

Making Reliability Reliable: A Systematic Approach to Reliability Coefficients

Organizational Research Methods
1-32

© The Author(s) 2016

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1094428116656239

orm.sagepub.com



Eunseong Cho¹

Abstract

The current conventions for test score reliability coefficients are unsystematic and chaotic. Reliability coefficients have long been denoted using names that are unrelated to each other, with each formula being generated through different methods, and they have been represented inconsistently. Such inconsistency prevents organizational researchers from understanding the whole picture and misleads them into using coefficient alpha unconditionally. This study provides a systematic naming convention, formula-generating methods, and methods of representing each of the reliability coefficients. This study offers an easy-to-use solution to the issue of choosing between coefficient alpha and composite reliability. This study introduces a calculator that enables its users to obtain the values of various multidimensional reliability coefficients with a few mouse clicks. This study also presents illustrative numerical examples to provide a better understanding of the characteristics and computations of reliability coefficients.

Keywords

reliability, coefficient alpha, structural equation modeling, tau-equivalency, hierarchical omega

Coefficient alpha (hereinafter alpha) is the most commonly used single-administration test score reliability coefficient (hereinafter reliability coefficient). Whereas previous studies such as Cortina (1993) and Schmitt (1996) offered influential lessons on alpha for organizational researchers, it is still commonly misconceived and widely misused (Cho & Kim, 2015; Dunn, Baguley, & Brunsten, 2013; Green & Yang, 2009a; Raykov, 2012; Sijtsma, 2009a; Yang & Green, 2011). Another study that focuses only on alpha is not likely to resolve the chronic misconceptions and misapplications. A better approach to ascertaining the characteristics of alpha is considering other reliability coefficients together with alpha. Once we know the commonalities and differences between alpha and other reliability coefficients, we can naturally discern the conditions under which it should or should

¹College of Business Administration, Kwangwoon University, Seoul, Republic of Korea

Corresponding Author:

Eunseong Cho, College of Business Administration, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul, 139-701, Republic of Korea.

Email: bene@kw.ac.kr

not be used. However, there is an obstacle that prevents organizational researchers from looking at the big picture.

The current conventions for reliability coefficients are haphazard and undisciplined. Our knowledge of reliability was not built in a day by a genius. For more than a hundred years, numerous researchers (e.g., Brown, 1910; Spearman, 1910) have developed reliability coefficients in various ways, but during this process, they were named, interpreted and expressed inconsistently. The present of reliability coefficients is locked in the past (i.e., path dependence). The way reliability coefficients are currently being named, computed, and used lacks *reliability*, which makes it difficult for new users of reliability to determine the whole picture.

This study attempts to improve the *reliability* of reliability coefficients. I describe my approach as *systematic* because I propose a system composed of reliability coefficients and that can conditionally suggest appropriate reliability coefficients depending on the characteristics of the data. The *system* includes most of the reliability coefficients commonly used in real-world data analyses or explained in research methods textbooks such as the Spearman–Brown formula, the Flanagan–Rulon formula, standardized alpha, alpha, stratified alpha, McDonald’s omega, and so-called composite reliability.

This study proves that various reliability coefficients are generated from measurement models nested within the bifactor measurement model. The idea of estimating reliability based on a measurement model in the large framework of structural equation modeling (SEM) is not new. For example, Miller (1995) employed an SEM path diagram to explain the meaning of alpha and its correct use. McDonald (1985, 1999) and Zinbarg, Revelle, and Yovel (2007) argued that alpha is a special case of omega when the data meet a certain prerequisite. This study is an extension of such previous studies, and it offers a more comprehensive analysis on a number of reliability coefficients and their algebraically equivalent variations instead of concentrating on one or two.

This study consists of three sections. The “problems” section examines the current practice and declares that alpha is ill positioned as a representative of reliability coefficients. The “a systematic approach” section claims that a successful repositioning of reliability coefficients may be based on their renaming and formula re-expressions. The “examples” section offers various computation examples and introduces a gadget that calculates various reliability coefficients with a few mouse clicks.

Problems in Current Practice

This study begins by determining what the problem is and assessing how widespread it is.¹ In previous studies criticizing alpha’s misuse, such misuse has typically consisted of one or both of the following types:

1. Alpha is most frequently used even though it is not the most accurate reliability coefficient (i.e., it is overused)
2. Alpha’s use is unqualified if its assumptions such as tau-equivalency are not examined (i.e., it is incorrectly used)

However, little research has empirically examined the fundamental premise that alpha is overused and/or incorrectly used in practice. One may raise a counterargument that alpha is not as severely misused as the existing literature suggests. For example, a reasonable argument is that although alpha was overused in the past, organizational researchers are increasingly switching from alpha to other reliability coefficients based on the influence of recent methodological studies that discourage the use of alpha. Another plausible expectation is that although articles appearing in less prestigious journals may use alpha without examining its assumptions, high-impact journal articles demonstrate exemplary use of reliability coefficients because the reviewers and editors demand

Table 1. The Current Use of Reliability Coefficients in Two Top Journals.

		AMJ (n = 42)		JAP (n = 96)	
		n	%	n	%
Used reliability coefficients	Alpha only	35	83.3	84	87.5
	CR only	1	2.4	1	1.0
	Alpha and CR	2	4.8	2	2.1
	Not identified	4	9.5	9	9.4
Tau-equivalency	Examined	0	0.0	0	0.0
	Not examined	42	100.0	96	100.0
Use of CFA/SEM	Reported	27	64.3	44	45.8
	Not reported	15	35.7	52	54.2
Name of alpha	Cronbach's alpha	13	35.1	30	34.9
	Coefficient alpha	7	18.9	15	17.4
	α	10	27.0	31	36.0
	Others	7	18.9	10	11.6

Note: AMJ is the *Academy of Management Journal*, JAP is the *Journal of Applied Psychology*, CR is composite reliability, CFA is confirmatory factor analysis, and SEM is structural equation modeling. The percentages in the "name of alpha" row are the ratios of each cell to the number of those studies that reported the use of alpha, and the percentages in other rows are the ratio of each cell to the total number of examined studies of each journal. Names such as internal consistency reliability, internal consistency, and reliability were classified as "not identified." Names such as Cronbach alpha, Cronbach's alpha reliability, and Cronbach's α were classified as "Cronbach's alpha." Names such as alpha, alpha reliability, and alpha coefficients were classified as "others."

higher standards of methodological rigor. To evaluate how seriously alpha is misused in organizational research, this study examined two elite journals, namely, the *Academy of Management Journal* (AMJ) and the *Journal of Applied Psychology* (JAP).

In addition to diagnosing whether alpha is overused and/or incorrectly used, this study addresses two other issues relevant to current practice. First, this study examines what terms are currently used to denote reliability coefficients. This is necessary because the study will later discuss the *unreliability* of reliability coefficients' names. Second, this study considers whether the use of confirmatory factor analysis (CFA) or SEM was reported. Composite reliability, which is based on a unidimensional SEM measurement model, is the second most frequently used reliability coefficient in organizational research (Peterson & Kim, 2013). This study predicts that the use of SEM will have an effect on the choice of reliability coefficient: Those studies that employ SEM are more likely to report composite reliability rather than alpha, and those studies that do not rely on SEM are more likely to use alpha.

Let us explain the method of data collection. After searching all articles except editorials that were published in AMJ and JAP during the years 2013 and 2014, I included empirical studies that reported single-administration test score reliability estimates and excluded studies that used other types of reliability (e.g., interrater reliability) and meta-analysis. The sample consisted of 42 AMJ articles from a total of 145 (29.0%) and 96 JAP articles from a total of 147 (65.3%). When multiple names were used to express the same reliability coefficient, the unabridged or more descriptive ones were recorded. For example, if both Cronbach's alpha and α were used, the former was coded as the name. If both internal consistency reliability and α were used, the latter was coded as the name.

