

Testing the Difference Between Reliability Coefficients Alpha and Omega

Educational and Psychological
Measurement

2017, Vol. 77(2) 185–203

© The Author(s) 2016

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0013164416658325

journals.sagepub.com/home/epm



Lifang Deng¹ and Wai Chan²

Abstract

Reliable measurements are key to social science research. Multiple measures of reliability of the total score have been developed, including coefficient alpha, coefficient omega, the greatest lower bound reliability, and others. Among these, the coefficient alpha has been most widely used, and it is reported in nearly every study involving the measure of a construct through multiple items in social and behavioral research. However, it is known that coefficient alpha underestimates the true reliability unless the items are tau-equivalent, and coefficient omega is deemed as a practical alternative to coefficient alpha in estimating measurement reliability of the total score. However, many researchers noticed that the difference between alpha and omega is minor in applications. Since the observed differences in alpha and omega can be due to sampling errors, the purpose of the present study, therefore, is to propose a method to evaluate the difference of coefficient alpha ($\hat{\alpha}$) and omega ($\hat{\omega}$) statistically. In particular, the current article develops a procedure to estimate the *SE* of ($\hat{\omega} - \hat{\alpha}$) and consequently the confidence interval (CI) for ($\omega - \alpha$). This procedure allows us to test whether the observed difference ($\hat{\omega} - \hat{\alpha}$) is due to sample error or $\hat{\omega}$ is significantly greater than $\hat{\alpha}$. The developed procedure is then applied to multiple real data sets from well-known scales to empirically verify the values of ($\hat{\omega} - \hat{\alpha}$) in practice. Results showed that in most of the comparisons the differences are significantly above zero but cases also exist where the CIs contain zero. An R program for calculating $\hat{\omega}$, $\hat{\alpha}$, and the *SE* of ($\hat{\omega} - \hat{\alpha}$) is also included in the present study so that the developed procedure is easily accessible to applied researchers.

¹Beihang University, Beijing, China

²The Chinese University of Hong Kong, Hong Kong, China

Corresponding Author:

Lifang Deng, Department of Psychology, Beihang University, Beijing 100191, China.

Email: Lifangdeng@buaa.edu.cn

Keywords

coefficient alpha, coefficient omega, standard errors, confidence intervals

Introduction

In social and behavioral sciences, most interesting attributes such as happiness, anxiety, and cognitive and social competence cannot be observed directly and have to be measured by multiple indicators that are subject to errors. Reliable measurements are key to social science research. When measurements are used quantitatively, we would like the observed differences between individuals to be due to the differences in true scores rather than due to measurement errors. The concept *reliability* was invented to quantify the quality of measurements for such a purpose (see, e.g., Allen & Yen, 1979; Raykov & Marcoulides, 2011). In particular, reliability of the observed score is defined as the ratio of the variance of the true score over the variance of the observed score. It is known that reliability depends on the formulation of items (how questions are phrased in questionnaires) as well as the targeted population (the composition of the participants; see, e.g., Thompson, 2003).

In practice, the total score across items is most widely used for analysis, and the reliability of the total score is of great interest. However, because the part of the true score within the total score is not observable, it is not straightforward to estimate its variance even if we have data with the observed scores. Consequently, multiple measures of reliability of the total score have been developed, including coefficient alpha (also referred to as Cronbach's alpha; Cronbach, 1951; see also Cortina, 1993; Raykov, 1997; Raykov & Marcoulides, 2011, 2015), coefficient omega (McDonald, 1999), the greatest lower bound reliability (Bentler, 1972; Bentler & Woodward, 1980; Li, Rosenthal, & Rubin, 1996), and others (Allen & Yen, 1979; Hunt & Bentler, 2015; Zinbarg, Revelle, Yovel, & Li, 2005). Among these, the coefficient alpha has been the most widely used measure of reliability, and it is reported in nearly every study involving the measure of a construct through multiple items in social and behavioral research. However, there exists criticism against coefficient alpha (see, e.g., Green, Lissitz, & Mulaik, 1977; Raykov, 1997; Sijtsma, 2009; Yang & Green, 2011). This is because, when the items are unidimensional (measuring the same latent trait), the sample coefficient alpha yields consistent estimate of reliability only when all the items have equal covariance with the true score, called tau-equivalence. But this assumption is seldom met in practice with educational and psychological scales (see, e.g., Green & Yang, 2009; Jöreskog, 1971; Lord & Novick, 1968). A measure that overcomes the deficiencies of alpha is coefficient omega, which is based on a one-factor model. In particular, when the covariance among the items can be approximately accounted for by a one-factor model, the formulation of coefficient omega closely matches the definition of reliability (McDonald, 1999). Also, almost all free and commercial statistical software outputs the parameter estimates of a one-factor model that allow the calculation of coefficient omega (e.g., Dunn, Baguley, & Brunnsden, 2014; Zhang & Yuan, 2016). In particular, the coefficient omega is

deemed as a practical alternative to coefficient alpha in estimating measurement reliability of the total score (Dunn et al., 2014). However, most applied researchers still choose to report only the coefficient alpha.

The reason behind the wide usage of coefficient alpha can be due to multiple reasons. First, Cronbach's coefficient alpha is well known but poorly understood by many applied researchers. Majority of researchers and users of psychometrical scales might not understand the differences between α and ω very well. A great number of researchers' understanding of reliability analysis is generally low, and it remains likely that fewer than half of all postgraduate courses in psychology offer in-depth coverage of methods of reliability analysis (Dunn et al., 2014), and thus alpha is widely misapplied in social science research (Cho & Kim, 2015; Green & Yang, 2009). In addition, articles that condemn alpha tend to be very technical, if implementation of an alternative is offered, it is usually presented in a manner too complex for applied researchers to easily implement. Raykov and colleagues made a great deal of effort for a balanced treatment about the criticism and misapplication of alpha (Raykov, 1997, 2012; Raykov & Marcoulides, 2011, 2015). And above all, although the advantages of omega have been illustrated by various authors, the difference between alpha and omega has been reported to be small in applications (e.g., Maydeu-Olivares, Coffman, & Hartmann, 2007; Raykov, 1997). Such an observation made the use of coefficient omega less appealing and indirectly promoted the use of coefficient alpha. However, the observation was primarily based on the direct comparison of the nominal values of the two estimates without making appropriate statistical inference.

