

## Research Article

# On Cronbach's Alpha as the Mean of All Possible $k$ -Split Alphas

**Matthijs J. Warrens**

*Institute of Psychology, Unit of Methodology and Statistics, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands*

Correspondence should be addressed to Matthijs J. Warrens; [warrens@fsw.leidenuniv.nl](mailto:warrens@fsw.leidenuniv.nl)

Received 13 July 2014; Accepted 17 September 2014; Published 30 September 2014

Academic Editor: Steve Su

Copyright © 2014 Matthijs J. Warrens. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Coefficient alpha is the most commonly used internal consistency reliability coefficient. Alpha is the mean of all possible  $k$ -split alphas if the items are divided into  $k$  parts of equal size. This result gives proper interpretations of alpha: interpretations that also hold if (some of) its assumptions are not valid. Here we consider the cases where the items cannot be split into parts of equal size. It is shown that if a  $k$ -split is made such that the items are divided as evenly as possible, the difference between alpha and the mean of all possible  $k$ -split alphas can be made arbitrarily small by increasing the number of items.

## 1. Introduction

In test theory, test scores are used to summarize the performance of participants on a test. An important concept of a test score is its reliability, which concerns the precision of the administration of a participant's score. Reliability can be conceptualized in different ways. In layman's terms, a test score is said to be reliable if it produces similar outcomes for participants under similar administration conditions. A more formal definition of reliability comes from classical test theory; namely, reliability is defined as the ratio of the true score variance and the total score variance [1–3]. The true score variance cannot be observed directly. Therefore, it has to be estimated from the data. If there is only one test administration reliability can be estimated using the so-called internal consistency reliability coefficients [4]. The internal consistency coefficient that is most commonly used in psychology and other behavioral sciences is coefficient alpha [5–10].

Coefficient alpha was already discussed in Guttman [11]. The coefficient was later popularized by Cronbach [6]. Since then it has been applied in thousands of research studies [5, 9]. It has been pointed out that alpha is not a measure of the one-dimensionality of a test score [6, 9, 12]. Furthermore, a problem is that the value of alpha is affected by the length of the test. Alpha is often high if the test consists of 15 items or more [5, 13, 14]. Moreover, the derivation of alpha is based on several assumptions from classical test theory. For example,

alpha is an estimate of the reliability if the items of the test are essentially tau-equivalent [1, 2]. These assumptions may not hold in practice. Various authors have therefore studied alpha's robustness to violations of its assumptions [12, 15–17].

Other authors have presented results that give a meaning to alpha and to its estimate from a sample in case its assumptions do not hold [6, 18]. The items of a test can be split into two parts. This will be called a 2-split. For each part, the sum score of the items can be determined. If we apply Cronbach's alpha to the two sum scores, we obtain the Flanagan-Rulon split-half estimate of the reliability [19, 20]. The split-half reliability is an alternative approach of estimating the reliability of a test when there is only one administration [3, 7, 8]. The problem with a 2-split of a test is that there are multiple ways to divide the items into two parts. The value of the 2-split alpha therefore depends on the way the 2-splits are made [7]. Cronbach [6] showed that if the 2-splits are made such that the parts have equal size, then the overall alpha is the mean of all possible 2-split alphas. Cronbach's result was generalized by Raju [18] to other types of splits. Let  $k \geq 2$  be an integer. Raju [18] showed that if the  $k$ -split is such that the  $k$  parts have the same number of items, then alpha is the mean of all possible  $k$ -split alphas.

The algebraic results by Cronbach [6] and Raju [18] are important since they provide a proper interpretation of alpha and of its estimate from a sample in the case that its assumptions do not hold [2, 5]. Thus, if the number of items is a composite number  $n$ , then there exists an interpretation of

alpha for any proper divisor of  $n$ . Raju [18] also showed that if the  $k$  parts do not have equal size, then the mean of all  $k$ -split alphas is strictly smaller than the overall alpha. Moreover, this difference can be substantial [18]. Thus, at present, it is unknown if there is an (approximate) interpretation of alpha as a mean of  $k$ -split alphas in case the number of items is a prime number. This is the topic of investigation of this paper.