Table 1 shows the results. More than 80% of the studies used alpha. No studies reported the use of any reliability coefficients other than alpha and composite reliability. Approximately 10% of the studies did not clearly indicate what reliability coefficients were used. No studies examined the

assumption of tau-equivalency. A total of 16 versions of alpha's name were recorded. Among them, frequently used names were Cronbach's alpha, coefficient alpha and α .

Alpha's overuse was more serious than anticipated. First, the use of SEM had little effect on the choice of reliability coefficients. Most studies that employed SEM still used alpha instead of SEM-based reliability coefficients such as composite reliability. Second, 1 in 10 studies unexpectedly did not specify the name of the utilized reliability coefficients. A plausible explanation for why researchers report reliability estimates namelessly is that they take the use of alpha for granted and thus feel little need to report a commonplace thing.

Alpha's incorrect use was also more severe than expected. Not a single study examined tau-equivalency. That is, organizational researchers are automatically using alpha without considering whether its assumptions are satisfied. Such unconditional choice probably stems from one or some combination of the following misconceptions:

1. Alpha is a versatile reliability coefficient that is applicable to any type of data
2. Alpha is robust to any violation of its assumptions (i.e., even a serious violation of the assumptions has an insignificant effect on the value of the reliability estimate)
3. A high value of alpha itself verifies that its assumptions are satisfied
4. Identifying whether alpha's assumptions are satisfied is difficult
5. Alternative methods that can be used when its assumptions are violated are difficult to use

This study claims that each of the above statements is incorrect by providing formula derivations (i.e., Misconception 1), counterexamples (i.e., Misconceptions 2 and 3), illustrative examples, and an easy-to-use solution (i.e., Misconceptions 4 and 5). For example, by showing various computation examples, this study demonstrates that alpha can produce estimation errors as large as .14 when it is misapplied to data that violate one of its assumptions. In addition to disproving such misconceptions, understanding what caused them is necessary for finding a fundamental solution.

A brief review of alpha's history is useful for identifying the underlying source of its misuse. Its popularity did not originate from its technical superiority. It became a *de facto* standard even at the existence of a mathematically superior predecessor (λ_2 : Guttman, 1945) for several reasons that seemed important at the time of Cronbach (1951) but that are trivial from a modern view (Cho & Kim, 2015). First, alpha's computation was simpler. Second, Cronbach (1951) positioned alpha as a reliability coefficient, whereas Guttman (1945) described λ_2 as lower-bound estimators of reliability, which was mathematically correct but represented an unpopular description. Third, Cronbach's (1951) proof that alpha equals the average of the reliability values (λ_4 : Guttman, 1945) that are calculated for all possible split-halves positioned it as a *general* reliability coefficient. Although it seems intuitively attractive, the average is not as meaningful as the maximum (Osburn, 2000) or the minimum (Revelle, 1979) of the λ_4 values that are obtained from all possible split-halves (for a modern interpretation, see Hunt & Bentler, 2015). A positive feedback loop and past popularity bred today's situation. Once the habit of unconditionally using alpha was formed, it prospered despite the development of more sophisticated methods.

Alpha's habitual use is a matter not of mathematics but of marketing. Alpha ranked consistently low in previous comparison studies that examined the accuracy of reliability coefficients (Kamata, Turhan, & Darandari, 2003; Osburn, 2000; Revelle & Zinbarg, 2009; Tang & Cui, 2012; van der Ark, van der Palm, & Sijtsma, 2011). What differentiates alpha from other reliability coefficients is that the awareness of the name alpha outdistances that of any other reliability coefficients. In other words, its name is the main cause of its immense use. Analyzing the reason for the phenomenal citation record of Cronbach (1951), Cronbach (1978) echoed this argument by stating that "I am sure the paper is cited mostly because I put a brand name on a common-place coefficient" (p. 263).

His comments capture the essence; alpha is a *brand name*. The reason a researcher automatically chooses alpha without understanding its formula is analogous to the reason why a laundry detergent

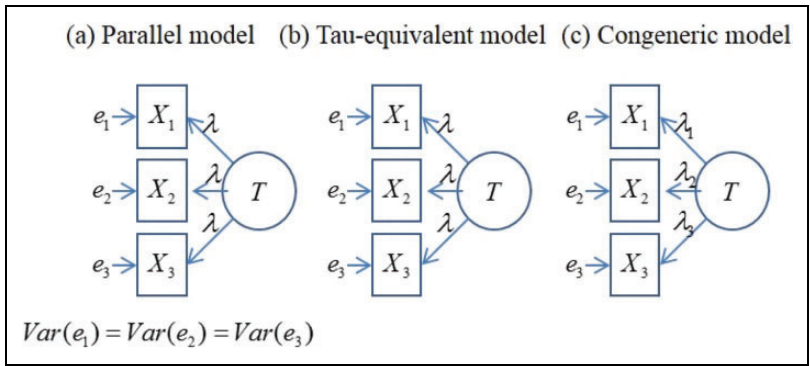


Figure 1. The parallel, tau-equivalent, and congeneric measurement models.

consumer habitually selects a familiar brand not knowing its chemical composition. Even if a comparison study were to reveal that the most popular brand underperforms its competitors, the top-of-mind brand (e.g., Gillette) would not lose significant market share. This phenomenon is what is occurring with alpha. Its use frequency remains unchanged despite the unfavorable results of several performance tests. However, if the brand name becomes null and void for some reason, its sales volume will rapidly decrease. Previous studies that placed sole reliance on a mathematical approach failed to change the 65-year-old habit. A deep-rooted problem requires a more comprehensive solution that includes a radical cure. Rebranding alpha is an efficient way to reposition it into the place where it belongs.

A Systematic Approach to Reliability Coefficients

Measurement Models

Before proceeding with the discussion, let us explain the measurement models used in this study starting from unidimensional models. Figure 1 provides a description of SEM styles regarding unidimensional parallel, tau-equivalent and congeneric measurement models. The term *unidimensional* will be omitted when little possibility of confusion exists. The modifiers strictly and essentially indicate whether item means are constrained to be equal. For example, an essentially tau-equivalent model includes a constant, whereas a strictly tau-equivalent model does not. Although the addition of a constant has an effect on the mean, it does not affect the variances, covariances or the value of reliability. This study focuses on essentially parallel/tau-equivalent/congeneric models and omits the term *essentially* for simplicity. Manifest variables (X_1, X_2, \dots) have a common latent variable (F) and errors (e_1, e_2, \dots). Errors are assumed to be purely random and independent of each other. To determine the scale, the variance of the latent variable is set to a nonzero number (typically 1.0). The congeneric model does not have additional constraints. The tau-equivalent model is the same as the congeneric model, only with the constraint that all the factor loadings are equal. The parallel model is the tau-equivalent model with the constraint that the error variances are all equal.

Table 2 presents interitem covariance matrices of data that perfectly satisfy the conditions of being parallel, tau-equivalent and congeneric. Covariances (i.e., off-diagonal elements of the covariance matrices) between observed item scores are determined only by the common latent variable, whereas variances (i.e., diagonal elements of the covariance matrices) of item scores are determined by the common latent variable and errors. Parallel data have equal interitem covariances and equal item variances. Tau-equivalent data have equal interitem covariances, but they may have different item variances. Congeneric data do not require the equality constraints about variances and covariances. Any parallel data are also tau-equivalent, and any tau-equivalent data are also congeneric.

Table 2. Covariance Matrices of Data That Satisfy the Conditions of Being Parallel, Tau-equivalent, and Congeneric.

