

# Chapter 6

## A New Expression and Interpretation of Coefficient Omega Under the Congeneric One-Factor Model



David J. Hessen

**Abstract** A new expression for the communality of the total score under the one-factor model is presented. In general, the communality of the total score is a lower bound to the reliability of the total score. Under the one-factor model, the communality of the total score also assesses the validity of the total score as a measure of the common factor. Conditions are given under which the new expression equals coefficient alpha. Furthermore, new expressions for the communality of an arbitrary item score and the proportion of total variance explained are derived under the one-factor model. For all new communality expressions, closed-form distribution-free estimates are provided. In an example, the closed-form estimates are calculated for a classic data set.

### 6.1 Introduction

Coefficient alpha (Guttman 1945; Cronbach 1951) is a very popular lower bound to the reliability of the total score (the unweighted sum of the item scores). Sijsma (2009) criticized the use of coefficient alpha for assessing the reliability of the total score and recommended the use of greater lower bounds, such as the greatest lower bound (Woodhouse and Jackson 1977; ten Berge et al. 1981) and coefficient lambda-2 (Guttman 1945). Despite the existence of greater lower bounds to the reliability of the total score, coefficient alpha continues to be used in practice. For an overview of many other lower bounds to the reliability of the total score, see Revelle and Zinbarg (2009).

It has been shown that coefficient alpha equals the communality of the total score if the item scores follow the essentially tau-equivalent model (Bentler 2009). The essentially tau-equivalent model is the unrealistic special case of the one-factor model in which all item scores have the same factor loading (Lord and

---

D. J. Hessen (✉)  
Utrecht University, Utrecht, The Netherlands  
e-mail: [d.j.hessen@uu.nl](mailto:d.j.hessen@uu.nl)

Novick 1968). Under the essentially tau-equivalent model, coefficient alpha is only equal to the reliability of the total score if all unique factors only contain random measurement error. However, under the essentially tau-equivalent model, the communality of the total score equals the proportion of variance of the total score explained by the common factor. This means that under the essentially tau-equivalent model, coefficient alpha assesses the validity of the total score as a measure of the common factor.

A more realistic model for the measurement of a single factor by a set of items than the essentially tau-equivalent model is the one-factor model (Spearman 1950). Under the one-factor model, factor loadings are not restricted to be equal. Since, in practice, the items of a subtest are usually constructed to measure one and the same latent factor, the item scores of a subtest are often assumed to follow the one-factor model. Under the one-factor model, the communality of the total score is given by coefficient omega (Heise and Bohrnstedt 1970; McDonald 1978) and equals the proportion of variance of the total score explained by the common factor. So under the one-factor model, coefficient omega assesses the validity of the total score as a measure of the common factor.

In this chapter, a new expression for the communality of the total score under the one-factor model is presented. Whereas coefficient omega expresses the communality of the total score in terms of factor model parameters, this new expression is in terms of the variances of the item scores, the covariances between the item scores, and the number of item scores. It is shown that the new expression equals coefficient alpha if the item scores follow the essentially tau-equivalent model. Furthermore, new expressions of the communality of an arbitrary item score and the proportion of total variance explained are derived under the one-factor model. Since all new expressions are functions of the population variances of the item scores and the population covariances between the item scores, distribution-free closed-form estimates are obtained by replacing the population parameters with sample analogues.

First, however, the one-factor model is briefly outlined in the next section. Subsequently, the new communality expressions and their closed-form estimates are presented. Finally, the closed-form estimates of the new communality expressions are calculated for a classic example data set.