The paper is organized as follows. Notation and Definitions are presented in Section 2. We also present a formula of the mean of all possible  $k$ -split alphas in Section 2. In Section 3, the nonnegative difference between alpha and the mean of all  $k$ -split alphas is studied. It is shown that if we divide the items as evenly as possible, the difference is strictly decreasing in the number of items. Furthermore, the difference goes to zero if the number of items increases. Hence, the difference can be made arbitrarily small by increasing the number of items. In Section 4, we determine for three criteria and for small values of  $k$  how many items are needed. Finally, Section 5 contains a conclusion.

## 2. Notation and Definitions

Suppose we have a test that consists of  $n \geq 2$  items. Let  $\sigma_{ii'}$  for  $1 \leq i, i' \leq n$  denote the covariance between items  $i$  and  $i'$ , and let  $\sigma_X^2$  denote the variance of the total score. Guttman's lambda3 or Cronbach's alpha is defined as

$$\alpha := \frac{n}{n-1} \cdot \frac{\sum_{i \neq i'} \sigma_{ii'}}{\sigma_X^2}. \quad (1)$$

In the summation  $\sum_{i \neq i'} \sigma_{ii'}$  both counters  $i$  and  $i'$  run from 1 to  $n$ , but they are never equal.

Next, suppose we split the test into  $k$  parts of sizes  $n_1, n_2, \dots, n_k$  with  $1 \leq n_j < n$  for  $1 \leq j \leq k$  and

$$\sum_{j=1}^k n_j = n. \quad (2)$$

Let  $p_j = n_j/n$  for  $1 \leq j \leq k$  denote the proportion of items in part  $j$ . Furthermore, let  $\sigma_{jj'}$  denote the covariance between the sum scores of parts  $j$  and  $j'$ . For a  $k$ -split coefficient alpha is defined as [18]

$$\alpha_k := \frac{k}{k-1} \cdot \frac{\sum_{j \neq j'} \sigma_{jj'}}{\sigma_X^2}. \quad (3)$$

We have  $\alpha_n = \alpha$ . Furthermore, for fixed  $k$  and  $n_1, n_2, \dots, n_k$ , the number of different  $k$ -splits is given by [21, page 823]

$$S = \frac{n!}{n_1! n_2! \cdots n_k!}. \quad (4)$$

For each  $k$ -split, there is an associated  $\alpha_k$ . The mean of all  $k$ -split alphas is denoted by  $E(\alpha_k)$ . An expression for  $E(\alpha_k)$  is presented in Lemma 1. The result is not derived in Raju [18] and is believed to be new.

**Lemma 1.** *It holds that*

$$E(\alpha_k) = \frac{n}{n-1} \cdot \frac{k}{k-1} \cdot \frac{(\sum_{j \neq j'} p_j p_{j'}) (\sum_{i \neq i'} \sigma_{ii'})}{\sigma_X^2}. \quad (5)$$

*Proof.* The variance of the total score  $\sigma_X^2$  is not affected by the  $k$ -split. Furthermore,  $k/(k-1)$  is fixed since  $k$  is fixed. Hence,

$$E(\alpha_k) = \frac{k}{k-1} \cdot \frac{E(\sum_{j \neq j'} \sigma_{jj'})}{\sigma_X^2}, \quad (6)$$

where, using (4),

$$E\left(\sum_{j \neq j'} \sigma_{jj'}\right) = \frac{1}{S} \sum_{\ell=1}^S \left(\sum_{j \neq j'} \sigma_{jj'}\right). \quad (7)$$

We want to determine how often each covariance  $\sigma_{ii'}$  between items  $i$  and  $i'$  occurs in the summation on the right-hand side of (7). Raju [18, page 559] showed that this amount is given by

$$2S \cdot \frac{n}{n-1} \cdot \sum_{j \neq j'} p_j p_{j'}. \quad (8)$$

Since each covariance  $\sigma_{ii'}$  occurs twice in the summation  $\sum_{i \neq i'} \sigma_{ii'}$ , (7) is given by

$$E\left(\sum_{j \neq j'} \sigma_{jj'}\right) = \frac{n}{n-1} \left(\sum_{j \neq j'} p_j p_{j'}\right) \left(\sum_{i \neq i'} \sigma_{ii'}\right). \quad (9)$$