	A. Observed Test Score Variance			B. True Score Variance			C. Error Variance					
Parallel reliability		X1	X2	X3		X1	X2	X3		X1	X2	X3
$\hat{\rho}_P = (3^2 \cdot 6) / (3^2 \cdot 6 + 3 \cdot 4) = .8181$	X1	10	6	6	X1	6	6	6	X1	4	0	0
	X2	6	10	6	X2	6	6	6	X2	0	4	0
	X3	6	6	10	X3	6	6	6	X3	0	0	4
Tau-equivalent reliability		X1	X2	X3		X1	X2	X3		X1	X2	X3
$\hat{\rho}_T = (3^2 \cdot 5) / 60 = .75$	X1	8	5	5	X1	5	5	5	X1	3	0	0
	X2	5	10	5	X2	5	5	5	X2	0	5	0
	X3	5	5	12	X3	5	5	5	X3	0	0	7
Congeneric reliability		X1	X2	X3		X1	X2	X3		X1	X2	X3
$\hat{\rho}_C = (2 + 3 + 4)^2 / 100 = .81$	X1	10	6	8	X1	2 ²	3·2	4·2	X1	6	0	0
	X2	6	16	12	X2	2·3	3 ²	4·3	X2	0	7	0
	X3	8	12	22	X3	2·4	3·4	4 ²	X3	0	0	6

Note: If tau-equivalent reliability (e.g., alpha) is misapplied to the above congeneric data, it underestimates the reliability by $.03$ ($\hat{\rho}_T = \hat{\alpha} = (3^2 \cdot ((6 + 8 + 12)/3)) / 100 = .78$).

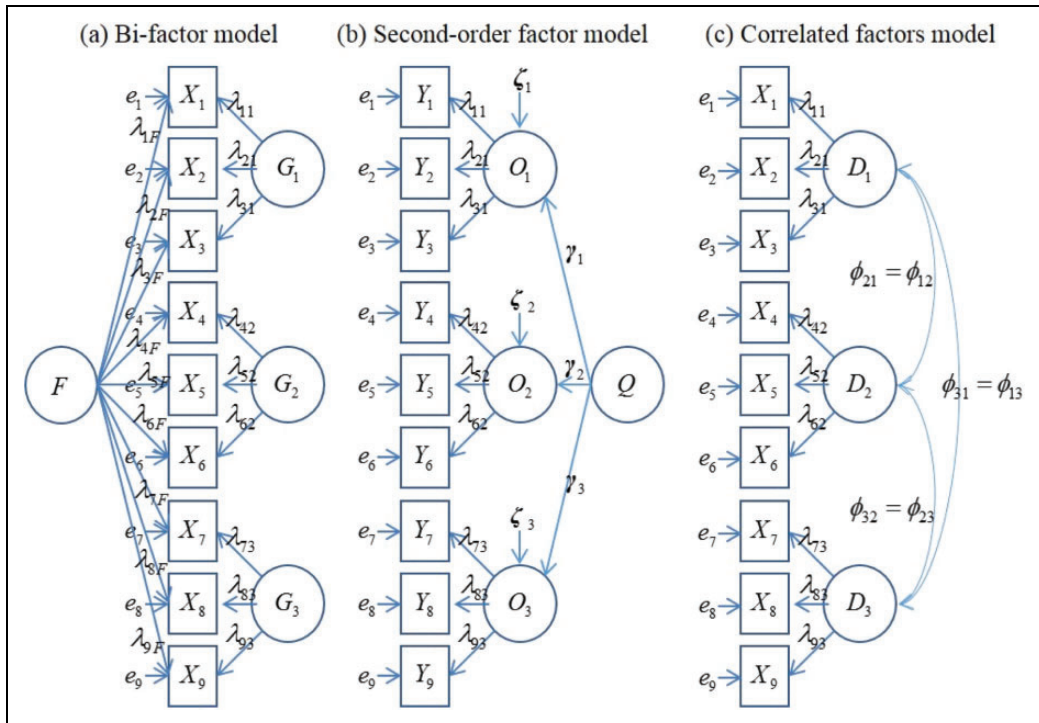


Figure 2. Three major multidimensional measurement models.

Now, let us consider multidimensional models. Three major models conceptualize multidimensionality in SEM (Figure 2): the correlated factor model, the second-order factor model, and the bifactor model. A unidimensional model consists of only a common construct (i.e., *T* in Figure 1)

and omits subtest constructs. A correlated factor model includes only subtest constructs (i.e., D_p in Figure 2) and neglects a common construct. The common construct of a bifactor model is called a general factor (i.e., F), and its subtest constructs are called group factors (i.e., G_p). The common construct in a second-order factor model is called a second-order factor (i.e., Q), and its subtest constructs are called first-order factors (i.e., O_p). To determine the scale, the variances of common constructs and subtest constructs are usually set to 1.0 (i.e., $Var(G_p) = Var(F) = Var(O_p) = Var(Q) = Var(D_p) = 1$ for all p).

The bifactor model is a generalization of the second-order factor model, and the latter is nested within the former. Mathematically speaking, the latter is equivalent to the former only under the proportionality constraint (Yung, Thissen, & McLeod, 1999). Before explaining the proportionality constraint, this study notes that group factors are defined to be independent of a general factor but that first-order factors are dependent on a second-order factor. Disturbances (i.e., ζ_i) are mathematically analogous to group factors because both explain the variances that are not explained by a common construct. The variance due to the disturbance is proportional to the variance due to the second-order factor between manifest variables that have the same first-order factor. For example, let us consider Y_1 and Y_2 . Any effect of the second-order factor (i.e., Q) or the disturbance on Y_1 or Y_2 must be mediated by the coefficients λ_{11} or λ_{21} . Y_1 's ratio of the variance due to the disturbance to the variance due to the second-order factor (i.e., $\lambda_{11}^2 \sigma_{\zeta_1}^2 / \lambda_{11}^2 \gamma_1^2$) is equal to that of Y_2 (i.e., $\lambda_{21}^2 \sigma_{\zeta_1}^2 / \lambda_{21}^2 \gamma_1^2$). The corresponding ratio of X_1 (i.e., $\lambda_{11}^2 / \lambda_{1F}^2$) may differ from that of X_2 (i.e., $\lambda_{21}^2 / \lambda_{2F}^2$) in a bifactor model.

This study offers a direct formula that computes the omega coefficient of a second-order factor model without a transformation. McDonald derived omega from a bifactor model. Applying its formula to a second-order factor model requires a Schmid–Leiman transformation (Schmid & Leiman, 1957) of the parameter estimates (Brunner, Nagy, & Wilhelm, 2012; Yung et al., 1999), which is unfamiliar to typical organizational researchers. A direct formula provides an easier computation and a better understanding of its meaning.

This study introduces multidimensional parallel models and multidimensional tau-equivalent models. The conditions of being parallel and tau-equivalent have only been discussed in unidimensional models in the literature. If such restrictions were so useful in deriving meaningful reliability coefficients (e.g., alpha and standardized alpha) from unidimensional models, they must be equally advantageous to multidimensional models. The second-order factor parallel model (Figure 3) requires four restrictions: The path coefficients between the second-order factor and all first-order factors are restricted to be equal to each other (i.e., $\gamma_p = \gamma$ for all p), the first-order factor loadings of all items are restricted to be equal to each other (i.e., $\lambda_i = \lambda$ for all i), all first-order factors are restricted to have equal numbers of items (i.e., $n_p = n$ for all p), and the errors of all items are restricted to be equal to each other (i.e., $\sigma_{e_i}^2 = \sigma_e^2$ for all i). The second-order factor tau-equivalent model requires three restrictions that are identical to those of the second-order factor parallel model, except for the equality constraint of the error variances (i.e., $\sigma_{e_i}^2 \neq \sigma_{e_j}^2$). Appendix B shows that the reliability coefficients that are derived from bifactor parallel/tau-equivalent models are identical to the reliability coefficients that are derived from second-order factor parallel/tau-equivalent models.