Raykov and Marcoulides (2015) provided a direct approach to point and interval estimation of Cronbach's coefficient alpha using Mplus. They concluded that "alpha and the reliability of a considered scale can be treated as practically identical at large" (p. 152) when the following four conditions hold: (a) items are unidimensional or there is no correlated errors when fitted by the one-factor model; (b) the average loading is above .7; (c) all the differences between the individual factor loadings and the average loading are less than .2; and (d) each item has zero specificity or each uniqueness is solely from measurement errors. When the four conditions in (a) to (d) hold, alpha can also be treated as practically identical to omega, and there is no need for additional development, as presented in the current article. However, in practice, it is very likely that the four conditions may not hold simultaneously. This will be further noted in our analysis of real data sets from well-known psychological scales. In such practical situations, coefficient omega may enjoy some advantage over coefficient alpha (Dunn et al., 2014; Zhang & Yuan, 2016), and then the difference between alpha and omega may become nontrivial. The technique developed in this article allows us to statistically evaluate whether the observed difference ($\hat{\omega} - \hat{\alpha}$) is due to sampling error. Of course, a significant ($\hat{\omega} - \hat{\alpha}$) does not imply that $\hat{\omega}$ is a consistent estimate of the true reliability in the population, nor is $\hat{\alpha}$. It still needs conditions (a) and (d) to hold for omega to equal the true population reliability,

whereas it needs all the four conditions (a to d) to hold for alpha to be practically identical to the population reliability.

Evaluating the difference of coefficient alpha and omega statistically will help enhance applied researchers' awareness of the advantage of omega and facilitate the shift from alpha to omega. Many studies discussed problems with coefficient alpha and pointed out the assumptions underlying coefficient alpha are unlikely to hold in practice, and violation of these assumptions can result in nontrivial negative or positive bias, coefficient omega has been shown to be a more sensible index of internal consistency (Dunn et al., 2014; Green & Yang, 2009; Zhang & Yuan, 2016). However, alpha continues to be widely applied by social science research to assess internal consistency reliability. Actually, Zinbarg et al. (2005) reported that even when the assumptions of essentially tau-equivalent model are met, omega performs at least as well as alpha. But under violations of tau-equivalence conditions likely to be the norm in psychology, omega outperforms alpha and is clearly the preferred choice. To change the practice, it is necessary to develop a rational, scientific, and convincing method to compare the difference of coefficient alpha and omega statistically in applications, and shift applied researchers' attention from alpha to omega (Dunn et al., 2014).

The purpose of the present study, therefore, is to propose a method to evaluate the difference of the sample coefficient alpha ($\hat{\alpha}$) and omega ($\hat{\omega}$) statistically. Since the estimates of the two coefficients are obtained from the same sample and consequently are correlated, we cannot judge the significance of their difference by the standard error (*SE*) of either coefficient, or their combination. The covariance of $\hat{\omega}$ and $\hat{\alpha}$ has to be included as well. Consider the wide application of $\hat{\omega}$ and $\hat{\alpha}$ in social research as well as the ongoing debate between the two coefficients, obtaining the *SE* of $(\hat{\omega} - \hat{\alpha})$ and consequently the confidence interval (*CI*) for $(\omega - \alpha)$ will have a great impact. In this study, we aim to (a) develop a formula to estimate the *SE* of $(\hat{\omega} - \hat{\alpha})$ and consequently the *CI* for $(\omega - \alpha)$, which can be easily implemented in any software with exploratory factor analysis or confirmatory factor analysis already built in. (b) Apply the proposed procedure to multiple real data sets from well-known scales and empirically see how often the difference between the two reliability coefficients is significant. In particular, in developing a method to estimate the *SE* of $(\hat{\omega} - \hat{\alpha})$, we will consider that data are not necessarily normally distributed in practice. (c) An R program for calculating $\hat{\omega}$, $\hat{\alpha}$, and the *SE* of $(\hat{\omega} - \hat{\alpha})$ is also included in the present study so that the developed procedure is easily accessible to applied researchers. R code is free and everybody has access to it.

The method to be developed allows us to test whether the observed difference $(\hat{\omega} - \hat{\alpha})$ is due to sampling error or ω is truly greater than α , it is a rational and scientific approach for applied researcher to examine the difference between coefficient alpha and omega. The present study also applies the developed procedure to major or well-known psychological scales to empirically verify their differences. In particular, if in most applications the 95% confidence intervals contain 0, then the criticism against coefficient alpha is not warranted. By contrast, if most of the confidence

intervals do not contain 0, then they will convey a clear message to applied researchers, and will also more effectively promote the applications of coefficient omega. In any case, the knowledge about the *SE* of $(\hat{\omega} - \hat{\alpha})$ will allow us to more effectively quantify and evaluate the difference between the two reliability coefficients. We will describe how the *SE* of $(\hat{\omega} - \hat{\alpha})$ is obtained in the next section. Then we will introduce four real data sets. Results of applying the developed procedures to these data sets are presented and discussed in the following section. Technical details leading to the formulations of the *SE* is given in the appendix. An R program for calculating $\hat{\omega}$, $\hat{\alpha}$, and the *SE* of $(\hat{\omega} - \hat{\alpha})$ can be downloaded at https://www.psy.cuhk.edu.hk/psy-media/WChan_Page/alpha-omega.txt, so that applied researchers can easily implement the developed procedure.

Methodology

In this section, we will first describe how the *SE* of $(\hat{\omega} - \hat{\alpha})$ is obtained, then introduce four data sets. Each data set is either classical or from well-known scales.

The Formulations of $\hat{\omega}$, $\hat{\alpha}$, and the *SE* of $(\hat{\omega} - \hat{\alpha})$

Let $\mathbf{S} = (s_{jk})$ be the sample covariance matrix based on p items, and $\Sigma = (\sigma_{ij})$ be the population counterpart of \mathbf{S} . The formula of sample coefficient alpha is

$$\hat{\alpha} = \frac{P}{p - 1} \left(1 - \sum_{j=1}^p s_{jj} / \sum_{j=1}^p \sum_{k=1}^p s_{jk} \right). \tag{1}$$

It is a consistent estimate of its population counterpart α when s_{jk} is replaced by σ_{jk} . In Equation (1), $\hat{\alpha}$ is a function of the elements of \mathbf{S} , and consequently the sampling properties of $\hat{\alpha}$ is determined by those of \mathbf{S} .

The computation of $\hat{\omega}$ involves a one-factor model, where exploratory factor model and confirmatory factor model are the same. Let the factor variance be fixed at 1.0 for identification purpose, and the factor loadings and error variances for the p items be $\lambda_1, \lambda_2, \lambda_p$, and $\psi_{11}, \psi_{22}, \dots, \psi_{pp}$, respectively. Then the formula¹ of sample coefficient omega is

$$\hat{\omega} = \frac{\left(\sum_{j=1}^p \hat{\lambda}_j \right)^2}{\left(\sum_{j=1}^p \hat{\lambda}_j \right)^2 + \sum_{j=1}^p \hat{\psi}_{jj}}, \tag{2}$$

where $\hat{\lambda}_j$ and $\hat{\psi}_{jj}$ are, respectively, the estimates of λ_j and ψ_{jj} obtained by fitting the (one-factor) model to data via minimizing a discrepancy function between the model implied covariance matrix $\hat{\Sigma}$ and the sample covariance matrix \mathbf{S} . In this article, we will use the normal-distribution-based maximum likelihood method to estimate the factor model, because it is the default method in most commercial and free software, and also the most widely used in practice. In the estimation process, the estimates $\hat{\lambda}_j$

and $\hat{\psi}_{jj}$ will take values so that the model implied covariance matrix best matches the sample covariance matrix. When elements of \mathbf{S} changes, the elements of the model implied covariance matrix also change accordingly. Thus, the estimates $\hat{\lambda}_j$ and $\hat{\psi}_{jj}$ are functions of \mathbf{S} (see, e.g., Yuan, Marshall, & Bentler, 2003). Consequently, $\hat{\omega}$ is also a function of \mathbf{S} , and the sampling properties of $\hat{\omega}$ is determined by those of \mathbf{S} .