Using (9) in (6), we obtain (5).  $\square$

## 3. Difference between $\alpha$ and $E(\alpha_k)$

In this section, we study the difference between alpha and the mean of all  $k$ -split alphas. Using (1), we can write (5) as

$$E(\alpha_k) = \alpha \cdot \frac{k}{k-1} \sum_{j \neq j'} p_j p_{j'}. \quad (10)$$

Cronbach [6] showed that  $\alpha = E(\alpha_2)$  if  $p_1 = p_2$ . A proof can also be found in Lord and Novick [1]. Raju [18, page 560] showed that  $\alpha = E(\alpha_k)$  if  $p_1 = p_2 = \dots = p_k$  and that  $\alpha > E(\alpha_k)$  if  $p_1, p_2, \dots, p_k$  are not all equal. Hence, the difference

$$\alpha - E(\alpha_k) = \alpha \left(1 - \frac{k}{k-1} \sum_{j \neq j'} p_j p_{j'}\right) \quad (11)$$

is nonnegative. Difference (11) can be used to study how close the values of alpha and the mean of all  $k$ -split alphas are. Since  $0 \leq \alpha \leq 1$ , we have the inequality

$$\alpha - E(\alpha_k) \leq 1 - \frac{k}{k-1} \sum_{j \neq j'} p_j p_{j'}. \quad (12)$$

The right-hand side of (12) is an upper bound of  $\alpha - E(\alpha_k)$  for all values of  $\alpha$  ( $0 \leq \alpha \leq 1$ ). It only depends on  $k$  and the proportions of items in the parts.

It turns out that difference (11) can be quite large [18]. This is the case when the variance between  $p_1, p_2, \dots, p_k$  is large, that is, if  $p_1, p_2, \dots, p_k$  are very different. To illustrate this property of difference (11), suppose we have  $m+k$  items,

where  $m \geq 3$  is an integer. Furthermore, suppose we distribute the items such that we have  $k - 1$  parts with 1 item and 1 part with  $m + 1$  items. The proportions of items per part are given by

$$p_1 = \frac{m + 1}{m + k}, \quad p_2 = \dots = p_k = \frac{1}{m + k}. \quad (13)$$

Using the proportions in (13) we have

$$\sum_{i \neq j} p_i p_j = \frac{(k - 1)(k - 2)}{(m + k)^2} + \frac{2(k - 1)(m + 1)}{(m + k)^2}. \quad (14)$$

Due to the  $m^2$ -term in the denominator of the right-hand side of (14) we have, for fixed  $k < m$ , that  $\sum_{i \neq j} p_i p_j \rightarrow 0$  as  $m \rightarrow \infty$ .

In this paper, we are not interested in an arbitrary  $k$ -split of the items but in a  $k$ -split in which the items are distributed as evenly as possible over the parts. Suppose we have  $km + c$  items where  $m, c \geq 1$  are integers with  $0 < c < k$ . If we want to distribute the items as evenly as possible over the  $k$  parts, an optimal  $k$ -split is to have  $k - c$  parts with  $m$  items and  $c$  parts with  $m + 1$  items. In this case the proportions of items per part are given by

$$p_1 = \dots = p_c = \frac{m + 1}{km + c}, \quad p_{c+1} = \dots = p_k = \frac{m}{km + c}. \quad (15)$$

Lemma 2 shows that for a split of the type in (15) the quantity  $\sum_{j \neq j'} p_j p_{j'}$  approaches the ratio  $(k - 1)/k$  as the counter  $m$  goes to infinity.