Systematic Nomenclature

Table 3 presents the names of reliability coefficients that are currently used in the literature. The conventional names of reliability coefficients are not user friendly. First, they do not deliver meaningful information to the users. For example, names such as Spearman–Brown or Flanagan–Rulon provide only the names of those who discovered these formulas without expanding on their characteristics. Although respect for these scholars is displayed, the names do not consider the needs of users.

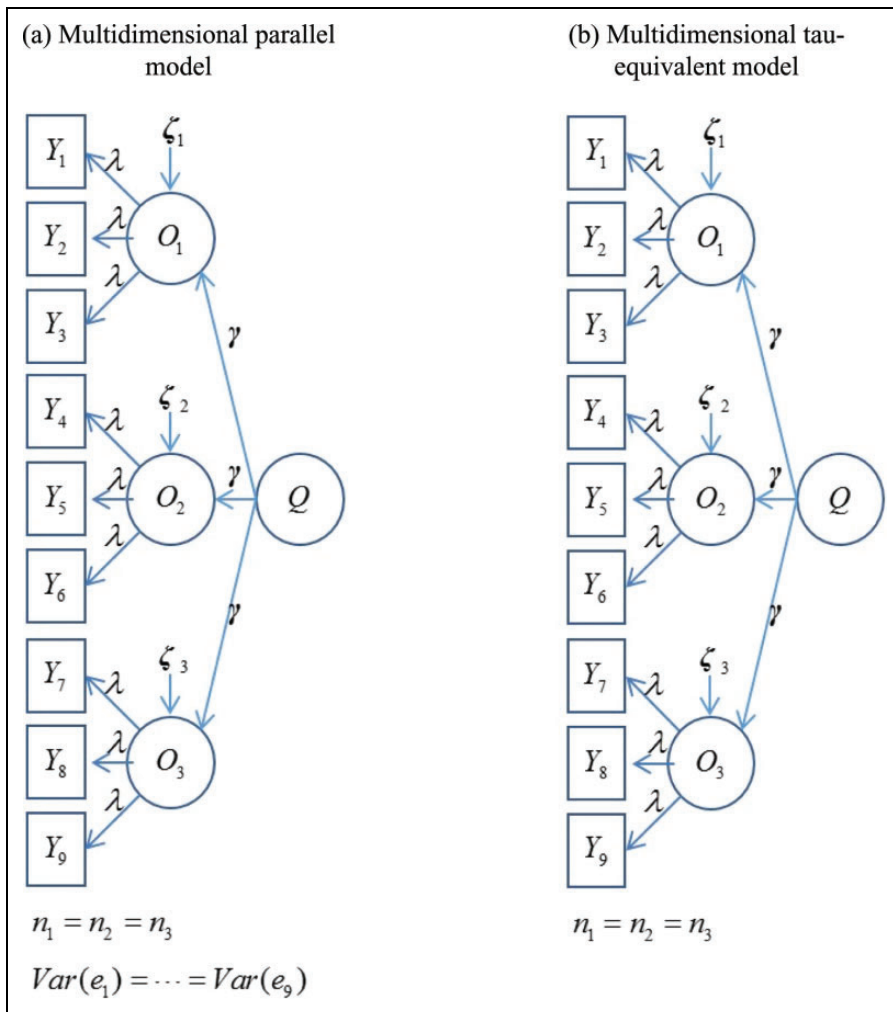


Figure 3. Two suggested multidimensional measurement models.

Second, the names are inconsistent. Some are called formulas (e.g., Spearman–Brown and Flanagan–Rulon), some are called coefficients (e.g., alpha), and others are called lower bounds (e.g., Guttman’s λ_i). Some bear their originators’ names (e.g., Spearman–Brown), and some use a combination of the first and second developers’ names (e.g., Flanagan–Rulon). One estimator goes by the name of the fourth person to propose it (Cronbach, 1951), and others do not bear the name of any developers.

Third, they are not mutually exclusive. Formulas that are algebraically equivalent have different names, such as the Flanagan–Rulon formula and Guttman’s λ_4 . Without background knowledge, users may accept these names as referring to different formulas. On the other hand, one name is used to represent multiple formulas. McDonald (1978) defined omega in a multidimensional context and later used the term regardless of the dimensionality (McDonald, 1999). Previous studies are increasingly using omega as a general name for various SEM-based reliability coefficients (Brunner et al., 2012; Cho & Kim, 2015; Dunn et al., 2013; Green & Yang, 2015; Lucke, 2005; Padilla & Divers, 2013; Revelle & Zinbarg, 2009). Whenever readers encounter the term *omega*, they must understand the context to know exactly which formula was used. Although methodologists may accept such

Table 3. Names of Reliability Coefficients Currently Used in the Literature.

	Unidimensional		Multidimensional
	Split-Half	General	General
Parallel	Spearman–Brown formula	Standardized alpha	(Not yet published)
Tau-equivalent	Flanagan–Rulon formula Flanagan formula Rulon formula Guttman’s λ_4	Cronbach’s alpha Coefficient alpha Guttman’s λ_3 Hoyt method KR-20	Stratified alpha
Congeneric	Raju (1970) coefficient Angoff–Feldt coefficient Angoff coefficient	Composite reliability Construct reliability Congeneric reliability Omega Unidimensional omega Raju (1977) coefficient Classical congeneric reliability coefficient	Omega Omega total McDonald’s omega Multidimensional omega

mixed use as being convenient, it can confuse nonexperts. Raju coefficients also require special attention; Raju (1970, 1977) coefficients require the specification of the years of publication to avoid confusion because the researcher proposed two reliability coefficients.

Fourth, the current nomenclature is not expandable. Not all reliability coefficients have names. Table 3 indicates that a reliability estimator based on a multidimensional parallel model is a theoretically possible reliability coefficient that no one has yet formally proposed, thereby leading to its lack of a name. While generally accepted rules for naming a newly developed reliability coefficient do not exist, a naming method that was popular in the past is choosing one of the Greek letters, for example, α (Cronbach, 1951), β (Revelle, 1979), λ (Guttman, 1945), μ (Ten Berge & Zegers, 1978), θ (Armor, 1974), and ω (McDonald, 1978). This naming strategy is not sustainable because we are running short of Greek letters. Greek letters such as σ , ε , ρ , and τ are almost exclusively used for frequently used methodological notions, and most of the remaining Greek letters (e.g., ϕ , ζ , γ , and ξ) are habitually used in the SEM literature. Using one of them as the name of a reliability coefficient would confuse users.

Table 4 shows the names proposed by this study. A systematic nomenclature should be informative, consistent, mutually exclusive and expandable. It should effectively and economically convey the characteristics of each method as well as their commonalities and differences. A systematic name of a reliability coefficient should also be combined with its use. For example, imagine that Fisher (1925) named the formula of two-way ANOVA (analysis of variance) after himself. We would spend a long time memorizing when to use the Fisher formula. If we rename alpha as tau-equivalent reliability, we do not need to learn the conditions under which it should be used once we know the tau-equivalent measurement model. A systematic nomenclature becomes easier once we become accustomed to it.

This study provides additional cautionary tales on the expressions Cronbach’s alpha and composite reliability because they are the most commonly used but misleading names. First, let us consider *Cronbach’s alpha*. At the time of Cronbach’s (1951) publication, this formula was already “a common-place coefficient” (Cronbach, 1978, p. 263) and usually called KR-20. Kuder and Richardson (1937) proposed several new reliability formulas but did not name them. The designation “KR-20” referred to its being the 20th formula in their article. Cronbach (1951) claimed that KR-20 was an awkward name for something that would be used frequently; thus, he proposed a new

Table 4. Names and Notations of Reliability Coefficients Suggested in This Study.

