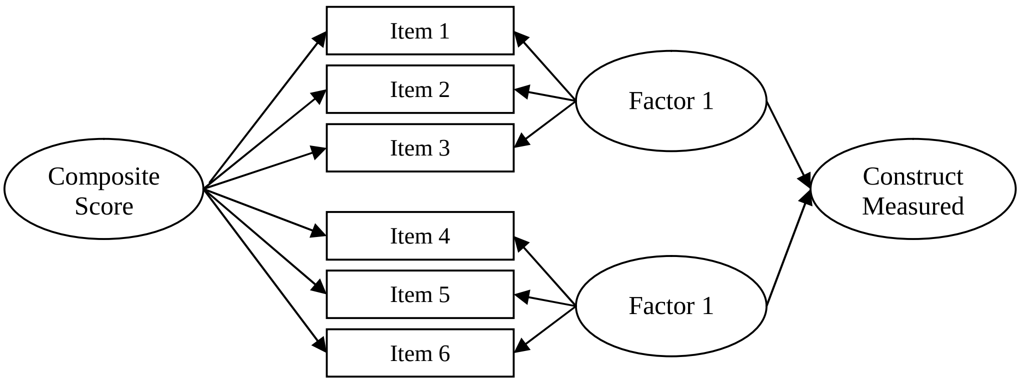

Figure 1. Structural model to estimate reliability of composite score

There has been no computer programs that help compute the this composite reliability. Applying this estimation require to calculate a complex syntax based program structural equation modelling (SEM) with EQS or LISREL. The model that used to estimate composite reliability coefficient can be seen in figure 1. The figure shows a measurement instrument consist six items that reflected in multidimensional model that consist two factors. Both of factors predict the amount of construct variance and each indicator predict the amount of the composite variance. Composite reliability obtained from the division of construct variance and composite variance.

4. McDonald's Construct Reliability

McDonald (1981) developed a reliability coefficient based on the analysis confirmatory factor analysis as part of structural equation modelling, that called construct reliability or omega coefficient (ω). Construct reliability explains how much proportion of indicators can explain a construct to measured.

$$\omega = \frac{\left(\sum_{i=1}^{i} \lambda_i\right)^2}{\left(\sum_{i=1}^{i} \lambda_i\right)^2 + \left(\sum_{i=1}^{i} 1 - \lambda_i^2\right)} \qquad (4)$$

λ_i = unstandardized factor loading of *i*-indiators

This coefficient is widely used by researchers who use structural equation modelling analysis using multidimensional construct (Segars, 1997). Applying this coefficient to parallel or tau equivalent model, which assumes each indicator has an equal amount to measure, will obtain reliability estimation in simliar value to alpha coefficient. On the other hand, if this coefficient apply to congeneric model, which assumes each indicator has a different amount of measure, reliability estimation that obtained will higher than alpha coefficient (Yurdugul, 2006).

5. Hancock and Mueller's Weighted Construct Reliability

Construct reliability coefficient was introduced by Hancock and Mueller (2001) which identify how well the indicator could reflect construct to be measured. This coefficient modify omega coefficient by McDonald which unable to accommodate different weights in each dimensions. A result of this modification is a new formula called weighted construct reliability.

$$\Omega_w = \frac{\sum_{i=1}^{p} \frac{l_i^2}{(1-l_i^2)}}{1 + \sum_{i=1}^{p} \frac{l_i^2}{(1-l_i^2)}} \qquad (5)$$

l_i = standardized coefficient in *i* dimension

To obtain reliability estimation, we should find loading factor square score of each dimension from confirmatory factor analysis. This score can be found if we use structural equation modelling analysis from computing programs such as EQS, LISREL or AMOS. For example, we try to estimate reliability from measurement that contains two dimensions which each dimension consist three items.

Table 5. Statistics of the data

Dimension	Items	l_i^2	$l_i^2/(1-l_i^2)$
A Dimension	1	0.38	0.612903
	2	0.53	1.12766
	3	0.59	1.439024
B Dimension	1	0.47	0.886792
	2	0.39	0.639344
	3	0.66	1.941176
TOTAL (Σ)			6.6469

$$\Omega_w = \frac{6,6469}{1+6,6469} = 0,8692$$

Construct weighted reliability coefficient can be interpreted as the square of the correlation between the dimensions of the optimum linear composite so that some author call it the maximum reliability.

Conclusion

The coefficient of reliability for multidimensional measurements presented in this paper can be distinguished into two types based on the information needed to calculate the reliability: reliability that based on variance and covariance factor that does not require the factor analysis procedures, such as the alpha stratified reliability and Mosier's composite score reliability and reliability with confirmatory factor analysis using structural equation modelling approaches, such as McDonald's construct reliability and Raykov's composite reliability. Reliability of multidimensional measures also can be distinguished: the amount of scale score variance that is accounted for by all underlying factors (composite reliability) and the degree to which the scale score reflects one particular factor (construct reliability) (Brunner & Sub, 2005).

Which coefficient reliability will be used, depends on the approach and assumption that elaborated. Before researcher estimate reliability, I suggest to conducting preliminary analysis to explore dimensionality through factor analysis or similar methods. If results suggest that the data contain multidimensional measures, reliability estimation through multidimensional reliability coefficient should be utilized. Researchers who use the structural equation modelling approach should apply the structural equation modelling based coefficients, while research that uses classical approaches can choose the coefficient of reliability that is not based structural equation modelling.

Brunner, M., & Sub, H.-M. (2005). Analyzing the Reliability of Multidimensional Measures: An Example from Intelligence Research. *Educational and Psychological Measurement, 65*(2), 227-240.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.

Cronbach, L. J., Schoneman, P., & McKie, D. (1965). Alpha coefficient for stratified-parallel tests. *Educational & Psychological Measurement, 25*, 291-312.

Drolet, A. L., & Morrison, D. G. (2001). Do We Really Need Multiple-Item Measures in Service Research? *Journal of Service Research, 3*(3), 196-204.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In S. d. T. Cudeck & D. Sorbom (Eds.), *A festschrift in honor of Karl Jöreskog (pp. 195-216)*. Lincolnwood, IL: Scientific Software International.

Kamata, A., Turhan, A., & Darandari, E. (2003). *Estimating Reliability for Multidimensional Composite Scale Scores*. Paper presented at the Annual meeting of American Educational Research Association.

Mosier, C. (1943). On the reliability of a weighted composite. [10.1007/BF02288700]. *Psychometrika, 8*(3), 161-168.

Raykov, T. (1997). Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement, 21*(2), 173-184.

Raykov, T., & Shrout, P. E. (2002). Reliability of Scales With General Structure: Point and Interval Estimation Using a Structural Equation Modeling Approach. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 195 - 212.