**Lemma 2.** *It holds that*

$$\sum_{i \neq j} p_i p_j \rightarrow \frac{k - 1}{k} \quad \text{as } m \rightarrow \infty. \quad (16)$$

*Proof.* Using the proportions in (15), we have the identity

$$\begin{aligned} \sum_{i \neq j} p_i p_j &= \frac{c(c - 1)(m + 1)^2}{(km + c)^2} + \frac{2c(k - c)m(m + 1)}{(km + c)^2} \\ &\quad + \frac{(k - c - 1)(k - c)m^2}{(km + c)^2}. \end{aligned} \quad (17)$$

Using the identity

$$c(c - 1) + 2c(k - c) + (k - c - 1)(k - c) = k(k - 1), \quad (18)$$

we can write (17) as

$$\begin{aligned} \sum_{i \neq j} p_i p_j &= \frac{c(c - 1)(2m + 1) + 2c(k - c)m + k(k - 1)m^2}{(km + c)^2} \\ &= \frac{c(c - 1) + 2c(k - 1)m + k(k - 1)m^2}{(km + c)^2} \\ &= \frac{c(c - 1) + (k - 1)(2cm + km^2)}{(km + c)^2}. \end{aligned} \quad (19)$$

For  $m \rightarrow \infty$ , we have

$$\sum_{i \neq j} p_i p_j \rightarrow \frac{k(k - 1)}{k^2} = \frac{k - 1}{k}. \quad (20)$$

□

It follows from Lemma 2 that the right-hand-side of (12), and thus the difference  $\alpha - E(\alpha_k)$ , goes to zero as  $m$  increases. In other words, if the  $k$ -split is of the type in (15), the difference between alpha and the mean of the best  $k$ -splits vanishes as the number of items becomes large. It is not clear from Lemma 2 that  $\sum_{j \neq j'} p_j p_{j'}$  is strictly increasing in  $m$ . This is shown in Lemma 3.

**Lemma 3.** *For split (15)  $\sum_{j \neq j'} p_j p_{j'}$  is strictly increasing in  $m$ .*

*Proof.* Using (19), the first order partial derivative is

$$\begin{aligned} \frac{\partial \sum_{i \neq j} p_i p_j}{\partial m} &= \frac{2(k - 1)(km + c)^2 - [c(c - 1) + (k - 1)(2cm + km^2)]2k}{(km + c)^3} \\ &= \frac{2c(k - c)}{(km + c)^3}. \end{aligned} \quad (21)$$

Since  $k - c > 0$ , (21) is strictly positive, and it follows that  $\sum_{j \neq j'} p_j p_{j'}$  is strictly increasing in  $m$ . □

## 4. Numerical Results

Lemmas 2 and 3 from the previous section show that the nonnegative difference  $\alpha - E(\alpha_k)$  decreases strictly to zero as  $m$  increases, if the  $k$ -split is of the type in (15). Hence, for sufficiently large  $m$ , the difference  $\alpha - E(\alpha_k)$  can be made smaller than the specified criteria. In this section, we determine the numbers of items that are needed so that the difference  $\alpha - E(\alpha_k)$  is less than 0.010, 0.005, and 0.001. Since a difference of 0.010 is negligible for most practical purposes, we may say that  $\alpha$  is approximately equal to  $E(\alpha_k)$  if the difference is less than 0.010, that is, if

$$\alpha - E(\alpha_k) \leq 0.010. \quad (22)$$

The values 0.005 and 0.001 can be used if one wants a stronger definition of approximately equal.

If we are interested in a  $k$ -split but the number of items is not a multiple of  $k$ , there are  $k - 1$  different scenarios to consider. For example, if  $k = 3$  but the number of items is not a multiple of 3, we either have  $3m + 1$  items or have  $3m + 2$  items. Inequality (22) may hold for different  $m$  depending on whether we have  $3m + 1$  items or  $3m + 2$  items. We therefore consider the number of items of all  $k - 1$  variants of a  $k$ -split separately. For all possibilities of  $k$  and  $c$  with  $2 \leq k \leq 7$  and  $1 \leq c < k$ , we applied the following numerical strategy. For fixed  $k$  and  $c$ , we increased  $m$  in steps of 1, starting at 1, until inequality (22) was satisfied. The results for  $k = 2, 3, \dots, 7$  are

TABLE 1: The minimum number of items per split variant for three criteria.

