Journal of Engineering
Research and Education
Vol. 7 (2013) 37-44

Effect of sample size multidimensional reliability of National Scale Test of Mathematics in DKI Jakarta Science Program High Schools

Diana Suzana Mandar

Doctoral Student at Psychometrics Study Program, Jakarta State University (UNJ), Jakarta, Indonesia
suzanad23@yahoo.com

ABSTRACT

Estimates of internal consistency reliability of the approach are based on data from one form of the imposition of a measure on a group of subjects (single trial administration). There is a relationship between the two sides approach with consistency in estimating the reliability of the mathematical estimation. Difference between the two approaches is the unit of analysis in which the two sides approach compares the first half with the other parts of the test while the internal consistency approach compares each item with every other item on the test. The advantage of using internal consistency approach is, to avoid the problems that are usually caused by a re-test approach and the approach of parallel tests. The shape and nature of the measuring instrument and the many parts that are made will determine the reliability coefficient estimation technique. Among the computational techniques of internal consistency reliability model of a multidimensional approach in this paper discusses the Stratified Alpha, Mosier, and Wang Reliability Coefficient formula. Interpretation of the reliability coefficient is relative. There is no absolute limit which showed the lowest reliability coefficients that must be achieved so that a measurement can be called reliable. Many researchers found that evaluation of the psychometric property measurements using psychometric formulas without verifying the assumptions underlying the formulas can be questioned its reliability. If these assumptions are not met then the Alpha coefficient reliability produced is below the limit value estimate (underestimate). Psychometric experts have warned that most of the formulas contained psychometric assumptions that need to be met in order for these formulas yield accurate information on the measurements made.

Keywords: *Reliability, Stratified Alpha, Mosier, Wang.*

INTRODUCTION

Understanding the reliability of measuring instruments and the reliability of measurement results can be considered equal. The concept of reliability in terms of reliability measure is closely related to the problem of measurement error. Measurement error itself refers to the extent of the inconsistency of measurement results occur when repeated measurements are made on the same subject. High and low reliability, is empirically demonstrated by a number called, the Wang reliability coefficient. Firstly, the level of high-reliability mirrored by the low correlation, between the distribution of scores from the two parallel measuring devices worn on the same group of individuals.

For researchers who required to identify the reliability of multidimensional measurement, it is recommended to use a reliability coefficient that can accommodate a multidimensional model. To measure the scores of high school students in the science program of the National Examination of Mathematics, consisting of more than one component in the Graduate Competency Standards that measured the ability is multidimensional.

RESEARCH METHODOLOGY

a) Stratified Alpha Reliability Coefficient

Alpha reliability coefficients appropriate measurements that are imposed on unidimension is indicated by factor analysis, items in the scale would produce a factor. Conversely if the result of factors compound the measurements made are multidimensional measurement.

Due to the Cronbach Alpha formula is to measure only one dimension, a formula that measures a combination of several dimensions is developed as below:

The formula is effected by:

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

Variance items in the i-th component
Reliability of i-th component
Variance total test scores

Formula Alpha reliability coefficient of Stratified composites has several characteristics, among others:

Alpha coefficient is precisely Stratified charged in the case of a multidimensional composite scores such as battery test is multidimensional. Variance and reliability of each component obtained from computing a set of points in one dimension while the variance component or total item obtained from the computation of all items.

Stratified Alpha coefficient (Stratified Alpha) was introduced by Cronbach et al. (1965) are useful for estimating the reliability of the instrument that consists of several subtest. Just as the coefficient Alpha, Alpha coefficient is a measure of internal consistency Stratified by involving the components of the test.

b) Wang Reliability Coefficient

The formula is effected by:

$$r_{xx'} = \frac{\sum_{i-1}^{n} w_i^2 r_{ii'} + \sum_{i-1}^{n}\sum_{j-1}^{n} w_i w_j r_{ij}}{\sum_{i-1}^{n} w_i^2 + \sum_{i-1}^{n}\sum_{j-1}^{n} w_i w_j r_{ij}}$$