Since both $\hat{\alpha}$ and $\hat{\omega}$ are functions of \mathbf{S} , their difference ($\hat{\omega} - \hat{\alpha}$) is also a function of \mathbf{S} , and the sampling properties of ($\hat{\omega} - \hat{\alpha}$) is determined by those of \mathbf{S} as well. In the appendix, the *SE* of ($\hat{\omega} - \hat{\alpha}$) is obtained by utilizing such a relationship. In particular, the variance of ($\hat{\omega} - \hat{\alpha}$) is obtained by deducting two times the covariance from the sum of the variances of $\hat{\omega}$ and $\hat{\alpha}$. Each term is computed by approximating $\hat{\omega}$ and $\hat{\alpha}$ using linear functions of \mathbf{S} , and such a technique has been widely used in studying properties of parameter estimates in structural equation models and elsewhere (e.g., Bentler & Dijkstra, 1985; Yuan, Guarnaccia, & Hayslip, 2003). Actually, the formulas of variances or *SEs* of $\hat{\alpha}$ and $\hat{\omega}$ have been described in the literature (e.g., Yuan & Bentler, 2002), and the contribution of the development in the appendix is mostly on computing the covariance between $\hat{\alpha}$ and $\hat{\omega}$.

The complete details leading to the asymptotic distribution of $\hat{\omega} - \hat{\alpha}$ and consequently the formula of confidence interval for ($\omega - \alpha$) are given in the appendix, and an R program calculating $\hat{\omega}$, $\hat{\alpha}$, and the *SE* of ($\hat{\omega} - \hat{\alpha}$) is provided online at https://www.psy.cuhk.edu.hk/psy_media/WChan_Page/alpha-omega.txt.

The Description of the Four Real Data Sets

We now describe the real data sets, and each also contains multiple subscales. Results on ($\hat{\omega} - \hat{\alpha}$) and its *SE* applying to each of the subscales will be reported in the following section.

Data Set 1. This data set was adopted from Holzinger and Swineford (1939), who developed a battery of 26 items, aiming to evaluate 5 cognitive traits of middle school students. The 5 traits are *spatial* (Items 1 to 4, Items 25 and 26), *verbal* (Items 5 to 9), *speed* (Items 10 to 13), *memory* (Items 14 to 19), and *math* (Items 20 to 24). Holzinger and Swineford reported two data sets. We will use the one with $N = 145$ students from the Grant-White school to examine the differences between the reliability coefficient for each of the 5 subscales.

Data Set 2. The data set consists of 44 items of the Big Five Inventory (BFI; John & Srivastava, 1999) with 5 subscales: *neuroticism* (*neuro*, 8 items), *extraversion* (*extra*, 8 items), *conscientiousness* (*cons*, 9 items), *openness* (*open*, 10 items), and *agreeableness* (*agree*, 9 items). Data are from administering the BFI questionnaires to college students from a midwestern private university in the United States, with $N = 190$ complete cases (Deng, Wang, & Zhao, 2016). Participants were recruited by campus flyers. A consent form was signed by each participant before data collection.

Data Set 3. The Humor Styles Questionnaire (Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003) has 4 subscales: *affiliative* (*affili*, 8 items), *self-enhancing* (*self-enha*, 8 items), *aggressive* (*aggress*, 8 items), and *self-defeating* (*self-defe*, 8 items). Each item is rated on a 5-point scale where 1 = *Never or very rarely true*, 2 = *Rarely true*, 3 = *Sometimes true*, 4 = *Often true*, 5 = *Very often or always true* (−1 = *Did not select an answer*). This data set is publicly available online (http://personality-testing.info/_raw_data) and was downloaded in the spring of 2015, with $N = 993$ complete cases.

Data Set 4. The data set comes from the Family Adaptability and Cohesion Evaluation Scales (FACES II; Olson, Portner, & Bell, 1982) that has 2 subscales: *cohesion* (16 items) and *adaptability* (*adapt*, 13 items), and each item is rated on a 5-point scale where 1 = *Never or very rarely true*, 2 = *Rarely true*, 3 = *Sometimes true*, 4 = *Often true*, 5 = *Very often or always true*. The data set is from administering the FACES scale to students from six colleges in Beijing, with $N = 852$ complete cases (Deng & Zheng, 2012).

Results of the Analysis of Real Data

In this section, we present the results of applying the methodology described in the previous section to the four data sets. The results include $\hat{\omega}$, $\hat{\alpha}$, $(\hat{\omega} - \hat{\alpha})$, and their *SEs*. In particular, two *SEs* are reported for each estimate. One is based on the assumption of normally distributed data and the other is based on a sandwich-type variance and is asymptotically distribution free (see, e.g., Maydeu-Olivares et al., 2007; Yuan et al., 2003). A CI for $(\omega - \alpha)$ corresponding to each *SE* is reported as well. Before computing $\hat{\omega}$, a one-factor model was fitted to each subscale to evaluate its unidimensionality. In particular, both the likelihood ratio statistic (T_{ml}) and the Satorra and Bentler (1994) rescaled statistic (T_{rml}) are included and so are their corresponding fit indices, comparative fit index (CFI; Bentler, 1990) and root mean square error of approximation (RMSEA; Steiger & Lind, 1980). These measures allow us to see whether the difference between omega and alpha is related to unidimensionality of the items.

The results are presented in Tables 1 to 4 corresponding to the Data Sets 1 to 4, respectively. Each table contains two parts. The upper panel contains fit statistics T_{ml} and T_{rml} , and their corresponding p value, CFI, and RMSEA.² The lower panel contains the estimate of $\hat{\omega}$, $\hat{\alpha}$, $(\hat{\omega} - \hat{\alpha})$, and their *SEs*. In particular, all the CIs are obtained with a nominal coverage rate of 95%.

For the five cognitive subscales (*spatial*, *verbal*, *speed*, *memory*, *math*) of Holzinger and Swineford's (1939) data, with results in Table 1, the p values corresponding to T_{rml} are respectively .206, .023, .007, .118, and .562, and RMSEA = .049, .105, .165, .063, and .000, respectively, suggesting that the items measuring subscales *spatial*, *memory*, and *math* can be regarded as approximately unidimensional whereas those measuring *verbal* and *speed* are not. However, according to the respective values of CFI (.977, .956, .927, .956, 1.000), the items on subscales *verbal* and *speed* can also be regarded as approximately unidimensional.

Table I. Data Set I: Twenty-Six Cognitive Items With Five Subscales (N = 145).