	Unidimensional		Multidimensional
	Split-Half	General	General
Parallel	Split-half parallel reliability (ρ_{SP})	Parallel reliability (ρ_P)	Multidimensional parallel reliability (ρ_{MP})
Tau-equivalent	Split-half tau-equivalent reliability (ρ_{ST})	Tau-equivalent reliability (ρ_T)	Multidimensional tau-equivalent reliability (ρ_{MT})
Congeneric	Split-half congeneric reliability (ρ_{SC})	Congeneric reliability (ρ_C)	<u>Bifactor model</u> Bifactor reliability (ρ_{BF}) <u>Second-order factor model</u> Second-order factor reliability (ρ_{SOF}) <u>Correlated factors model</u> Correlated factors reliability (ρ_{CF})

name: coefficient alpha. This name, however, had the potential to be confusing because the name alpha was regularly used in research methodology textbooks to denote other concepts such as significance level (e.g., $\alpha = .05$). Thus, calling it *Cronbach's* alpha would have provided the users more discernibility than simply referring to it as alpha. Such convention, however, gives a misleading impression that Cronbach first developed the formula, and Cronbach (2004) opposed the expression *Cronbach's* alpha.

Next, let us consider composite reliability. At least seven names are used to represent a reliability estimator based on a congeneric model in the literature: composite reliability, construct reliability (e.g., Hair, Black, Babin, & Anderson, 2010), congeneric reliability (e.g., Graham, 2006; Lucke, 2005), classical congeneric reliability coefficient (Feldt & Brennan, 1989), omega (McDonald, 1999), unidimensional omega (Cho & Kim, 2015) and Raju (1977) coefficient. Among them, composite reliability is the least appropriate for the name of a specific reliability coefficient because it is shorthand for the reliability of composite scores (Cho & Kim, 2015). This name may even cause erroneous associations such as with complex or synthesized reliability. Below is my criticism in terms of history.

The originator of this formula intended to use the term *congeneric*, not *composite*. The term *composite reliability* first appeared in Werts, Rock, Linn, and Jöreskog (1978), in which the authors called it simply "reliability" but referred to it as "the composite reliability" once when they compared it with single-item reliability. Few pioneers of this formula called it composite reliability or any other term. It is ironic that the lack of an appropriate alternative caused the unintended name to become widely used. Jöreskog (1971) also did not suggest a special name for the reliability coefficient he developed and simply called it reliability. However, the term *congeneric* is what he coined to describe its measurement model. Congeneric reliability is a name that honors his contribution.

Readers should consider the parallel use of conventional and systematic names if necessary. Some conventional names, such as alpha, are too deeply embedded in our memory to remove in the short term. Even if a researcher uses only a systematic name in his or her study, the reviewers and readers are likely to be unfamiliar with it. Using both conventional and systematic names may enhance the fidelity of communication. For example, tau-equivalent reliability can be denoted as "tau-equivalent reliability (i.e., alpha)." If the systematic nomenclature becomes more common in the future, it will substitute the conventional nomenclature in the long term.

Systematic Derivation of the Formula

Although understanding how a formula was derived is the best way to understand its essence, the literature on reliability often overlooks the derivations of formulas. Four problems are associated with this practice. First, few articles comprehensively and systematically address the derivation of formulas regarding various reliability coefficients. Second, due to the lack of related research, the various original studies that discovered these formulas are being used as the sole source of derivations. Third, the derivational methods used in the original studies were not able to utilize later, more efficient methodologies, which means that the original studies use more complex methods even though simpler derivational methods have since been developed. Fourth, most reliability coefficients were derived using different sets of logic, leading to a lack of consistency in derivational methods. For example, to derive formulas that are algebraically equivalent, Hoyt (1941) used the ANOVA approach, whereas Cronbach (1951) calculated the average of the split-half reliability coefficient obtained from all possible split halves.

The current study proves that various reliability coefficients can be derived from measurement models nested within the bifactor model. The goal of the current study is to propose a derivational method for formulas that clearly reveal the basic assumptions of each reliability coefficient and the difference between each reliability coefficient. The derivations based on the unidimensional model are displayed here, and the derivations based on multidimensional models will appear in Appendix B.

The definition of reliability. Let us define test score reliability in the unidimensional model or classical true score model (Lord & Novick, 1968). Consider a test that is composed of k dichotomously or polytomously scored items. This study defines the test score X as the sum of k observed scores X_i : $X = \sum_{i=1}^k X_i$. The unidimensional model assumes that X_i consists of two unobserved scores: the true score T_i and the error e_i . This study assumes that no specific factors exist to circumvent the controversial issue of specific factors and errors (Bollen, 1989). T_i is further deconstructed into two components such that $T_i = \mu_i + \lambda_i T$. Thus, the observed score of item i is expressed as $X_i = T_i + e_i = \mu_i + \lambda_i T + e_i$. The constant μ_i has essentially no effect on variances or covariances and thus has no effect on the value of the reliability. Factor loading λ_i is interpreted as the importance, discrimination or effective length of item i . The variance of a latent variable is fixed to unity (i.e., $Var(T) = 1$) without loss of generality to determine the scale. Errors are purely random and are not correlated with the true score (i.e., $Cov(T_i, e_i) = Cov(T_i, e_j) = 0$). This study further assumes that errors are not correlated with each other (i.e., $Cov(e_i, e_j) = 0$). The reliability is defined as the ratio of the true score variance to the test score variance: $\rho_{XX'} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2}$.

The unidimensional parallel model. The unidimensional parallel model restricts the factor loading and error variance of each item to be equal (i.e., $\lambda_i = \lambda$ and $\sigma_{e_i}^2 = \sigma_e^2$ for all i). The variance of item i is $\lambda^2 + \sigma_e^2$, and the covariance between item i and j is λ^2 . Let c denote $\lambda^2 + \sigma_e^2$.

Special case ($k = 2$). Let ρ_{12} denote the ratio of λ^2 to c . ρ_{12} equals the product-moment correlation between the first and second items (or split half). The interitem covariance matrix is as follows:

$$\Sigma_{sp} = \begin{vmatrix} \lambda^2 + \sigma_e^2 & \lambda^2 \\ \lambda^2 & \lambda^2 + \sigma_e^2 \end{vmatrix} = c \begin{vmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{vmatrix} = c \begin{vmatrix} \rho_{12} & \rho_{12} \\ \rho_{12} & \rho_{12} \end{vmatrix} + c \begin{vmatrix} 1 - \rho_{12} & 0 \\ 0 & 1 - \rho_{12} \end{vmatrix}$$

The sums of the second, third, and fourth matrices are σ_X^2 , σ_T^2 , and σ_E^2 , respectively. Therefore, $\sigma_T^2 = 4c\rho_{12}$, and $\sigma_E^2 = 2c(1 - \rho_{12})$. Brown (1910) and Spearman (1910) independently and simultaneously developed an algebraically equivalent formula for split-half parallel reliability. The

derived version (ρ_{SP}) and the original version ($\tilde{\rho}_{SP}$) are as follows:

$$\rho_{SP} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} = \frac{4\rho_{12}}{4\rho_{12} + 2(1 - \rho_{12})}, \text{ and } \tilde{\rho}_{SP} = \frac{2\rho_{12}}{1 + \rho_{12}}$$

