Effect of Variability on Cronbach Alpha Reliability in Research Practice

Muhammad Amirrudin[1], Khoirunnisa Nasution[2], Supahar[3]

Abstract

This study aims to describe the effects of variability through data simulation to determine which aspect of variability that maximizes coefficient of Cronbach Alpha reliability. Cronbach Alpha is widely used for estimation of reliability, in recent still. This study served a conceptual and practical simulation for estimating the profound aspect of Cronbach Alpha coefficient relating to the variability of the data. This study carried out with data simulated using the rand between method by Microsoft Excel then simulate different categorical data responses to different range of items by manipulating sample size, range, number of items, variance and standard deviation. The results show that number of variance and standard deviation of data had the most profound aspect of Cronbach Alpha's reliability other than range. The increasing number on some aspect shows that standard deviation and variance has the stability to shows the positive correlation with the coefficient of Cronbach Alpha reliability other than range.

Keywords: Cronbach Alpha, reliability, standard deviation, variability, variance.

1. INTRODUCTION AND PRELIMINARIES

Reliability is pertinent of an instrument to estimate consistently [25]. Mathematically, reliability interpreted as the ratio of the real score of variance to the total score of variance [16] [21]. Reliability is also described as an essential fountain of evidence when establishing the validity of the inferences one predicate on scores from measures and tests [30]. Nunnally [18] stated that the reliability of an instrument does not depend on its validity, but an instrument cannot be valid unless it is reliable. More studies on reliability are needed to make stronger conclusions, as the number of studies addressing this psychometric property was limited [17].

There are multiple definitions and types of reliability measurement (e.g., inter-rater, internal consistency, test-retest), and multiple ways to attain reliability coefficients or ratio (e.g., by using correlation or covariance matrices and via different estimation methods). The importance and complexities of the concept of reliability lead to repeatedly debates. It strays the interpretations and also on purposes of various types of reliability, on the favour and disadvantages of different reliability ratio, and on the methods for attaining them [2][21] [22].

_____

[1,2,3]*Penelitian dan Evaluasi Pendidikan, Universitas Negeri Yogyakarta*

Email : [1]muhammadamirrudin.2019@student.uny.ac.id,
[2]khoirunnisanasution.2019@student.uny.ac.id, [3]supahar@uny.ac.id

Reliability in measurement implies consistency and precision, lack of reliability implies inconsistency and imprecision, both of which equated with measurement error. In the context of testing, measurement error could be defined as any variation in scores that results from factors related to the measurement process that is unrelated to what measured. Reliability, then, is a quality of test scores that suggests they are sufficiently consistent and free from mensuration error to be useful.

It is feasible to estimate the reliability of an instrument and define the meaning of Cronbach's Alpha, the extensively used objective estimation of reliability. Cronbach's alpha is also known as a measure of internal consistency used in the context of multi-item measurement instruments and has wider application than it was early in its development [3] [17] [20][7]. Reason to choose Cronbach Alpha reliability estimation is a topic related with it that has attracted particular attention in the psychometric literature is Cronbach's Alpha [3], which remains the most widely and frequently used reliability indices [22] and used since the 1950s [17] [20]. Cronbach Alpha remains popular lately despite many misunderstanding, misinterpreted, misapply and even justification of not able to properly measure the reliability. [2] [5] [19][28].

Coefficient alpha [3] is undoubtedly one of the most essential and pervasive statistics in research involving test construction and use [2]. Amid the methods to measure internal consistency, Cronbach Alpha reliability has earned more attention than others, such as split-half reliability. Cortina reported that in a review of the Social Sciences Citations Index (SCCI) for the literature from 1966 to 1990, Cronbach's 1951 article had been cited nearly 60 times every year in a total of 278 different journals that wrap a variety of research fields including not only psychology but also education, sociology, statistics, medicine, counselling, nursing, political science, and economics [2].