Subscale	# of items	Statistic T_{nm}					Rescaled statistic T_{nm}				
		df	T_{nm}	p_v	CFI	RMS	T_{nm}	p_v	CFI	RMS	
Spatial	6	9	13.723	.133	0.973	0.060	12.132	.206	0.977	0.049	
Verbal	5	5	14.698	.012	0.977	0.116	12.996	.023	0.966	0.105	
Speed	4	2	11.002	.004	0.947	0.177	9.857	.007	0.927	0.165	
Memory	6	9	14.784	.097	0.960	0.067	14.109	.118	0.956	0.063	
Math	5	5	4.057	.541	1.000	0.000	3.912	.562	1.000	0.000	
Subscale	# of items	Reliability	Estimate (SE_{nm} & SE_{sw})			$[L_{-nm}, U_{nm}]$			$[L_{-sw}, U_{sw}]$		
Spatial	6	ω	0.773 (0.031 & 0.033)	[0.713, 0.833]		[0.709, 0.837]		[0.709, 0.837]		[0.709, 0.837]	
		α	0.736 (0.029 & 0.032)	[0.680, 0.792]		[0.673, 0.799]		[0.673, 0.799]		[0.673, 0.799]	
Verbal	5	$\omega - \alpha$	0.037 (0.007 & 0.006)	[0.023, 0.052]		[0.025, 0.050]		[0.025, 0.050]		[0.025, 0.050]	
		ω	0.884 (0.017 & 0.018)	[0.852, 0.917]		[0.849, 0.919]		[0.849, 0.919]		[0.849, 0.919]	
		α	0.839 (0.016 & 0.017)	[0.809, 0.870]		[0.807, 0.872]		[0.807, 0.872]		[0.807, 0.872]	
Speed	4	$\omega - \alpha$	0.045 (0.008 & 0.009)	[0.030, 0.060]		[0.028, 0.062]		[0.028, 0.062]		[0.028, 0.062]	
		ω	0.779 (0.032 & 0.032)	[0.717, 0.842]		[0.717, 0.842]		[0.717, 0.842]		[0.717, 0.842]	
		α	0.754 (0.030 & 0.032)	[0.696, 0.813]		[0.691, 0.818]		[0.691, 0.818]		[0.691, 0.818]	
Memory	6	$\omega - \alpha$	0.025 (0.009 & 0.010)	[0.007, 0.043]		[0.005, 0.045]		[0.005, 0.045]		[0.005, 0.045]	
		ω	0.707 (0.040 & 0.048)	[0.629, 0.784]		[0.613, 0.800]		[0.613, 0.800]		[0.613, 0.800]	
		α	0.694 (0.036 & 0.043)	[0.624, 0.764]		[0.610, 0.778]		[0.610, 0.778]		[0.610, 0.778]	
Math	5	$\omega - \alpha$	0.013 (0.009 & 0.008)	[-0.005, 0.030]		[-0.003, 0.028]		[-0.003, 0.028]		[-0.003, 0.028]	
		ω	0.763 (0.035 & 0.035)	[0.694, 0.833]		[0.694, 0.833]		[0.694, 0.833]		[0.694, 0.833]	
		α	0.698 (0.030 & 0.031)	[0.640, 0.756]		[0.636, 0.760]		[0.636, 0.760]		[0.636, 0.760]	
		$\omega - \alpha$	0.065 (0.015 & 0.015)	[0.035, 0.095]		[0.037, 0.094]		[0.037, 0.094]		[0.037, 0.094]	

Note. p_v = p value; RMS = RMSEA; nm = based on normal distribution assumption; sw = sandwich-type covariance matrix.

Clearly, except for the subscale *memory*, the CI of $(\omega - \alpha)$ for the other four subscales do not contain 0, suggesting that there are significant difference between the two reliability coefficients. That is, $\hat{\omega}$ is significantly greater than $\hat{\alpha}$ in subscales *spatial*, *verbal*, *speed*, and *math*.

The results for the Big Five Inventory are in Table 2, and the items on each of the five subscales (*neuroticism*, *extraversion*, *conscientiousness*, *openness*, *agreeableness*) are not well fitted by a one-factor model (CFI = .855, .943, .915, .736, and .825, respectively; and RMSEA = .136, .108, .087, .138, and .087, respectively). Except for the subscale *openness*, the CIs for $(\omega - \alpha)$ for the four other subscales do not contain 0 (95% CI: [.001, .007], [.001, .008], [-.011, -.001], [.004, .022]), suggesting that $\hat{\omega}$ is significantly greater than $\hat{\alpha}$ in subscales *neuroticism*, *extraversion*, and *agreeableness*, but the opposite holds for *conscientiousness*.

The results for the Humor style questionnaires are in Table 3, where the numbers indicate that there are significant difference between the two reliability coefficients in three out of the four subscales (*affiliative*, *self-enhancing*, *aggressive*, *self-defeating*). Only the CI for $(\omega - \alpha)$ corresponding to *affiliative* literally covers zero. Moreover, the fit indices RMSEA and CFI suggested that the one-factor model fits the items of each of the four subscales reasonably well although not excellent (CFI = .918, .912, .940, and .953, respectively; RMSEA = .082, .096, .067, and .069, respectively).

Results with the last data set (family functioning with 2 subscales: *cohesion* and *adaptability*) are in Table 4, where the values of RMSEA (.058 and .049) suggest that the items on each of the subscales might be regarded as approximately unidimensional. However, the values of CFI (.887 and .925, respectively) suggests that items on *cohesion* is poorly fitted by the one-factor model. The CIs for $(\omega - \alpha)$ for neither of the two subscales contain 0, suggesting that $\hat{\omega}$ is significantly greater than $\hat{\alpha}$ in both the subscales.

Discussion and Conclusion

Measurement reliability plays an important role in understanding the quality of educational and psychological variables. Alpha, conceived as an “internal consistency” coefficient, is the most widely used reliability coefficient in social science reach. However, the properties of coefficient alpha are not well understood by applied researchers, as previous studies such as Green and Yang (2009, p. 121) pointed out that “the general use of coefficient alpha to assess reliability should be discouraged on a number of grounds”. Similarly, Cho and Kim (2015, p. 207) clarified “six common misconceptions about coefficient alpha:

- (1) Alpha was first developed by Cronbach.
- (2) Alpha equals reliability.
- (3) A high value of alpha is an indication of internal consistency.
- (4) Reliability will always be improved by deleting items using “alpha if item deleted” [an option in SPSS].
- (5) Alpha should be greater than or equal to .7 (or, alternatively, .8).
- (6) Alpha is the best choice among all published reliability coefficients.

Table 2. Data Set 2: Big Five Inventory With Five Subscales ($N = 190$).