## 6.2 The One-Factor Model

Let the random variables  $X_1, X_2, \dots, X_J$  be  $J$  item scores for a randomly selected individual from a population. The means of  $X_1, X_2, \dots, X_J$  are denoted by  $b_1, b_2, \dots, b_J$ ; the variances are denoted by  $\sigma_1^2, \sigma_2^2, \dots, \sigma_J^2$ ; and the covariance between two arbitrary item scores  $X_j$  and  $X_k$  is denoted by  $\sigma_{jk}$ , for all  $j$  and  $k \neq j$ . In the one-factor model, it is assumed that

$$X_j = b_j + a_j\xi + U_j, \quad \text{for all } j, \quad (6.1)$$

where  $a_j$  is a constant factor loading, for all  $j$ ,  $\xi$  is the common factor, and  $U_j$  is a unique factor, for all  $j$ . Note that  $U_j = S_j + E_j$ , where  $S_j$  is an item-specific factor (only varying between persons) and  $E_j$  is random measurement error (varying between persons and within persons), so that  $T_j = b_j + a_j\xi + S_j$  is the item true score. The common factor  $\xi$  is assumed to be independent of all unique factors  $U_1, U_2, \dots, U_J$ . The unique factors are assumed to be mutually independent. To identify the model, the variance of  $\xi$  is set to one. Let  $var(U_j) = \delta_j$ , for all  $j$ . Then, it follows that  $\sigma_j^2 = a_j^2 + \delta_j$ , for all  $j$ , and  $\sigma_{jk} = a_j a_k$ , for all  $j$  and  $k \neq j$ . Usually, the items are constructed such that it can be assumed that  $a_j > 0$ , for all  $j$ . If the items are not constructed this way, a transformation can be applied to some of the item scores such that it can be assumed that  $a_j > 0$ , for all  $j$ . Note that if  $a_j > 0$ , for all  $j$ , then  $\sigma_{jk} > 0$ , for all  $j$  and  $k \neq j$ .

### 6.3 Communality/Validity

Let  $C_j = b_j + a_j\xi$ , for all  $i$ . Then, the total score is given by  $X = \sum_j X_j = C + U$ , where  $C = \sum_j C_j = \sum_j b_j + \sum_j a_j\xi$  and  $U = \sum_j U_j$ . The communality of  $X$  is given by coefficient omega (Heise and Bohrnstedt 1970; McDonald, 1978) and is the squared correlation between  $X$  and  $C$ , that is,

$$\omega = \rho_{XC}^2 = \frac{var(C)}{\sigma_X^2} = \frac{\left(\sum_j a_j\right)^2}{\sum_j \sigma_j^2 + \sum_j \sum_{k \neq j} \sigma_{jk}} = \frac{\sum_j a_j^2 + \sum_j \sum_{k \neq j} a_j a_k}{\sum_j a_j^2 + \sum_j \delta_j + \sum_j \sum_{k \neq j} a_j a_k}, \quad (6.2)$$

where  $\sigma_X^2 = var(X)$ . Note that under the one-factor model,  $\rho_{XC}^2$  is equal to the squared correlation between  $X$  and  $\xi$  given by

$$\rho_{X\xi}^2 = \frac{\{cov(X, \xi)\}^2}{\sigma_X^2} = \frac{[E\{(\sum_j a_j\xi + \sum_j U_j)\xi\}]^2}{\sigma_X^2} = \frac{(\sum_j a_j)^2}{\sigma_X^2}. \quad (6.3)$$

From this, it can be concluded that under the one-factor model, the communality coefficient  $\rho_{XC}^2$  also assesses the validity of  $X$  as a measure of  $\xi$ . Now, since  $\sigma_{jk} = a_j a_k$ , for all  $j \neq k$ , it follows that

$$\frac{\sigma_{jk}\sigma_{jl}}{\sigma_{kl}} = \frac{a_j a_k a_j a_l}{a_k a_l} = a_j^2, \quad \text{for all } j, k \neq j, \text{ and } l \neq j, k.$$

Taking the average over all  $k \neq j$  and  $l \neq j, k$  gives

$$\frac{1}{(J-1)(J-2)} \sum_{k \neq j} \sum_{l \neq j, k} \frac{\sigma_{jk}\sigma_{jl}}{\sigma_{kl}} = a_j^2, \quad \text{for all } j. \quad (6.4)$$