General case. Let $\bar{\rho}_{ij}$ denote the ratio of λ^2 to c . $\bar{\rho}_{ij}$ equals the average product-moment correlation between items. The interitem covariance matrix in the unidimensional parallel model is as follows:

$$\Sigma_{up} = c \begin{vmatrix} 1 & \bar{\rho}_{ij} & \cdots & \bar{\rho}_{ij} \\ \bar{\rho}_{ij} & 1 & \cdots & \bar{\rho}_{ij} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\rho}_{ij} & \bar{\rho}_{ij} & \cdots & 1 \end{vmatrix} = c \begin{vmatrix} \bar{\rho}_{ij} & \bar{\rho}_{ij} & \cdots & \bar{\rho}_{ij} \\ \bar{\rho}_{ij} & \bar{\rho}_{ij} & \cdots & \bar{\rho}_{ij} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\rho}_{ij} & \bar{\rho}_{ij} & \cdots & \bar{\rho}_{ij} \end{vmatrix} + c \begin{vmatrix} 1 - \bar{\rho}_{ij} & 0 & \cdots & 0 \\ 0 & 1 - \bar{\rho}_{ij} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 - \bar{\rho}_{ij} \end{vmatrix}$$

The sum of the second $k \times k$ matrix (i.e., σ_T^2), of which all elements are $\bar{\rho}_{ij}$, is $ck^2\bar{\rho}_{ij}$. The sum of the third $k \times k$ matrix (i.e., σ_E^2), of which all diagonal elements are $1 - \bar{\rho}_{ij}$, is $ck(1 - \bar{\rho}_{ij})$. The derived version (ρ_P) of parallel reliability is algebraically equivalent to the formula known as standardized alpha ($\tilde{\rho}_P$). Specifically,

$$\rho_P = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} = \frac{k^2\bar{\rho}_{ij}}{k^2\bar{\rho}_{ij} + k(1 - \bar{\rho}_{ij})}, \text{ and } \tilde{\rho}_P = \frac{k\bar{\rho}_{ij}}{1 + (k - 1)\bar{\rho}_{ij}}$$

The unidimensional tau-equivalent model. The unidimensional tau-equivalent model (Novick & Lewis, 1967) restricts the factor loadings of each item to be equal (i.e., $\lambda_i = \lambda$ for all i). The observed score of item i is expressed as $X_i = \mu_i + \lambda F + e_i$. The unidimensional parallel model is nested within the unidimensional tau-equivalent model, and the only difference between them is the equality restriction of error variances. The variance of item i is $\lambda^2 + \sigma_{e_i}^2$, and the covariance between item i and j is λ^2 .

Special case ($k = 2$). Let $\sigma_1^2 (= \lambda^2 + \sigma_{e_1}^2)$ and $\sigma_2^2 (= \lambda^2 + \sigma_{e_2}^2)$ denote the variance of the first and second items, respectively, or the split-half, and let $\sigma_{12} (= \lambda^2)$ denote the covariance between them. The interitem covariance matrix and split-half tau-equivalent reliability (ρ_{ST}) are as follows:

$$\Sigma_{st} = \begin{vmatrix} \lambda^2 + \sigma_{e_1}^2 & \lambda^2 \\ \lambda^2 & \lambda^2 + \sigma_{e_2}^2 \end{vmatrix} = \begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{vmatrix} = \begin{vmatrix} \sigma_{12} & \sigma_{12} \\ \sigma_{12} & \sigma_{12} \end{vmatrix} + \begin{vmatrix} \sigma_1^2 - \sigma_{12} & 0 \\ 0 & \sigma_2^2 - \sigma_{12} \end{vmatrix}, \text{ and}$$

$$\rho_{ST} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{4\sigma_{12}}{\sigma_X^2}.$$

Flanagan (1937), Guttman (λ_4 , 1945), Rulon (1939), and Mosier (1941) proposed formulas that are algebraically equivalent to ρ_{ST} :

$$\tilde{\rho}_F = \frac{4\rho_{12}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2} = \frac{4\sigma_{12}}{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}} = \frac{4\sigma_{12}}{\sigma_X^2} = \rho_{ST},$$

$$\tilde{\rho}_{Rulon} = 1 - \frac{\sigma_{X_1 - X_2}^2}{\sigma_X^2} = \frac{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}) - (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})}{\sigma_X^2} = \frac{4\sigma_{12}}{\sigma_X^2} = \rho_{ST},$$

$$\tilde{\rho}_M = \frac{4(\rho_{1X}\sigma_1\sigma_X - \sigma_1^2)}{\sigma_X^2} = \frac{4(\sigma_1^2 + \sigma_{12} - \sigma_1^2)}{\sigma_X^2} = \frac{4\sigma_{12}}{\sigma_X^2} = \rho_{ST}, \text{ and}$$

$$\tilde{\lambda}_4 = 2 \left(1 - \frac{\sigma_1^2 + \sigma_2^2}{\sigma_X^2} \right) = 2 \left(\frac{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}) - (\sigma_1^2 + \sigma_2^2)}{\sigma_X^2} \right) = \frac{4\sigma_{12}}{\sigma_X^2} = \rho_{ST}.$$

General case. Let $\bar{\sigma}_{ij}$ denote λ^2 . $\bar{\sigma}_{ij}$ equals the average covariance between items. The interitem covariance matrix is as follows:

$$\Sigma_{uu} = \begin{vmatrix} \sigma_1^2 & \bar{\sigma}_{ij} & \cdots & \bar{\sigma}_{ij} \\ \bar{\sigma}_{ij} & \sigma_2^2 & \cdots & \bar{\sigma}_{ij} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\sigma}_{ij} & \bar{\sigma}_{ij} & \cdots & \sigma_k^2 \end{vmatrix} = \begin{vmatrix} \bar{\sigma}_{ij} & \bar{\sigma}_{ij} & \cdots & \bar{\sigma}_{ij} \\ \bar{\sigma}_{ij} & \bar{\sigma}_{ij} & \cdots & \bar{\sigma}_{ij} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\sigma}_{ij} & \bar{\sigma}_{ij} & \cdots & \bar{\sigma}_{ij} \end{vmatrix} + \begin{vmatrix} \sigma_1^2 - \bar{\sigma}_{ij} & 0 & \cdots & 0 \\ 0 & \sigma_2^2 - \bar{\sigma}_{ij} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k^2 - \bar{\sigma}_{ij} \end{vmatrix}.$$

Cronbach ($\tilde{\alpha}$, 1951), Guttman ($\tilde{\lambda}_3$, 1945), Hoyt (1941), and Kuder and Richardson (1937) proposed formulas that are algebraically equivalent to tau-equivalent reliability. Kuder and Richardson (1937) expressed the formula (i.e., KR-20) slightly differently from others because they applied the formula to dichotomously score items (i.e., $X_i = 0$ or 1). Let p_i denote the percentage of correct responses for item i . We obtain the formula of tau-equivalent reliability (ρ_T) by dividing the sum of the second matrix by σ_X^2 . The conventional version ($\tilde{\rho}_T$) and KR-20 ($\tilde{\rho}_{KR-20}$) are all algebraically equivalent to ρ_T . Specifically,

$$\rho_T = \frac{\sigma_T^2}{\sigma_X^2} = \frac{k^2 \bar{\sigma}_{ij}}{\sigma_X^2}, \tilde{\rho}_T = \tilde{\alpha} = \tilde{\lambda}_3 = \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_X^2} \right), \text{ and}$$

$$\tilde{\rho}_{KR-20} = \frac{k}{k-1} \left(1 - \frac{\sum p_i(1-p_i)}{\sigma_X^2} \right) = \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_X^2} \right) = \frac{k}{k-1} \left(\frac{\sum_{i \neq j} \sum_j \sigma_{ij}}{\sigma_X^2} \right) = \frac{k^2 \bar{\sigma}_{ij}}{\sigma_X^2} = \rho_T.$$

The unidimensional congeneric model. The unidimensional congeneric model (Jöreskog, 1971) is an unrestricted base model. The observed score of item i is expressed as $X_i = \mu_i + \lambda_i T + e_i$. The unidimensional tau-equivalent model is nested within the unidimensional congeneric model, and the only difference between them is the equality restriction on the factor loadings. The variance of item i is $\lambda_i^2 + \sigma_{e_i}^2$, and the covariance between items i and j is $\lambda_i \lambda_j$.