Subscale	# of items	Statistic T_{ml}				Rescaled statistic T_{ml}			
		df	T_{ml}	p_v	RMS	T_{ml}	p_v	CFI	RMS
Neuro	8	20	96.555	.000	0.851	90.321	.000	0.855	0.136
Extra	8	20	67.732	.000	0.940	64.354	.000	0.943	0.108
Cons	9	27	69.845	.000	0.915	65.224	.000	0.915	0.087
Open	10	35	186.907	.000	0.717	161.832	.000	0.736	0.138
Agree	9	27	76.503	.000	0.824	65.676	.000	0.825	0.087
Subscale	# of items	Reliability	Estimate (SE_{nm} & SE_{sw})		$[L_{nm}, U_{nm}]$		$[L_{sw}, U_{sw}]$		
Neuro	8	ω	0.845 (0.017 & 0.018)		[0.811, 0.878]		[0.810, 0.880]		
		α	0.841 (0.017 & 0.018)		[0.807, 0.875]		[0.805, 0.876]		
Extra	8	$\omega^{-\alpha}$	0.004 (0.002 & 0.001)		[0.001, 0.007]		[0.001, 0.007]		
		ω	0.900 (0.011 & 0.010)		[0.879, 0.922]		[0.880, 0.920]		
Cons	9	α	0.895 (0.011 & 0.010)		[0.873, 0.917]		[0.875, 0.916]		
		$\omega^{-\alpha}$	0.005 (0.002 & 0.002)		[0.001, 0.008]		[0.001, 0.008]		
Open	10	ω	0.818 (0.020 & 0.020)		[0.779, 0.857]		[0.778, 0.858]		
		α	0.824 (0.019 & 0.019)		[0.787, 0.861]		[0.787, 0.861]		
Agree	9	$\omega^{-\alpha}$	-0.006 (0.003 & 0.002)		[-0.011, -0.001]		[-0.011, -0.001]		
		ω	0.782 (0.024 & 0.025)		[0.736, 0.828]		[0.733, 0.831]		
Agree	9	α	0.787 (0.023 & 0.023)		[0.742, 0.832]		[0.742, 0.832]		
		$\omega^{-\alpha}$	-0.005 (0.004 & 0.004)		[-0.013, 0.002]		[-0.013, 0.003]		
Agree	9	ω	0.710 (0.031 & 0.034)		[0.648, 0.771]		[0.643, 0.776]		
		α	0.697 (0.033 & 0.035)		[0.632, 0.761]		[0.627, 0.766]		
Agree	9	$\omega^{-\alpha}$	0.013 (0.005 & 0.005)		[0.003, 0.023]		[0.004, 0.022]		

Note. p_v = p value; RMS = RMSEA; nm = based on normal distribution assumption; sw = sandwich-type covariance matrix.

Table 3. Data Set 3: Humor Styles Questionnaire With Four Subscales (N = 993).

Subscale	# of items	Statistic T_{rml}				Rescaled statistic T_{rml}				
		df	T_{rml}	p_v	CFI	RMS	T_{rml}	p_v	CFI	RMS
affili	8	20	235.405	.000	0.919	0.104	152.130	.000	0.918	0.082
self-enha	8	20	246.524	.000	0.904	0.107	201.237	.000	0.912	0.096
aggress	8	20	134.529	.000	0.932	0.076	110.052	.000	0.940	0.067
self-defe	8	20	139.865	.000	0.949	0.078	115.775	.000	0.953	0.069
Subscale	# of items	Reliability				Estimate (SE_{nm} & SE_{sw})				
Affili	8	ω	0.841 (0.008 & 0.009)	[0.825, 0.856]	[0.823, 0.858]					
		α	0.839 (0.008 & 0.009)	[0.824, 0.854]	[0.822, 0.857]					
		$\omega - \alpha$	0.001 (0.001 & 0.001)	[0.000, 0.002]	[-0.000, 0.002]					
Self-enha	8	ω	0.828 (0.008 & 0.009)	[0.812, 0.844]	[0.812, 0.845]					
		α	0.822 (0.008 & 0.009)	[0.805, 0.838]	[0.805, 0.839]					
		$\omega - \alpha$	0.007 (0.001 & 0.001)	[0.004, 0.009]	[0.004, 0.009]					
Aggress	8	ω	0.793 (0.010 & 0.010)	[0.773, 0.812]	[0.772, 0.813]					
		α	0.790 (0.010 & 0.011)	[0.770, 0.810]	[0.769, 0.811]					
		$\omega - \alpha$	0.003 (0.001 & 0.001)	[0.002, 0.004]	[0.002, 0.004]					
Self-defe	8	ω	0.823 (0.008 & 0.009)	[0.807, 0.840]	[0.807, 0.840]					
		α	0.819 (0.009 & 0.009)	[0.802, 0.836]	[0.802, 0.837]					
		$\omega - \alpha$	0.004 (0.001 & 0.001)	[0.002, 0.006]	[0.002, 0.006]					

Note. p_v = p value; RMS = RMSEA; nm = based on normal distribution assumption; sw = sandwich-type covariance matrix.

Table 4. Data Set 4: Family Functioning Scale With Two Subscales ($N = 852$).

Subscale	# of items	df	Statistic T_{ml}				Rescaled statistic T_{rml}			
			T_{ml}	p_v	CFI	RMS	T_{rml}	p_v	CFI	RMS
cohesion	16	104	539.231	.000	0.888	0.070	402.858	.000	0.887	0.058
	13	65	243.217	.000	0.928	0.057	198.555	.000	0.925	0.049
Subscale	# of items		Reliability	Estimate (SE_{nm} & SE_{sw})			$[L_{nm}$ $U_{nm}]$		$[L_{sw}$ $U_{sw}]$	
Cohesion	16		ω	0.866 (0.007 & 0.008)		$[0.853, 0.879]$		$[0.850, 0.882]$		
			α	0.864 (0.007 & 0.008)		$[0.851, 0.877]$		$[0.848, 0.879]$		
			$\omega - \alpha$	0.002 (0.001 & 0.001)		$[0.001, 0.003]$		$[0.001, 0.003]$		
Adapt	13		ω	0.805 (0.010 & 0.011)		$[0.786, 0.825]$		$[0.785, 0.826]$		
			α	0.799 (0.010 & 0.011)		$[0.779, 0.818]$		$[0.777, 0.820]$		
			$\omega - \alpha$	0.007 (0.001 & 0.001)		$[0.004, 0.009]$		$[0.004, 0.010]$		

Note. p_v = p value; RMS = RMSEA; nm = based on normal distribution assumption; sw = sandwich-type covariance matrix.

More and more researchers suggest that alpha is not the best choice within current research practice and advocate the switch from alpha to omega, especially when the tau-equivalent assumption is violated. However, the majority of applied researches still tend to choose alpha because they are more familiar with alpha than omega, and also because the difference between alpha and omega was believed to be small. So there is an increasing need to develop a convincing method to compare the difference of coefficient alpha and omega statistically, which will offer an updated perspective for the ongoing debate on the issue.

In this article, we developed a methodology for estimating the *SE* of $(\hat{\omega} - \hat{\alpha})$ and consequently the CI for $(\omega - \alpha)$. We further applied the method to four real data sets from well-known scales, and the results indicated that in most of the cases $\hat{\omega}$ and $\hat{\alpha}$ are significantly different (13 of 16 scales, about 81.25%). These suggest that significant differences do exist between coefficient alpha and omega and, to some extent, support that “substituting alpha with a superior alternative is not merely a matter of personal choice but a matter of academia consciously responding to the issue” (Cho & Kim, 2015, p. 225). However, cases also exist in which the difference between coefficients alpha and omega is not significant (3 of 16 subscales, about 18.75%).