Substitution from Eq. 6.4 and  $\sigma_{jk} = a_j a_k$  into

$$\rho_{XC}^2 = \frac{\sum_j a_j^2 + \sum_j \sum_{k \neq j} a_j a_k}{\sigma_X^2} \quad (6.5)$$

yields the new expression of the communality of the total score  $X$  given by

$$\rho_{XC}^2 = \left\{ \frac{1}{(J-1)(J-2)} \sum_j \sum_{k \neq j} \sum_{l \neq j, k} \frac{\sigma_{jk} \sigma_{jl}}{\sigma_{kl}} + \sum_j \sum_{k \neq j} \sigma_{jk} \right\} / \sigma_X^2. \quad (6.6)$$

Note that substitution from Eq. 6.4 into  $\sigma_j^2 = a_j^2 + \delta_j$  and solving for  $\delta_j$  yields

$$\delta_j = \sigma_j^2 - \frac{1}{(J-1)(J-2)} \sum_{k \neq j} \sum_{l \neq j, k} \frac{\sigma_{jk} \sigma_{jl}}{\sigma_{kl}}, \quad \text{for all } j. \quad (6.7)$$

Also note that if  $a_j > 0$ , for all  $i$ , then it follows from Eq. 6.4 that

$$a_j = \sqrt{\frac{1}{(J-1)(J-2)} \sum_{k \neq j} \sum_{l \neq j, k} \frac{\sigma_{jk} \sigma_{jl}}{\sigma_{kl}}}, \quad \text{for all } j. \quad (6.8)$$

Under the essentially tau-equivalent model,  $\sigma_{jk} = a^2$ , for all  $j$  and  $k \neq j$ , so that

$$\sum_{l \neq j, k} \frac{\sigma_{jk} \sigma_{jl}}{\sigma_{kl}} = (J-2) \sigma_{jk}, \quad \text{for all } j \text{ and } k \neq j. \quad (6.9)$$

Substitution from Eq. 6.9 into Eq. 6.6 and factoring  $\sum_j \sum_{k \neq j} \sigma_{jk}$  yields coefficient alpha given by

$$\alpha = \frac{J}{J-1} \sum_j \sum_{k \neq j} \sigma_{jk} / \sigma_X^2. \quad (6.10)$$

In addition to the communality of the total score, the communalities of the individual item scores might be of interest in practice. The communality of item score  $X_j$  is defined as the squared correlation between  $X_j$  and  $C_j$ . Under the one-factor model, the communality of item score  $X_j$  is given by

$$h_j^2 = \rho_{X_j C_j}^2 = \frac{\text{cov}(X_j, C_j)^2}{\sigma_j^2 \text{var}(C_j)} = \frac{\text{var}(C_j)}{\sigma_j^2} = \frac{a_j^2}{a_j^2 + \delta_j}, \quad \text{for all } j. \quad (6.11)$$

Note that under the one-factor model,  $\rho_{X_j C_j}^2$  is equal to  $\rho_{X_j \xi}^2$ . So under the one-factor model, the communality of item score  $X_j$  also assesses the validity of  $X_j$  as

a measure of  $\xi$ . Now, dividing the left-hand side of Eq. 6.4 by  $\sigma_j^2$  yields

$$h_j^2 = \frac{1}{(J-1)(J-2)\sigma_j^2} \sum_{k \neq j} \sum_{l \neq j, k} \frac{\sigma_{jk}\sigma_{jl}}{\sigma_{kl}}, \text{ for all } j. \tag{6.12}$$

The total variance is defined as  $\sum_j \sigma_j^2$ . Under the one-factor model, the proportion of total variance explained by the common factor is given by

$$\pi = \frac{\sum_j a_j^2}{\sum_j a_j^2 + \sum_j \delta_j} = \frac{\sum_j a_j^2}{\sum_j \sigma_j^2} \tag{6.13}$$

and assesses the extent to which the items measure the common factor relative to the unique factors. Substitution from Eq. 6.4 into Eq. 6.13 yields

$$\pi = \frac{1}{(J-1)(J-2)} \sum_j \sum_{k \neq j} \sum_{l \neq j, k} \frac{\sigma_{jk}\sigma_{jl}}{\sigma_{kl}} / \sum_j \sigma_j^2. \tag{6.14}$$