Weight of the j-th dimension
Reliability dimension j
Correlation between the dimensions of the i-th and j-th

Wang reliability coefficient formula has several characteristics, among others:

1) The lowest reliability can be achieved from the reliability coefficient is the lowest value of the reliability dimension.
2) The reliability of the composite can be higher than the reliability of each dimension once there is a high correlation value between one dimension with another dimension.
3) If the reliability coefficients of each dimension are equal, then the reliability of the composite can reach a maximum value when the weighting of each dimension are equal.

c) Mosier Reliability Coefficient

The formula is effected by:

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

Weight of the j-th dimension
Reliability dimension j
Correlation between the dimensions of the i-th and j-th
Dimension of the j-th variance

Reliability coefficient formula Mosier has several characteristics, among others:

1) Reliability coefficient may be worth 1.00 if all the component reliability is well worth the 1.00.
2) The greater the correlation between the dimensions of the resulting value of the greater reliability.
3) The reliability tended to be larger than the average reliability of each component, except on the condition of the components have reliability, variance and the same weight as well as the correlation between the components is zero. This last condition will produce an average reliability of the reliability of each component.

RESULTS AND DISCUSSION

Theoretically, these problems will become the foundation in thinking framework:

1. Stratified Alpha reliability coefficient is effected by:

   In constantly variance total test scores, then:

   a) Reliability of i-th component:
      If reliability of i-th component multiply with a positive number which is bigger than one, then Stratified Alpha reliability coefficient becomes bigger.

      If reliability of i-th component multiply with a positive number which is smaller than one, then Stratified Alpha reliability coefficient becomes smaller.

   b) Variance items in the i-th component:
      If variance items in the i-th component multiply with a positive number which is bigger than one, then Stratified Alpha reliability coefficient becomes smaller.

      If variance items in the i-th component multiply with a positive number which is smaller than one, then Stratified Alpha reliability coefficient becomes bigger.

2. Wang reliability coefficient is effected by:

   In constantly weight of the j-th dimension, then:

   a) Correlation between the dimensions of the i-th and j-th:

If correlation between the dimensions of the i-th and j-th multiply with a positive number which is bigger than one, then Wang reliability coefficient becomes bigger.

If correlation between the dimensions of the i-th and j-th multiply with a positive number which is smaller than one, then Wang reliability coefficient becomes smaller.

b) Reliability dimension j:
If reliability dimension j multiply with a positive number which is bigger than one, then Wang reliability coefficient becomes smaller.

If reliability dimension j multiply with a positive number which is smaller than one, then Wang reliability coefficient becomes bigger.

3. Mosier reliability coefficient is effected by:

In constantly weight of the j-th dimension, then:

a) Correlation between the dimensions of the j-th and k-th:
If correlation between the dimensions of the j-th and k-th multiply with a positive number which is bigger than one, then Wang reliability coefficient becomes bigger.

If correlation between the dimensions of the j-th and k-th multiply with a positive number which is smaller than one, then Wang reliability coefficient becomes smaller.

b) Reliability dimension j:
If reliability dimension j multiply with a positive number which is bigger than one, then Wang reliability coefficient becomes smaller.

If reliability dimension j multiply with a positive number which is smaller than one, then Wang reliability coefficient becomes bigger.

c) Dimension of the j-th variance:
If dimension of the j-th variance multiply with a positive number which is bigger than one, then Wang reliability coefficient becomes smaller.

If dimension of the j-th variance multiply with a positive number which is smaller than one, then Wang reliability coefficient becomes bigger.

CONCLUSIONS

1. Stratified Alpha reliability coefficient value higher than the value of Mosier reliability coefficient.
2. Stratified Alpha reliability coefficient value higher than the value of the Wang reliability coefficient.
3. Mosier reliability coefficient value higher than the value of Wang reliability coefficient.

REFERENCES

[1] Anastasi, Anne & Susana Urbina. (1997). *Psychological Testing.* Seventh edition. Upper Saddle River, NJ: Prentice-Hall International.