This study also evaluated unidimensionality properties of items on 16 subscales. Most scales are fitted reasonably well by the one-factor model. However, the fit indices CFI and RMSEA do not always agree with the goodness of fit. This has been pointed out by Kim and Markland (Kim, 2005; Markland, 2005), and our analysis reconfirmed it. The main purpose of this study is to see how often the differences between the two reliability coefficients is significant. Our results indicate that there is no apparent association between the difference of the two reliability coefficients and the unidimensionality properties of the items although omega is calculated on the estimation of the one-factor model. There are still significant difference between $\hat{\omega}$ and $\hat{\alpha}$ regardless whether the items were not fitted well by the one-factor model.

This article provides a scientific method and an R program for computing the *SE* of $(\hat{\omega} - \hat{\alpha})$ and consequently the CI for $(\omega - \alpha)$. The development offers a sound procedure for applied researchers to compare the difference between alpha and omega. The results from the analysis of the well-known scales may also offer certain solid evidence for researchers who might consider a shift from alpha to omega in future. In addition, the development in this article is a necessary supplement to that developed in Raykov and Marcoulides (2015) when the average loading is below .7 or when differences between certain individual factor loadings and the average loading are greater than .2. Four conditions (a) to (d) were noted earlier in this article when discussing the relationship between alpha, omega, and the true reliability (see also Raykov & Marcoulides, 2015). These conditions might be hard to verify in practice. As were seen earlier in this article, in 13 of 16 scales $\hat{\omega}$ and $\hat{\alpha}$ are significantly different, which may also imply that some of the conditions in (a) to (d) are violated. In particular, we are literally unable to conclude that items are unidimensional even if a test statistic is not significant. What we can conclude is that there is not enough evidence to reject the one-factor model.

Following McDonald (1999), the coefficient omega in this article is defined on the estimation of a one-factor model, although the items may not be unidimensional. Alternatively, we can fit the items on each subscale by a multifactor model or by including correlated errors (e.g., Bentler, 2007; Yang & Green, 2010). However, it might be difficult to label the factors of each subscale within well-known or well-developed instruments. Such a difficulty may also pass onto the interpretation of the resulting reliability estimates. More studies in such a direction might be needed in order to understand the difference between reliability estimates based on one and multiple factor models.

In this article, the formula for the SE of $(\hat{\omega} - \hat{\alpha})$ is obtained by asymptotics, via the sandwich-type covariance matrix of the parameter estimates of factor loadings and error variances. Alternatively, we may also obtain an estimate of the SE by the bootstrap methodology, parallel to the development in Chan (2009) and Raykov and Marcoulides (2015). Under the condition of identically distributed or exchangeable observations, results in Yuan and Hayashi (2006) indicate that the two methods yield essentially the same results even at small sample sizes. But the bootstrap method may fail when the observations are not exchangeable or not identically distributed (Wu, 1986). In contrast, SE s based on sandwich-type covariance matrices are still consistent as long as the observations are independent (White, 1980). Results in Jones and Waller (2013) indicate that confidence intervals based on asymptotics can be more accurate than those based on the bootstrap. Also, with bootstrap different people will get different results due to different bootstrap replications, or starting seed, such a difference can be confounded with the difference between sample alpha and omega. So the development of the analytical formula for the SE of $(\hat{\omega} - \hat{\alpha})$ in this article not only allows a more reliable assessment of the difference between alpha and omega but also represents an advancement in assessment methodology.

Appendix A

This appendix contains the details leading to the asymptotic distribution of $\hat{\omega} - \hat{\alpha}$ and consequently the formula of confidence interval for $\omega - \alpha$. For such a purpose, we will need to obtain the asymptotic expansions of $\hat{\alpha}$ and $\hat{\omega}$ as functions of $\mathbf{s} = \text{vech}(\mathbf{S})$, respectively. Deng, Marcoulides, and Yuan (2015) gave a procedure for obtaining the standard error of the difference between two alphas or two omegas of correlated samples, but not for the difference between the estimates of omega and alpha computed using the same sample. Thus, the development here is parallel to that in Deng et al. (2015) but in a different direction.

Notice that $\mathbf{\Sigma}$ is a symmetric matrix with duplicated elements, and let $\boldsymbol{\sigma} = \text{vech}(\mathbf{\Sigma})$ be a $p^* = p(p+1)/2$ -dimensional vector that contains the low-triangular part of $\mathbf{\Sigma}$. Let \mathbf{a} be a p^* -dimensional vector whose elements are 1.0 corresponding to the position of σ_{jj} ($j = 1, 2, \dots, p$) in $\boldsymbol{\sigma}$ and 0 elsewhere; and \mathbf{b} be also a p^* -dimensional vector whose elements are 1.0 corresponding to the position of σ_{jj} ($j = 1, 2, \dots, p$) in $\boldsymbol{\sigma}$ and 2.0 elsewhere. Then corresponding to Equation (1) we can rewrite α as

$$\alpha = \alpha(\boldsymbol{\sigma}) = \frac{p}{p-1} \left(\mathbf{1} - \frac{\mathbf{a}'\boldsymbol{\sigma}}{\mathbf{b}'\boldsymbol{\sigma}} \right).$$

Using standard calculus, the differential of α is given by

$$d\alpha = -\frac{p}{p-1} \left[\frac{\mathbf{a}'d\boldsymbol{\sigma}}{\mathbf{b}'\boldsymbol{\sigma}} - \frac{\mathbf{a}'\boldsymbol{\sigma}}{(\mathbf{b}'\boldsymbol{\sigma})^2} (\mathbf{b}'d\boldsymbol{\sigma}) \right].$$

Consequently, the derivative of $\alpha(\boldsymbol{\sigma})$ is

$$\dot{\alpha}(\boldsymbol{\sigma}) = \frac{d\alpha(\boldsymbol{\sigma})}{d\boldsymbol{\sigma}} = \frac{p}{p-1} \left[\frac{\mathbf{a}'\boldsymbol{\sigma}}{(\mathbf{b}'\boldsymbol{\sigma})^2} \mathbf{b} - \frac{\mathbf{1}}{\mathbf{b}'\boldsymbol{\sigma}} \mathbf{a} \right].$$

It follows from the mean value theorem that

$$\sqrt{n}(\widehat{\alpha} - \alpha) = \sqrt{n}[\alpha(\mathbf{s}) - \alpha(\boldsymbol{\sigma})] = \dot{\alpha}'(\overline{\boldsymbol{\sigma}}) \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma}) = \dot{\alpha}'(\boldsymbol{\sigma}) \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma}) + o_p(1), \quad (\text{A1})$$

where $\overline{\boldsymbol{\sigma}}$ is a p^* -dimensional vector whose elements are between \mathbf{s} and $\boldsymbol{\sigma}$; and $o_p(1)$ denotes a term that approaches 0 in probability as n increases.