Cronbach [4] also revealed that the article of "Coefficient Alpha and The Internal Coefficient of Test" in 1951 had been cited roughly 325 times per year and had been mentioned no less than 5,590 times the Social Sciences Citation Index (SCCI) in recent years. In Google Scholar Index until 2019 December it has been cited for 42878 times. In the world of education, Taber [24] surveyed four leading educational journals (International Journal of Science Education, Journal of Research in Science Teaching, Research in Science Education, and Science Education) in 2015, that as many as 69 articles explicitly used Cronbach Alpha in his research.

In practice, the reliability of a test score must estimate from the data of a study. The estimation that is most frequently used to measure reliability in social science and behavioural is coefficient alpha [8] [9] [29]. Alpha requires a sole administration of a test only and its labelled as internal consistency reliability [19]. Alpha was proposed by Kuder and Richardson [14] to estimate dichotomous items. Jackson and Hoyt [12], Guttman [10] and Jackson [13] developed the general version of Alpha's formula as an alternative for the split-half method. Cronbach [3] with his uniqueness proposed formula, famous known as Cronbach Alpha.

Here general formulas of reliability are shown in formula (1)

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_X^2}\right) \qquad (1)$$

Formula 1.1 *General formulas of reliability*

Where α is instrument reliability coefficient, k is the number of items, s_i^2 is the variance of individual item i where $i = 1 \ldots, k$, and s_X^2 is the variance for all items on the scale. This formula is often reported in reduced form as $\alpha = (k^2 \bar{s}_{ij})/s_X^2$ where $\bar{s}_{ij}$ is the mean covariance between all pairs of items on the scale [25]. In other words, the higher the coefficient α, the item has the same covariance and can measure (same consept)[23]. It is the most general known

formula of Cronbach alpha although most researcher have difficulties interpretating it[11]. It also assumes that all items have the same variance of 1.00. However, this assumption is rarely encountered in practice[6].

The size of the reliability coefficient is directly related to the standard deviation of obtained scores for any sample of subjects since the reliability coefficient is a correlation coefficient, as shown in equation (1). Further, Nunnally [18] explained variance; the errors of measurement is approximately independent of the standard deviation of obtained scores. In other words, the standard deviation of measurement considered to be a fixed characteristic of any measure, regardless of the sample of the subject under investigation. It should be clear that the reliability coefficient will be larger in more variable samples.

One of major factor influenced the coefficient of reliability is the variability of data. This variability means it is the extent to which the distribution of characteristics of the group that sampled in calculating the reliability of measuring instruments. If the reliability testing carried out in a homogeneous or similar group, the variance in the group would be very small, when the reliability calculation performed on this measuring instrument – which done by calculating the correlation – the coefficient to be obtained will also be very low [1].

Many studies reported the effect of variability but indirectly spoken as the primary cause. Such as the report of the relationship between test length and reliability studied by Lord & Novick [15]. Cortina added dimensionality, inter-item correlation and number of items that determined coefficient of Alpha by her simulation [2]. However, she noted that Alpha could be high despite low item intercorrelations and multidimensionality; it means the multidimensional test does not has a straightforward or unambiguous interpretation.

Urbina [27] stated that the coefficient alpha is heavily dependent on the amount of inter-item variability within a test. It stands to reason that any lack of uniformity in the content of test items, such as content heterogeneity, will lower these coefficients.

The second undesirable notion is that the length of the test does not affect the value of Alpha. Alpha will increase by alteration the number of items if the average item covariance and variance are kept constant. Since test scores usually have higher reliability in more parts, this attribute makes sense. It may be that this property is usually not taken into consideration when the report and interpretation of administrating Alpha.

It is also insisted and often not realized that tests that subsist of a substantial number of items have a relatively large alpha simply because of the number of items (items alteration) [2][25]. Another result which proving variability has a positive impact is the study from Lozano. He altered with the number of option in a Likert scale which directly affected the coefficient of Alpha. When he reduced the number of choices, then he was reducing the variability of the scale too, so the reliability decrease.

2. METHODS

All of the theoretical above indicated that variability took the central role on the coefficient of Cronbach Alpha reliability. In this paper, we provide a simulation for estimating Cronbach Alpha reliability through changing different coefficients of variability to prove which variability aspect would have a profound effect on Cronbach Alpha reliability. Cronbach Alpha reliability chosen- rather than KR-20 or others- based on the theoretical reason above.