[2] Cortina, J. (1993). What is coefficient Alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.

[3] Crocker, Linda & James Algina. (1986) *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winston.

[4] Ebel, Robert L. (1979). *Essentials of Educational Measurement* (3rd Ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

[5] Gronlund, Norman E. (1985). *Measurement and Evaluation in Teaching.* New York: The Macmillan Publishing Company.

[6] Guilford, J. P. (1954). *Psychometric Methods* (2nd Ed.). Tokyo: Kogakusha Company Ltd.

[7] 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, Chicago, April 2003.

[8] Katsis, Athanassios & Limakopoulou, Aristea. (2005). *The Determination of the Optimal Sample Size for Reliability Scales in Social Sciences.* Paper in Greek, 435-440.

[9] Linn, Robert L. (Ed). (1989). *Educational Measurement* (3rd Ed.). New York: American Council on Education/Mac-Millan Publishing Company.

[10] Lyman, Howard B. (1991). *Test Scores and What They Mean* (5th Ed.). Boston: Allyn and Bacon.

[11] Masters, Geoffrey N. & John P. Keeves (Ed). (1999). *Advances in Measurement in Educational Research and Assessment*. Amsterdam: Pergamon.

[12] McDonald, Roderick P. (1999). *Test Theories: A Unified Treatment.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

[13] Mosier, C. I. (1943). On The Reliability of a Weighted Composite. *Psychometrika, 8,* 161-168. (6,11)

[14] Naga, Dali S. (1992). *Pengantar Teori Sekor pada Pengukuran Pendidikan*. Jakarta: Penerbit Gunadarma.

[15] Nunnally, J. C. (1978). *Psychometric Theory* (2nd Ed.). New York: McGraw-Hill.

[16] Rummel, R. J. (1970). *Applied Factor Analysis*. Evanston, ILL: Northwestern University Press.

[17] Rummel, R. J. (1967). Understanding Factor Analysis. *The Journal of Conflict Resolution, December*, 444-480.

[18] Socan, G. (2000). Assessment of Reliability When Test Items are not Essentially t-Equivalent. In A. Ferligoj & A. Mrvar (Eds.). *Developments in Survey Methodology Editors*. Ljubljana: FDV.

[19] Vehkalahti, K., Puntanen, S., & Tarkkonen, L. (2006). *Estimation of reliability: A Better Alternative for Cronbach's Alpha*. Retrieved 19 November, 2009, from mathstat.helsinki.fi/reports/Preprint430.pdf.

[20] Viswanathan, Madhu. (2005). *Measurement Error and Research Testing*. California: Sage Publications.

[21] Wainer and Thissen. (2001). *Test Scoring.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

[22] Wainer and Braun. (1988). *Test Validity.* Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

[23] Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample Size and Optimal Designs for Reliability Studies. *Statistics in Medicine, 17,* 101-110.

[24] Wang, T. (1998). Weight That Maximize Reliability Under A Congeneric Model. *Applied Psychological Measurement, 22*(2), 179-187.

[25] Widhiarso, W. *Aplikasi Pemodelan Persamaan Struktural dalam Pengujian Model Pengukuran Psikologi*.

[26] Widhiarso, W. *Koefisien Reliabilitas untuk Pengukuran Multidimensi*.

[27] Widhiarso, W. & Mardapi, D. *Komparasi Ketepatan Estimasi Koefisien Reliabilitas Teori Skor Murni Klasik*.

[28] Widhiarso, W. (2011). *Menghitung Koefisien Alpha Berstrata*.

[29] Wiersma, William & Stephen G. Jurs. (1990). *Educational Measurement and Testing* (2nd Ed.). Boston: Allyn and Bacon.

[30] Yurdugul, H. (2006). The Comparison of Reliability Coefficients in Parallel, Tauequivalent, and Congeneric Measurements. *Journal of Faculty of Educational Sciences, 39*(1), 15-37.