Similar to (A1), we need to obtain a formula for $\widehat{\omega}$ to be approximated by a linear combination of the elements of \mathbf{s} . As noted in the main body of the article, $\widehat{\omega}$ in Equation (2) is a function of $\widehat{\boldsymbol{\theta}}$, and $\widehat{\boldsymbol{\theta}}$ is a function of \mathbf{s} . It follows from the definition of ω , parallel to that of Equation (2), that its differential is given by

$$\begin{aligned} d\omega &= \frac{2(\mathbf{1}'\boldsymbol{\lambda})(\mathbf{1}'d\boldsymbol{\lambda})}{(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi})} - \frac{(\mathbf{1}'\boldsymbol{\lambda})^2 [2(\mathbf{1}'\boldsymbol{\lambda})(\mathbf{1}'d\boldsymbol{\lambda}) + \text{tr}(d\boldsymbol{\Psi})]}{[(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi})]^2} \\ &= \frac{2[(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi})](\mathbf{1}'\boldsymbol{\lambda})(\mathbf{1}'d\boldsymbol{\lambda}) - (\mathbf{1}'\boldsymbol{\lambda})^2 [2(\mathbf{1}'\boldsymbol{\lambda})(\mathbf{1}'d\boldsymbol{\lambda}) + \text{tr}(d\boldsymbol{\Psi})]}{[(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi})]^2} \\ &= \frac{2\text{tr}(\boldsymbol{\Psi})(\mathbf{1}'\boldsymbol{\lambda})(\mathbf{1}'d\boldsymbol{\lambda}) - (\mathbf{1}'\boldsymbol{\lambda})^2 \text{tr}(d\boldsymbol{\Psi})}{[(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi})]^2}, \end{aligned}$$

where $\mathbf{1}$ is a vector of p 1s. Thus,

$$\frac{\partial \omega}{\partial \boldsymbol{\lambda}} = \frac{2\text{tr}(\boldsymbol{\Psi})(\mathbf{1}'\boldsymbol{\lambda})}{[(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi})]^2} \mathbf{1} \quad \text{and} \quad \frac{\partial \omega}{\partial \boldsymbol{\Psi}} = -\frac{\omega^2}{(\mathbf{1}'\boldsymbol{\lambda})^2} \mathbf{1}.$$

Notice that

$$\omega(1 - \omega) = \frac{(\mathbf{1}'\boldsymbol{\lambda})^2 \text{tr}(\boldsymbol{\Psi})}{\left[(\mathbf{1}'\boldsymbol{\lambda})^2 + \text{tr}(\boldsymbol{\Psi}) \right]^2}.$$

We have

$$\dot{\omega}(\boldsymbol{\theta}) = \left(2 \frac{\omega(1 - \omega)}{(\mathbf{1}'\boldsymbol{\lambda})} \mathbf{1}', -\frac{\omega^2}{(\mathbf{1}'\boldsymbol{\lambda})^2} \mathbf{1}' \right).$$

It follows from the mean value theorem that

$$\sqrt{n}(\hat{\omega} - \omega) = \sqrt{n} \left[\omega(\hat{\boldsymbol{\theta}}) - \omega(\boldsymbol{\theta}) \right] = \dot{\omega}'(\bar{\boldsymbol{\theta}}) \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \dot{\omega}'(\boldsymbol{\theta}) \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + o_p(1), \tag{A2}$$

where $\hat{\boldsymbol{\theta}}$ is a p^* -dimensional vector whose elements are between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$.

We still need to relate $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$ to \mathbf{s} and $\boldsymbol{\sigma}$, respectively. Let $\dot{\boldsymbol{\sigma}}$ be the matrix of derivatives of $\boldsymbol{\sigma}$ with respect to $\boldsymbol{\theta}'$, and $\mathbf{W} = \mathbf{D}'_p(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\mathbf{D}_p/2$, where \mathbf{D}_p is the duplication matrix and \otimes is the notation for Kronecker product (Schott, 2005). Then it follows from Equation (6) of Yuan et al. (2003) that

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \mathbf{P}\sqrt{n}(\mathbf{s} - \boldsymbol{\sigma}) + o_p(1), \tag{A3}$$

where $\mathbf{P} = (\dot{\boldsymbol{\sigma}}'\mathbf{W}\dot{\boldsymbol{\sigma}})^{-1}\dot{\boldsymbol{\sigma}}'\mathbf{W}$. Substituting the $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ in (A2) by (A3) results in

$$\sqrt{n}(\hat{\omega} - \omega) = \dot{\omega}'\mathbf{P}\sqrt{n}(\mathbf{s} - \boldsymbol{\sigma}) + o_p(1). \tag{A4}$$

Combining (A1) and (A4) yields

$$\sqrt{n}[(\hat{\omega} - \hat{\alpha}) - (\omega - \alpha)] = (\dot{\omega}'\mathbf{P} - \dot{\alpha}')\sqrt{n}(\mathbf{s} - \boldsymbol{\sigma}) + o_p(1), \tag{A5}$$

where $\dot{\omega}$, $\dot{\boldsymbol{\sigma}}$, and \mathbf{W} are functions of $\boldsymbol{\theta}$; and $\dot{\alpha}$ is a function of $\boldsymbol{\sigma}$. Let $\boldsymbol{\Gamma} = \text{Cov}(\sqrt{n}\mathbf{s})$. It follows from (A5) that

$$\sqrt{n}[(\hat{\omega} - \hat{\alpha}) - (\omega - \alpha)] \xrightarrow{L} N(0, \tau^2), \tag{A6}$$

Where

$$\tau^2 = (\dot{\omega}'\mathbf{P} - \dot{\alpha}')\boldsymbol{\Gamma}(\mathbf{P}'\dot{\omega} - \dot{\alpha}).$$

A consistent estimate $\hat{\tau}^2$ of τ^2 will be obtained when $\dot{\omega}$, \mathbf{P} , $\dot{\alpha}$, and $\boldsymbol{\Gamma}$ are replaced by $\dot{\omega}(\hat{\boldsymbol{\theta}})$, $\hat{\mathbf{P}} = [\dot{\boldsymbol{\sigma}}'(\hat{\boldsymbol{\theta}})\mathbf{W}(\hat{\boldsymbol{\theta}})\dot{\boldsymbol{\sigma}}(\hat{\boldsymbol{\theta}})]^{-1}\dot{\boldsymbol{\sigma}}'(\hat{\boldsymbol{\theta}})\mathbf{W}(\hat{\boldsymbol{\theta}})$, $\dot{\alpha}(\mathbf{s})$, and $\hat{\boldsymbol{\Gamma}} = \mathbf{S}_y$, where \mathbf{S}_y is the sample covariance matrix \mathbf{S} available without raw data, we may estimate $\boldsymbol{\Gamma}$ by $\hat{\boldsymbol{\Gamma}}_1 = 2\mathbf{D}_p^+(\mathbf{S} \otimes \mathbf{S})\mathbf{D}_p^{+'}$ or $\hat{\boldsymbol{\Gamma}}_2 = 2\mathbf{D}_p^+(\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) \otimes \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}))\mathbf{D}_p^{+'}$, where $\mathbf{D}_p^+ = (\mathbf{D}'_p \mathbf{D}_p)^{-1}\mathbf{D}'_p$. Then we still get a consistent $\hat{\tau}^2$ when data are normally distributed or when in addition the one-factor model can be regarded as correctly specified.