Thompson [26] stated that simulation is a quantitative method with the primary key is randomness, and the approach uses an algorithm. The research was administrated by data

simulation with 20 different simulation groups. This study simulates different categorical data responses to different range of items by manipulating sample size, range, number of items, variance and standard deviation. Total data simulation administrated was 20 groups.

The primary purpose of this study is to (1) provide other Cronbach alpha reliability users, afterwards, can estimate well data's variability to maintain Cronbach alpha's reliability, (2) present an empirical example of effect variability on Cronbach Alpha's coefficient, (3) and understand which aspect of variability has a most profound effect on Cronbach alpha's coefficient. Commonly known variability of measures of statistical dispersion is the variance, standard deviation, and range.

3. MAIN RESULTS

In the first trial we set 200 persons to fill 25 items we maintain range as a dependent variable. The result described in Table 1.

Table 1. Estimating Alpha with different number of range

Series Number	N of Respondents	N of Items	Range	Variance	Standard Deviation	α	Mean
1	200	25	27.00	27.05	5.20	0.10	69.54
2	200	25	30.00	24.36	4.93	0.01	74.29
3	200	25	35.00	42.18	6.49	0.47	80.13
4	200	25	40.00	61.79	7.86	0.67	81.10
5	200	25	45.00	65.50	8.09	0.67	80.42
6	200	25	51.00	68.48	8.27	0.68	79.83
7	200	25	60.00	79.55	8.91	0.74	80.83
8	200	25	72.00	76.62	8.75	0.72	80.93
9	200	25	89.00	103.97	10.19	0.80	80.57
10	200	25	100.00	108.52	10.41	0.81	80.63

From Table 1, we can see that the higher the value of the Range makes the value of reliability also rises. However, anomalies occur in the data range 27.00 and 30.00, even though both Range is 30.00. However, its value of the reliability is not as high as the range 27.00.

Anomalies also found in the data range 60.00 and 72.00. The reliability value of data with a range of 72.00 is not higher than data with a range of 60.00. Does this indicate that there is no positive correlation between the magnitude of range variability and reliability? Wait, we still have another element of variability, that is variance and standard deviation.

We can check that the first anomaly data, the range 30.00 has a variance value and the standard deviation is lower than the data that has a range of 27.00. Then what about the second anomaly, it turns out the same thing found in the second anomaly data. In the data which has a range of 60.00, it turns out to have a higher variance and standard deviation than data with a range of 72.00. Is it true that variances and standard deviations can better represent a positive correlation with the magnitude of reliability? Check at the second data trial as a result shows in Table 2.

In the second data trial, there were still 200 respondents with 25 points but the range value was equal to 60.00.

Table 2. Estimating Alpha with same number of range

Number series	N of Respondents	N of Items	Range	Variance	Standard Deviation	α	Mean
1	200	25	60.00	79.55	8.91	0.74	80.83
2	200	25	60.00	83.72	9.36	0.75	80.72
3	200	25	60.00	87.74	9.36	0.76	80.82
4	200	25	60.00	89.16	9.44	0.76	80.86
5	200	25	60.00	104.21	10.20	0.80	80.67
6	200	25	60.00	121.30	11.01	0.83	80.09
7	200	25	60.00	130.26	11.41	0.84	79.88
8	200	25	60.00	136.21	11.67	0.84	79.01
9	200	25	60.00	144.37	12.01	0.85	79.32
10	200	25	60.00	147.72	12.15	0.85	79.36

In the Table 2 above there is no anomalies of data, but what can we derived from it?

Experiment 1 shows that a higher number of the Range also has a positive correlation with the coefficient of Cronbach Alpha. Similar to what was stated by Nunnally [18] that the size of the data variability affects the alpha reliability coefficient. The greater the variability, the greater the coefficient of the alphabet. However, it turns out when there is an anomaly that there is a greater range that has a smaller Cronbach alpha reliability value than data with a smaller range. Is Nunnaly's theory incorrect?