Split type	Split variant	Minimum number of items		
		$\leq 0.010$	$\leq 0.005$	$\leq 0.001$
2-split	$2m + 1$	11	15	33
3-split	$3m + 1$	10	16	34
	$3m + 2$	11	17	32
4-split	$4m + 1$	13	17	33
	$4m + 2$	14	18	38
	$4m + 3$	11	15	35
5-split	$5m + 1$	11	16	36
	$5m + 2$	17	22	42
	$5m + 3$	13	18	43
	$5m + 4$	14	19	34
6-split	$6m + 1$	13	19	37
	$6m + 2$	14	20	44
	$6m + 3$	15	21	45
	$6m + 4$	16	22	46
	$6m + 5$	11	17	35
7-split	$7m + 1$	15	15	36
	$7m + 2$	16	23	44
	$7m + 3$	17	24	45
	$7m + 4$	18	25	46
	$7m + 5$	19	19	47
	$7m + 6$	13	20	34

presented in Table 1. The table presents for each variant of a  $k$ -split the minimum number of items for which the difference  $\alpha - E(\alpha_k)$  is  $\leq 0.010$  (column 3),  $\leq 0.005$  (column 4), and  $\leq 0.001$  (column 5).

Table 2 summarizes for each  $k$ -split the minimum number of items required per criteria. For example, consider the case  $k = 2$ . If the number of items is even we have a perfect 2-split and  $\alpha = E(\alpha_2)$ . If the number of items is odd and we distribute the items as evenly as possible, Table 1 shows that the difference  $\alpha - E(\alpha_2)$  is no more than 0.010 if we have at least 11 items, no more than 0.005 if we have at least 15 items, and no more than 0.001 if we have at least 33 items. Hence, we may conclude that in the case that we distribute the items as evenly as possible the difference  $\alpha - E(\alpha_2)$  is no more than 0.010 if we have at least 10 items, no more than 0.005 if we have at least 14 items, and no more than 0.001 if we have at least 32 items. These numbers are presented in Table 2.

As another example, consider the case  $k = 3$ . If we have  $3m$  items, we have a perfect 3-split and  $\alpha = E(\alpha_3)$ . If we have  $3m + 1$  items, Table 1 shows that the difference  $\alpha - E(\alpha_3)$  is no more than 0.010 if we have at least 10 items, no more than 0.005 if we have at least 16 items, and no more than 0.001 if we have at least 34 items. If we have  $3m + 2$  items, Table 1 shows that the difference  $\alpha - E(\alpha_3)$  is no more than 0.010 if we have at least 11 items, no more than 0.005 if we have at least 17 items, and no more than 0.001 if we have at least 32 items. Hence, we may conclude that in the case that we distribute the items as evenly as possible, the difference  $\alpha - E(\alpha_3)$  is no more than 0.010 if we have at least 9 items, no more than 0.005 if we have

TABLE 2: The minimum number of items per split type for three criteria.

Split type	Minimum number of items		
	$\leq 0.010$	$\leq 0.005$	$\leq 0.001$
2-split	10	14	32
3-split	9	15	32
4-split	11	15	35
5-split	13	18	39
6-split	11	17	41
7-split	13	19	41

at least 15 items, and no more than 0.001 if we have at least 32 items. These numbers are presented in Table 2.

## 5. Conclusion

Alpha is the most commonly used coefficient for estimating reliability of a test score if there is only one test administration [5–9]. Cronbach [6] and Raju [18] have presented results that give a meaning to alpha and to its estimate from a sample, even if its assumptions do not hold. Cronbach [6] showed that if the 2-splits are made such that the parts have equal size, then alpha is the mean of all possible 2-split alphas. Raju [18] showed that if a  $k$ -split for  $k \geq 2$  is made such that the  $k$  parts have the same number of items, then alpha is the mean of all possible  $k$ -split alphas. The importance of these results lies in the fact that they provide a proper interpretation of alpha if (some of) its assumptions do not hold [2, 5].