It follows from (A6) that a confidence interval for $\omega - \alpha$ with Level $1 - 2\beta$ can be obtained as

$$[\hat{\omega} - \hat{\alpha} - z_{\beta}\hat{\tau}/\sqrt{n}, \hat{\omega} - \hat{\alpha} + z_{\beta}\hat{\tau}/\sqrt{n}],$$

where z_{β} is the $(1 - \beta)$ critical value under the standard normal distribution. For example, $z_{\beta} = 1.96$ with $\beta = 0.25$, corresponding to an interval with confidence level .95.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was partially supported by a grant from the Department of Psychology, The Chinese University of Hong Kong.

Notes

1. There is a formula for calculating coefficient omega with more than one latent dimension (see, e.g., Yang & Green, 2010). However, there might be operational difficulty in practice because cross-loadings and correlated errors are typically confounded.
2. Note that RMSEA and CFI under the rescaled statistic are computed by simply replacing T_{ml} by T_{rml} under the substantive model and the base model, respectively.

References

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks-Cole.
- Bentler, P. M. (1972). A lower-bound method for the dimension-free measurement of internal consistency. *Social Science Research, 1*, 343-357.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.
- Bentler, P. M. (2007). Covariance structure models for maximal reliability of unit-weighted composites. In S.-Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 1-19). Amsterdam, Netherlands: North-Holland.
- Bentler, P. M., & Dijkstra, T. K. (1985). Efficient estimation via linearization in structural models. In P. R. Krishnaiah (Ed.), *Multivariate analysis VI* (pp. 9-42). Amsterdam, Netherlands: North-Holland.
- Bentler, P. M., & Woodward, J. A. (1980). Inequalities in long lower bounds to reliability: With applications to test construction and factor analysis. *Psychometrika, 45*, 249-267.
- Chan, W. (2009). Bootstrap standard error and confidence intervals for the difference between two squared multiple correlation coefficients. *Educational and Psychological Measurement, 69*, 566-584.

- Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha well known but poorly understood. *Organizational Research Methods, 18*, 207-230.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Deng, L., Marcoulides, G., & Yuan, K.-H. (2015). Psychometric properties of measures of team diversity with Likert data. *Educational and Psychological Measurement, 75*, 512-534.
- Deng, L., Wang, L., & Zhao, Y. (2016). How creativity was affected by environmental factors and Individual characteristics: A cross-cultural comparison perspective. *Creativity Research Journal, 28*, 357-366.
- Deng, L., Zheng, R. C. (2012). Relationships among family functioning, emotional expression and loneliness in college students. *Studies of Psychology and Behavior, 11*, 223-238.
- Dunn, T. J., Baguley, T., & Brunnsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*, 399-412.
- Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*, 827-837.
- Green, S. B., & Yang, Y. (2009). Commentary on coefficient alpha: A cautionary tale. *Psychometrika, 74*, 169-173. doi:10.1007/s11336-008-9098-4
- Holzinger, K. J., & Swineford, F. (1939). *A study in factor analysis: The stability of a bi-factor solution* (Supplementary Educational Monographs, No. 48). Chicago, IL: University of Chicago.
- Hunt, T. D., & Bentler, P. M. (2015). Quantile lower bounds to reliability based on splits. *Psychometrika, 80*, 182-195.
- John, O. P., & Srivastava, S. (1999). The big five-trait taxonomy: History, measurement, and theoretical perspectives. In J. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 102-138). New York, NY: Guilford Press.
- Jones, J. A., & Waller, N. G. (2013). Computing confidence intervals for standardized regression coefficients. *Psychological Methods, 18*, 435-453. doi:10.1037/a0033269
- Jóreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 36*, 109-133.
- Kim, K. H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling, 12*, 368-390.
- Li, H., Rosenthal, R., & Rubin, D. B. (1996). Reliability of measurement in psychology: From Spearman-Brown to maximal reliability. *Psychological Methods, 1*, 98-107.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Maydeu-Olivares, A., Coffman, D. L., & Hartmann, W. M. (2007). Asymptotically distribution free (ADF) interval estimation of coefficient alpha. *Psychological Methods, 12*, 157-176.
- Markland, D. (2005). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modeling. *Personality and Individual Differences, 42*, 851-858.
- Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality, 37*, 48-75.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- Olson, D. H., Portner, J., & Bell, R. Q. (1982). *FACES II: Family Adaptability and Cohesion Evaluation Scales*. Minneapolis, MN: Family Social Science, University of Minnesota.

- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research, 32*, 329-353.
- Raykov, T. (2012). Scale construction and development using structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 472-492). New York, NY: Guilford Press.
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York, NY: Taylor & Francis.
- Raykov, T., & Marcoulides, G. A. (2015). A direct latent variable modeling based method for point and interval estimation of coefficient alpha. *Educational and Psychological Measurement, 75*, 146-156.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Newbury Park, CA: Sage.
- Schott, J. (2005). *Matrix analysis for statistics* (2nd ed.). New York, NY: Wiley.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*, 107-120.
- Steiger, J. H., & Lind, J. M. (1980, July). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Thompson, B. (Ed.). (2003). *Score reliability: Contemporary thinking on reliability*. Thousand Oaks, CA: Sage.
- Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics, 14*, 1261-1295.
- White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica, 48*, 817-838.
- Yang, Y., & Green, S. B. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling, 17*, 66-81.
- Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment, 29*, 377-392.
- Yuan, K.-H., & Bentler, P. M. (2002). On robustness of the normal-theory based asymptotic distributions of three reliability coefficient estimates. *Psychometrika, 67*, 251-259.
- Yuan, K.-H., Guarnaccia, C. A., & Hayslip, B. (2003). A study of the distribution of sample coefficient with the Hopkins symptom checklist: Bootstrap versus asymptotic. *Educational and Psychological Measurement, 63*, 5-23.
- Yuan, K.-H., & Hayashi, K. (2006). Standard errors in covariance structure models: Asymptotics versus bootstrap. *British Journal of Mathematical and Statistical Psychology, 59*, 397-417.
- Yuan, K.-H., Marshall, L. L., & Bentler, P. M. (2003). Assessing the effect of model misspecifications on parameter estimates in structural equation models. *Sociological Methodology, 33*, 241-265.
- Zhang, Z. Y., & Yuan, K.-H. (2016). Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data: methods and software. *Educational and Psychological Measurement, 76*, 387-411.
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α , Revelle's β , and McDonald's ω H: Their relations with each other and two alternate conceptualizations of reliability. *Psychometrika, 70*, 1-11.