We retested with the same number of respondents and the same range to see whether the other variability also did not correlate positively with the reliability coefficient. Respondents are still 200 with the set of range at 60.00.

The results turned out to be under the opinion of Nunnally [18] that variability has a significant effect on the magnitude of reliability. Here prove that the variance and standard deviation are better to show a positive correlation with the coefficient of Alpha consistently. It shows in Figure 1.
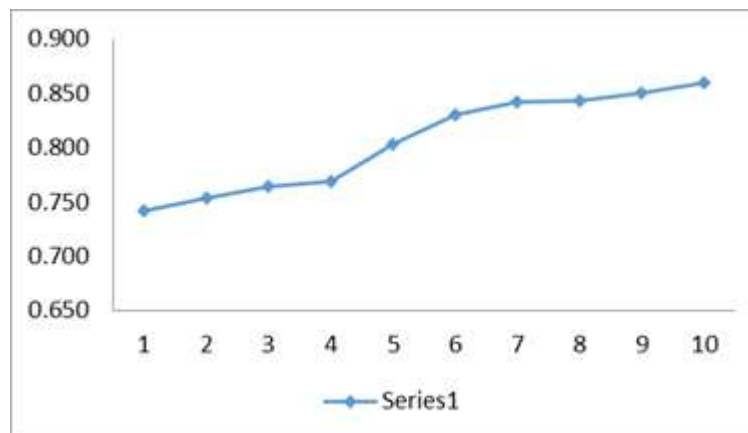
Figure 1. Experiment 2 illustration result

The number of variances and standard deviation sorted from the smallest according to the serial number/data. It will make it easier for us to see the correlation between the amount of variance and the standard deviation of the reliability value in the illustration below.

The greater the value of variance and standard deviation, the greater the value of reliability will make. This second trial can ensure that the correlation between the amount of variance and standard deviation can more positively represent the magnitude of the reliability.

Another factor that affects on coefficient alpha, test length [14], and the number of items[2] [25] basically will lead to data variability. For example, the more number of items, the greater the possibility of data variability could. So the most fundamental principle that affects the coefficient of internal consistency is the variability of the data. However, the most stable to represent the variability of data is the number of variances and standard deviation.

4. CONCLUSION

In real practice, testing measures are hardly consistent. Theories of standard test reliability have refined to estimate the effects of inconsistency on the accuracy of the mensuration. The influence of consistency and inconsistency factors is the fundamental starting point in the test result that reflected in any theories of test reliability.

Based on the formulation, reliability and variability simulation have a close relationship with the value of reliability. The greater the value of variability makes the coefficient of Cronbach Alpha reliability also rises. In the results of experiments or tests found that the correlation of variance and standard deviation can more positively represent the coefficient of Cronbach Alpha reliability than Range. This pattern-finding will help the researcher to formulate their research instrument, because the higher the coefficient, the higher the Cronbach's Alpha value, the more reliable the instrument will be used to measure the construct.

This research will open and lead to more extensive studying about the popularity of Cronbach Alpha reliability in research practice. Especially questions about how does Cronbach alpha's variability pattern-finding will lead to an undesirable belief of Cronbach alpha.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest

REFERENCES

[1] Anastasi, A., & Urbina, S. (1997). Psychological Testing (7th ed.). Upper Saddle River, NJ: Prentice Hall.

[2] Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78,* doi: 10.1037/0021-9010.78.1.98

[3] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests.