In this paper, we consider the cases where the items cannot be split into parts of equal size. In these cases, the value of alpha exceeds the mean of all possible  $k$ -split alphas [18]. However, it was shown that if a  $k$ -split is made such that the items are divided as evenly as possible, the difference between alpha and the mean of all possible  $k$ -split alphas is negligible for sufficiently many items (Lemmas 2 and 3). Using numerical simulations, it was shown, in the case of a 2-split, that the difference is no more than 0.010 if we have at least 10 items, no more than 0.005 if we have at least 14 items, and no more than 0.001 if we have at least 32 items and, in the case of a 3-split, that the difference is no more than 0.010 if we have at least 9 items, no more than 0.005 if we have at least 15 items, and no more than 0.001 if we have at least 32 items.

The results in Cronbach [6] and Raju [18] provide interpretations of alpha if the number of items is a composite number  $n$ , for any proper divisor of  $n$ . The results and simulations in this paper provide interpretations of alpha if the number of items is prime.

## Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This research was done while the author was funded by the Netherlands Organisation for Scientific Research, Veni Project 451-11-026.

## References

- [1] F. M. Lord and M. R. Novick, *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, Mass, USA, 1968.
- [2] R. P. McDonald, *Test Theory: A Unified Treatment*, Erlbaum, Mahwah, NJ, USA, 1999.
- [3] W. Revelle and R. E. Zinbarg, "Coefficients alpha, beta, omega, and the glb: comments on Sijtsma," *Psychometrika*, vol. 74, no. 1, pp. 145–154, 2009.
- [4] H. G. Osburn, "Coefficient alpha and related internal consistency reliability coefficients," *Psychological Methods*, vol. 5, no. 3, pp. 343–355, 2000.
- [5] J. M. Cortina, "What is coefficient alpha? An examination of theory and applications," *Journal of Applied Psychology*, vol. 78, no. 1, pp. 98–104, 1993.
- [6] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951.
- [7] A. Field, *Discovering Statistics Using SPSS*, Sage, Los Angeles, Calif, USA, 3rd edition, 2009.
- [8] R. M. Furr and V. R. Bacharach, *Psychometrics. An Introduction*, Sage, Los Angeles, Calif, USA, 2008.
- [9] K. Sijtsma, "On the use, the misuse, and the very limited usefulness of Cronbach's alpha," *Psychometrika*, vol. 74, no. 1, pp. 107–120, 2009.
- [10] M. J. Warrens, "Some relationships between Cronbach's alpha and the Spearman-Brown formula," *Journal of Classification*. In press.
- [11] L. Guttman, "A basis for analyzing test-retest reliability," *Psychometrika*, vol. 10, pp. 255–282, 1945.
- [12] D. Grayson, "Some myths and legends in quantitative psychology," *Understanding Statistics*, vol. 3, no. 2, pp. 101–134, 2004.
- [13] A. P. Keszei, M. Novak, and D. L. Streiner, "Introduction to health measurement scales," *Journal of Psychosomatic Research*, vol. 68, no. 4, pp. 319–323, 2010.
- [14] M. Tavakol and R. Dennick, "Making sense of Cronbach's alpha," *International Journal of Medical Education*, vol. 2, pp. 53–55, 2011.
- [15] S. B. Green and S. L. Hershberger, "Correlated errors in true score models and their effect on coefficient alpha," *Structural Equation Modeling*, vol. 7, no. 2, pp. 251–270, 2000.
- [16] S. B. Green and Y. Yang, "Commentary on coefficient alpha: a cautionary tale," *Psychometrika*, vol. 74, no. 1, pp. 121–135, 2009.
- [17] Y. Sheng and Z. Sheng, "Is coefficient alpha robust to non-normal data?" *Frontiers in Psychology*, vol. 3, article 34, 2012.
- [18] N. S. Raju, "A generalization of coefficient alpha," *Psychometrika*, vol. 42, no. 4, pp. 549–565, 1977.
- [19] J. C. Flanagan, "A proposed procedure for increasing the efficiency of objective tests," *Journal of Educational Psychology*, vol. 28, no. 1, pp. 17–21, 1937.
- [20] P. J. Rulon, "A simplified procedure for determining the reliability of a test by split-halves," *Harvard Educational Review*, vol. 9, pp. 99–103, 1939.
- [21] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, NY, USA, 1970.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