### 6.3.1 Estimates

Let  $x_{ij}$  be the observed score of individual  $i = 1, 2, \dots, N$  on item  $j = 1, 2, \dots, J$ . The sample mean score on item  $j$  is given by  $\bar{x}_j = \sum_{i=1}^N x_{ij}/N$ , for all  $j$ . The observed total score of individual  $i$  is then given by  $x_i = \sum_{j=1}^J x_{ij}$ , for all  $i$ , and the sample mean total score is then given by  $\bar{x} = \sum_{i=1}^N x_i/N$ . A closed-form estimate of  $\rho_{XC}^2$  is now given by

$$\hat{\rho}_{XC}^2 = \left\{ \frac{1}{(J-1)(J-2)} \sum_j \sum_{k \neq j} \sum_{l \neq j, k} \frac{s_{jk}s_{jl}}{s_{kl}} + \sum_j \sum_{k \neq j} s_{jk} \right\} / s^2, \tag{6.15}$$

where  $s_{jk} = \sum_{i=1}^N (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)/(N-1)$  is the estimate of  $\sigma_{jk}$ , for all  $j$  and  $k \neq j$ , and  $s^2 = \sum_{i=1}^N (x_i - \bar{x})^2/(N-1)$  is the estimate of  $\sigma_X^2$ . Note that  $s^2 = \sum_j s_j^2 + \sum_j \sum_{k \neq j} s_{jk}$ , where  $s_j^2 = \sum_{i=1}^N (x_{ij} - \bar{x}_j)^2/(N-1)$  is the estimate of  $\sigma_j^2$ , for all  $j$ . A closed-form estimate of  $\delta_j$  is given by

$$\hat{\delta}_j = s_j^2 - \frac{1}{(J-1)(J-2)} \sum_{k \neq j} \sum_{l \neq j, k} \frac{s_{jk}s_{jl}}{s_{kl}}, \text{ for all } j. \tag{6.16}$$

If  $s_{jk} > 0$ , for all  $j \neq k$ , then a closed-form estimate of factor loading  $a_j$  is given by

$$\hat{a}_j = \sqrt{\frac{1}{(J-1)(J-2)} \sum_{k \neq j} \sum_{l \neq j, k} \frac{s_{jk}s_{jl}}{s_{kl}}}, \text{ for all } j. \tag{6.17}$$

A closed-form estimate of the item communality  $\rho_{X_j C_j}^2 = h_j^2$  is given by

$$\hat{h}_j^2 = \frac{1}{(J-1)(J-2)s_j^2} \sum_{k \neq j} \sum_{l \neq j,k} \frac{s_{jk}s_{jl}}{s_{kl}}, \text{ for all } j. \quad (6.18)$$

Finally, a closed-form estimate of the proportion of total variance explained by the common factor is given by

$$\hat{\pi} = \frac{1}{(J-1)(J-2)} \sum_j \sum_{k \neq j} \sum_{l \neq j,k} \frac{s_{jk}s_{jl}}{s_{kl}} / \sum_j s_j^2. \quad (6.19)$$

## 6.4 An Example

The data in this example are taken from Lord and Novick (1968, p. 91) and are the entries of the sample covariance matrix of four measures of English as a foreign language. The sample covariance matrix is based upon a sample size of 1416 and is given by

$$\begin{bmatrix} s_1^2 & & & \\ s_{21} & s_2^2 & & \\ s_{31} & s_{32} & s_3^2 & \\ s_{41} & s_{42} & s_{43} & s_4^2 \end{bmatrix} = \begin{bmatrix} 94.7 & & & \\ 87.3 & 212.0 & & \\ 63.9 & 138.7 & 160.5 & \\ 58.4 & 128.2 & 109.8 & 115.4 \end{bmatrix}.$$