Psychometrika, 16, doi:10.1007/BF02310555

[4]   Cronbach, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. Educational and Psychological Measurement, 64,

[5]   Crutzen, R., & Peters, G. J. Y. (2017). Scale quality: alpha is an inadequate estimate and factor-analytic evidence is needed first of all. *Health Psychology Review*, *11*(3), 242–247. https://doi.org/10.1080/17437199.2015.1124240

[6]   Davenport, E. C., Davison, M. L., Liou, P. Y., & Love, Q. U. (2015). Reliability, Dimensionality, and Internal Consistency as Defined by Cronbach: Distinct Albeit Related Concepts. *Educational Measurement: Issues and Practice*, *34*(4), 4–9. https://doi.org/10.1111/emip.12095

[7]   de Vet, H. C. W., Mokkink, L. B., Mosmuller, D. G., & Terwee, C. B. (2017). Spearman–Brown prophecy formula and Cronbach's alpha: different faces of reliability and opportunities for new applications. *Journal of Clinical Epidemiology*, *85*, 45–49. https://doi.org/10.1016/j.jclinepi.2017.01.013

[8]   Field, A. (2009). Discovering statistics using SPSS. (3nd ed.). Los Angeles: Sage.

[9]   Furr, R. M., & Bacharach, V. R. (2008). Psychometrics. An introduction. Los Angeles: Sage.

[10]   Guttman, L. (1945). A basis for analyzing test-retest reliability. Psychometrika, 10

[11]   Hoekstra, R., Vugteveen, J., Warrens, M. J., & Kruyen, P. M. (2019). An empirical analysis of alleged misunderstandings of coefficient alpha. *International Journal of Social Research Methodology*, *22*(4), 351–364. https://doi.org/10.1080/13645579.2018.1547523

[12]   Hoyt, C. (1941). Test reliability estimated by analysis of variance. Psychometrika, 6,

[13]   Jackson, R. W. B., & Ferguson, G. A. (1941). Studies on the reliability of tests. Department of Educational Research, University of Toronto.

[14]   Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. Psychometrika, 2(3),

[15]   Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental lest scores.* Reading, MA: Addison-Wesley.

[16]   McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale NJ: Erlbaum.

[17]   Mcneish, D., & Mcneish, D. (2018). Psychological Methods Thanks Coefficient Alpha , We ' ll Take It From Here. *Psychological Methods*, *23*(3), 412–433.

[18] Nunnally J, Bernstein L.(1994). *Psychometric theory*. New York: McGraw-Hill Higher, INC;

[19] Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. Psychological Methods, 5,

[20] Poitras, I., Dupuis, F., Bielmann, M., Campeau-Lecours, A., Mercier, C., Bouyer, L. J., & Roy, J. S. (2017). Validity and reliability of Wearable Sensors For Joint Angle Estimation: A Systematic Review. *Psychological Methods*, *19*(7), 1–17. https://doi.org/10.3390/s19071555

[21] Revelle, W., & Zinbarg, R. E. (2009). *Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. Psychometrika,74,* doi: 10.1007/s11336-008-9102-z

[22] Sijtsma, K. (2009). *On the use, the misuse, and the very limited usefulness of Cronbach's Alpha. Psychometrika, 74,* doi: 10.1007/s11336-008-9101-0

[23] Singh, A. S. (2017). *Common Procedures for Development, Validity and Reliability of a Questionnaire. Swaziland: Department of AEM, Faculty of Agriculture, University of Swaziland. V*(5), 2017.

[24] Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

[25] Tavakol M, Mohagheghi MA, Dennick R.(2008). *Assessing the skills of surgical residents using simulation*. J Surg Educ.;65(2)

[26] Thompson, J.K. (2000). *Body Image, Eating Disorder, and Obesity an Integrative Guide for Asesment and Treatment.* Washington : American Psychological Association.

[27] Urbina, Susana. (2004). *Essentials of Psychological Testing*. New Jersey:John Wiley Sons.

[28] Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How Reliable are Measurement Scales? External Factors with Indirect Influence on Reliability Estimators. *Procedia Economics and Finance*, *20*(15), 679–686. https://doi.org/10.1016/s2212-5671(15)00123-9

[29] Warens, M. (2015). Some Relationships Between Cronbach's Alpha and The Spearman-Brown Formula. *Psychometrika*, *32*, 127–137.

[30] Zumbo, AM Gaderman. (2007). *Ordinal Varsions Of Coefficients Alpha And Theta For Likert Rating Scales.* Journal of Modern Applied Statistical Methods 6 (1), 4