Special case ($k = 2$). The interitem covariance matrix and split-half congeneric reliability (ρ_{SC}) are as follows:

$$\Sigma_{sc} = \begin{vmatrix} \lambda_1^2 + \sigma_{e_1}^2 & \lambda_1 \lambda_2 \\ \lambda_1 \lambda_2 & \lambda_2^2 + \sigma_{e_2}^2 \end{vmatrix} = \begin{vmatrix} \lambda_1^2 & \lambda_1 \lambda_2 \\ \lambda_1 \lambda_2 & \lambda_2^2 \end{vmatrix} + \begin{vmatrix} \sigma_{e_1}^2 & 0 \\ 0 & \sigma_{e_2}^2 \end{vmatrix}, \text{ and}$$

$$\rho_{SC} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{(\lambda_1 + \lambda_2)^2}{\sigma_X^2}.$$

This coefficient cannot be estimated without further constraints because the model is under-identified. Specifically, there are more unknowns (i.e., λ_1 , λ_2 , $\sigma_{e_1}^2$, and $\sigma_{e_2}^2$) than known pieces of information (i.e., σ_1^2 , σ_2^2 , and σ_{12}) (Feldt & Brennan, 1989; Haertel, 2006). Appendix A shows how previous studies addressed this issue.

General case. The interitem covariance matrix, the derived formula of congeneric reliability (ρ_C), and the conventional version ($\tilde{\rho}_C$) are as follows:

$$\Sigma_{uc} = \begin{vmatrix} \lambda_1^2 + \sigma_{e_1}^2 & \lambda_2\lambda_1 & \cdots & \lambda_k\lambda_1 \\ \lambda_1\lambda_2 & \lambda_2^2 + \sigma_{e_2}^2 & \cdots & \lambda_k\lambda_2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1\lambda_k & \lambda_2\lambda_k & \cdots & \lambda_k^2 + \sigma_{e_k}^2 \end{vmatrix} = \begin{vmatrix} \lambda_1^2 & \lambda_2\lambda_1 & \cdots & \lambda_k\lambda_1 \\ \lambda_1\lambda_2 & \lambda_2^2 & \cdots & \lambda_k\lambda_2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1\lambda_k & \lambda_2\lambda_k & \cdots & \lambda_k^2 \end{vmatrix} + \begin{vmatrix} \sigma_{e_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{e_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{e_k}^2 \end{vmatrix},$$

$$\rho_C = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\left(\sum \lambda_i\right)^2}{\sigma_X^2}, \text{ and } \tilde{\rho}_C = \frac{\left(\sum \lambda_i\right)^2}{\left(\sum \lambda_i\right)^2 + \sum \sigma_{e_i}^2}.$$

Systematic Expression of Formula

A formula may have multiple algebraically equivalent versions. For example, the definition of reliability can be expressed in several ways, including $\rho_{XX'} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} = 1 - \frac{\sigma_E^2}{\sigma_X^2}$. Henceforth, algebraically equivalent variations will be referred to as *versions* or *formula expressions*. When a formula is more complex, it is more difficult to observe the algebraic equivalence of the formula and its many versions. If no consistent rule exists for selecting which version to use, users will express the same formula in different ways, which will ultimately lead to an increased difficulty in understanding the meaning of the formula and discovering commonalities.

Table 5 presents a summary of conventional formula expressions. Excluding the relationship between split-half parallel reliability (i.e., the Spearman–Brown formula) and parallel reliability (i.e., standardized alpha), no commonalities or regularities between the formulas occur. Some formulas are a subtraction of a certain value from 1, and others are not. One formula has σ_X^2 as the denominator, whereas others have other elements in their denominators. It appears as if these formulas are not related to each other.

The current principle applied when expressing the formula of a reliability coefficient is that the formula expression used by the author of the original article is used. This convention is author friendly but not reader friendly. The formula expression chosen by the author and the expression easily understood by readers is different for three reasons. First, each formula is derived using different methods with various interpretations. The derivational method and the interpretation of the formulas also influenced their expression.

Second, authors are rewarded for the distinctiveness of their work, not consistency. Academia prefers studies that are different from previous studies; those that present results that are similar to existing findings are not highly regarded. Researchers may attempt to express different versions of formulas and propose different methods of interpretation.

Third, authors have preferred formula expressions that are computationally simple. Many reliability formulas were published in an era during which they could be calculated only by hand instead of with computers. Because of social inertia, computationally simpler formulas have been chosen even after computer-based computational processes became common. Thus, the most famous versions of formulas are typically those that require the fewest calculations (Falk & Savalei, 2011).

The three problems mentioned above suggest a direction for our proposed systematic formula expression. First, the system should include a consistent set of principles. Second, this study’s goal is to connect apparently disparate formulas to a degree in which users can easily recognize the common features of the formulas. Third, its goal is not to provide ease of calculation; rather, it

Table 5. Conventional Formula Expressions of Reliability Coefficients.

	Unidimensional		Multidimensional
	Split-Half	General	General
Parallel	$\frac{2\rho_{12}}{1+\rho_{12}}$	$\frac{k\bar{\rho}_{ij}}{1+(k-1)\bar{\rho}_{ij}}$	(Not yet published)
Tau-equivalent	$1 - \frac{\sigma_{x_1-x_2}^2}{\sigma_x^2}$	$\frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_x^2} \right)$	$1 - \frac{\sum_{p=1}^m \sigma_p^2 (1 - \alpha_p)}{\sigma_x^2}$
Congeneric	$\frac{\sigma_{12}}{\pi_1 \pi_2 \sigma_x^2}$	$\frac{\left(\sum_{i=1}^k \lambda_i \right)^2}{\left(\sum_{i=1}^k \lambda_i \right)^2 + \sum_{i=1}^k \sigma_{e_i}^2}$	$\frac{\mathbf{1}' \mathbf{c} \mathbf{c}' \mathbf{1} + \mathbf{1}' \mathbf{A} \mathbf{A}' \mathbf{1}}{\sigma_x^2}$

Note: σ_x^2 is observed test score variance, k is the number of items, $\bar{\rho}_{ij}$ is the average product-moment correlation between items, λ_i in the unidimensional models is the factor loading of item i , and m is the number of group factors. Other notations are explained in Appendices A and B.

wants to provide ease of understanding. In the current environment in which all calculations are dependent on computers, ease of calculation cannot be an important consideration for users.

Table 6 summarizes the formula expressions that the current study proposes. This study applies a consistent set of principles. More specifically, between matrix expressions and nonmatrix expressions, this study uses nonmatrix expressions. The denominator is made consistent to be σ_x^2 or its equivalent, and the numerator is made consistent to be equivalent to σ_T^2 . The sum is presented in the order of *true score + error* or *general factor + group factor + error*. This study does not reduce fractions to their lowest terms. For example, in the formula expression of parallel reliability, that is,

$$\frac{k^2 \bar{\rho}_{ij}}{k^2 \bar{\rho}_{ij} + k(1 - \bar{\rho}_{ij})}, k^2 \bar{\rho}_{ij} \text{ represents } \sigma_T^2 \text{ and } k(1 - \bar{\rho}_{ij}) \text{ denotes } \sigma_E^2.$$

The systematic formula expressions provide us with an intuitive understanding about the condition under which reliability coefficients have a value of less than zero. The definition of reliability

$$\left(\text{i.e., } \rho_{XX'} = \frac{\sigma_T^2}{\sigma_X^2} = 1 - \frac{\sigma_E^2}{\sigma_X^2} \right) \text{ tells us that the value of reliability always ranges between zero and one.}$$

Contrary to the popular belief that the value of a reliability estimate also always ranges between zero and one, it can be negative if the prerequisite of the reliability estimator is not satisfied. Table 6 reveals that the value of congeneric model-based reliability coefficients is always nonnegative, whereas the value of parallel or tau-equivalent model-based reliability coefficients can be negative.