Estimates of the parameters of the one-factor model are often obtained by maximum likelihood estimation under the assumption of multivariate normality of the item scores in the population. For comparison, both the maximum likelihood estimates and the closed-form estimates of  $a_j$ ,  $\delta_j$ , and  $h_j^2$ , for all  $j$ , and  $\omega$  and  $\pi$  are calculated. The maximum likelihood estimates are denoted by  $\tilde{a}_j$  and  $\tilde{\delta}_j \tilde{h}_j^2$ , for all  $j$ , and  $\tilde{\omega}$  and  $\tilde{\pi}$ . The estimates for all item parameters and coefficients are given in Table 6.1. Note that the item order given by the closed-form estimates  $\hat{h}_1^2 < \hat{h}_3^2 < \hat{h}_4^2 < \hat{h}_2^2$  is different from the item order given by the maximum likelihood estimates  $\tilde{h}_1^2 < \tilde{h}_3^2 < \tilde{h}_2^2 < \tilde{h}_4^2$ . The estimate of coefficient  $\alpha$  is  $\hat{\alpha} = .891$ . The maximum likelihood estimate of coefficient  $\omega$  is  $\tilde{\omega} = .909$ , and its closed-form estimate is  $\hat{\omega} = .912$ . So, about 91% of the sample variance of the total score is explained by the common factor. The maximum likelihood estimate of the total variance  $\pi$  is  $\tilde{\pi} = .725$ , and its closed-form estimate is  $\hat{\pi} = .735$ . So, about 73% of the total sample variance is explained by the common factor.

**Table 6.1** Estimates of all item parameters and coefficients under the one-factor model, for the Lord and Novick (1968) example data

$j$	Estimate					
	$\hat{a}_j$	$\tilde{a}_j$	$\hat{\delta}_j$	$\tilde{\delta}_j$	$\hat{h}_j^2$	$\tilde{h}_j^2$
1	6.164	6.133	56.708	57.020	.401	.397
2	13.455	12.939	30.975	44.432	.854	.790
3	10.653	10.882	47.015	41.963	.707	.738
4	9.791	9.937	19.534	16.581	.831	.856

## 6.5 Conclusion

Coefficient alpha has traditionally been used to assess the reliability of the total score. Since coefficient alpha equals the communality of the total score under the essentially tau-equivalent model, coefficient alpha is a lower bound to the reliability of the total score. The communality of the total score under the more realistic one-factor model is also a lower bound to the reliability of the total score. Under the one-factor model, however, the communality of the total score equals the proportion of variance of the total score explained by the common factor and therefore assesses the extent to which the common factor is measured by the total score. If items are constructed to measure one and the same latent factor, then the one-factor model can be used to study whether the items actually measure a single common factor. Once it has been concluded that the items measure a single common factor, it is of interest to assess how well the single common factor is measured by the item scores or the total score. To assess how well the single common factor is measured by the total score, coefficient omega and its new expression can be used. Under the one-factor model, coefficient omega and its new expression give the proportion of variance of the total score explained by the common factor. In practice, the maximum likelihood estimate of coefficient omega is often used as the estimate of the communality of the total score under the one-factor model. The closed-form estimate of omega now provides a distribution-free alternative.

## References

- Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, 74(1), 137–143. <https://doi.org/10.1007/s11336-008-9100-1>.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255–282. <https://doi.org/10.1007/BF02288892>.
- Heise, D. R., & Bohrnstedt, G. W. (1970). Validity, invalidity, and reliability. In E. F. Borgatta (Ed.), *Sociological methodology* (pp. 104–129). Jossey-Bass.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.
- McDonald, R. P. (1978). Generalizability in factorable domains: “domain validity and generalizability”: 1. *Educational and Psychological Measurement*, 38(1), 75–79. <https://doi.org/10.1177/001316447803800111>.

- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: comments on Sijtsma. *Psychometrika*, *74*(1), 145–154. <https://doi.org/10.1007/s11336-008-9102-z>.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107–120. <https://doi.org/10.1007/s11336-008-9101-0>.
- Spearman, C. (1950). *Human ability*. Macmillan.
- ten Berge, J. M. F., Snijders, T. A. B., & Zegers, F. E. (1981). Computational aspects of the greatest lower bound to reliability and constrained minimum trace factor analysis. *Psychometrika*, *46*(2), 201–213. <https://doi.org/10.1007/BF02293900>.
- Woodhouse, B., & Jackson, P. H. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: II: A search procedure to locate the greatest lower bound. *Psychometrika*, *42*(4), 579–591. <https://doi.org/10.1007/BF02295980>.