This study derives two hierarchical omega coefficients from multidimensional parallel and multidimensional tau-equivalent models. McDonald (1985, 1999) proposed two definitions of reliability that are applicable to multidimensional models, and Zinbarg, Revelle, Yovel, and Li (2005) more explicitly expressed McDonald’s proposal, recommending it be categorized into ω , which were relabeled as ω_r in Revelle and Zinbarg (2009), and ω_H (or ω_h). ω_H is called hierarchical omega or omega hierarchical. The formula of the McDonald–Zinbarg hierarchical omega includes only the variance due to a general factor in the numerator (i.e., $\omega_H = \frac{\sigma_{General}^2}{\sigma_X^2}$), whereas that of omega includes the variance due to general factor and the variance due to group factors in the numerator (i.e., $\rho_{BF} = \omega = \frac{\sigma_{General}^2 + \sigma_{Group}^2}{\sigma_X^2}$). Although obtaining bifactor or second-order factor hierarchical omega requires parameter estimates that are obtained from multidimensional SEM models, estimating parallel/tau-equivalent hierarchical omega does not rely on SEM procedures. Previous studies strongly have recommended the use of hierarchical omega (Revelle &

Table 6. Systematic Formula Expressions of Reliability Coefficients.

	Unidimensional		Multidimensional
	Split-Half	General	General
Parallel	$\frac{4\rho_{12}}{4\rho_{12} + 2(1 - \rho_{12})}$	$\frac{k^2 \bar{\rho}_{ij}}{k^2 \bar{\rho}_{ij} + k(1 - \bar{\rho}_{ij})}$	$\frac{k^2 \bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d)}{k^2 \bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d) + k(1 - \bar{\rho}_s)}$
Tau-equivalent	$\frac{4\sigma_{12}}{\sigma_x^2}$	$\frac{k^2 \bar{\sigma}_{ij}}{\sigma_x^2}$	$\frac{k^2 \bar{\sigma}_d + nk(\bar{\sigma}_s - \bar{\sigma}_d)}{\sigma_x^2}$
Congeneric	$\frac{(\lambda_1 + \lambda_2)^2}{\sigma_x^2}$	$\frac{\left(\sum_{i=1}^k \lambda_i\right)^2}{\sigma_x^2}$	<p>Bifactor reliability</p> $\frac{\left(\sum_{i=1}^k \lambda_{iF}\right)^2 + \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip}\right)^2}{\sigma_x^2}$ <p>Correlated factors reliability</p> $\frac{\sum_{p=1}^m \sum_{q=1}^m \phi_{pq} \left(\sum_{i=1}^k \lambda_{ip}\right) \left(\sum_{i=1}^k \lambda_{iq}\right)}{\sigma_x^2}$ <p>Second-order factor reliability</p> $\frac{\left(\sum_{p=1}^m \sum_{i=1}^k \lambda_{ip} \gamma_p\right)^2 + \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip} (1 - \gamma_p^2)^{1/2}\right)^2}{\sigma_x^2}$

Note: $\bar{\sigma}_{ij}$ is the average covariance between items, λ_{iF} is the general factor loading of item i , λ_{ip} is the factor loading of item i on p th subtest construct (i.e., G_p , O_p , or D_p), n is the number of items per group factor, $\bar{\rho}_d$ is the average product-moment correlation between items that have different group factors, $\bar{\rho}_s$ is the average product-moment correlation between items that have the same group factor, $\bar{\sigma}_d$ is the average covariance between items that have different group factors, $\bar{\sigma}_s$ is the average covariance between items that have the same group factor and m is the number of group factors. Other notations are explained in Appendices A and B.

Zinbarg, 2009; Zinbarg et al., 2005; Zinbarg et al., 2007; Zinbarg, Yovel, Revelle, & McDonald, 2006). If the concept of hierarchical omega is useful for bifactor models, it must be equally advantageous to multidimensional tau-equivalent and multidimensional parallel models. Table 7 displays their formulas.

This study classifies the reliability of multidimensional models into two hierarchical levels: the reliability of a measurement model and the reliability of subtest constructs. Offering both test reliability and subtest reliability estimates may provide more complete information about the measurement. Table 8 shows the formulas of subtest reliability. Formulas of subtest reliability coefficients of multidimensional parallel, multidimensional tau-equivalent and correlated factors models are a minor modification of those of parallel reliability, tau-equivalent reliability, and congeneric reliability. Subtest reliability coefficients of bifactor and second-order factor models do not have analogous unidimensional reliability coefficients. Brunner et al. (2012) extended the concept of hierarchical omega to the subtest reliability level; however, they used the term to denote the ratio of the variances due to group factors to test score variances (i.e., % group factor in Table 8), which is different from the way Zinbarg et al. (2005) defined hierarchical omega at the test reliability level (i.e., % general factor). Because using the term *hierarchical omega* at the level of subtest reliability can be confusing, this study employed generic terms.

Systematic Use

Although tau-equivalent reliability (i.e., alpha) is the most popular reliability coefficient among organizational researchers, experts treat it quite differently. Previous studies are practically unanimous in declaring that there must be an alternative to the current practice of indiscriminately using this coefficient,

Table 7. Four Coefficients of Hierarchical Omega.

	Description	Formula
Parallel	Parallel hierarchical omega (ω_{HP})	$\frac{k^2 \bar{\rho}_d}{k^2 \bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d) + k(1 - \bar{\rho}_s)}$
Tau-equivalent	Tau-equivalent hierarchical omega (ω_{HT})	$\frac{k^2 \bar{\sigma}_d}{\sigma_x^2}$
Second-order factor	Second-order factor hierarchical omega (ω_{HSOF})	$\frac{\left(\sum_{p=1}^m \sum_{i=1}^k \lambda_{ip} \gamma_p\right)^2}{\sigma_x^2}$
Bifactor	Bifactor hierarchical omega (ω_{HBF})	$\frac{\left(\sum_{i=1}^k \lambda_{iF}\right)^2}{\sigma_x^2}$

Note: The notations are explained in Table 4, and the derivations are listed in Appendix B.

Table 8. Formulas of Subtest Reliability in Multidimensional Models.

	Subtest Reliability	% General Factor	% Group Factor
Multidimensional Parallel (ρ_{MP})	$\frac{n^2 \bar{\rho}_{X_i \in G_p}}{n^2 \bar{\rho}_{X_i \in G_p} + n(1 - \bar{\rho}_{X_i \in G_p})}$	—	—
Multidimensional Tau-equivalent (ρ_{MT})	$\frac{n^2 \bar{\sigma}_{X_i \in G_p}}{\sigma_{X_i \in G_p}^2}$	—	—
Bifactor (ρ_{BF})	$\frac{\left(\sum_{i \in G_p} \lambda_{iF}\right)^2 + \left(\sum_{i \in G_p} \lambda_{ip}\right)^2}{\sigma_{X_i \in G_p}^2}$	$\frac{\left(\sum_{i \in G_p} \lambda_{iF}\right)^2}{\sigma_{X_i \in G_p}^2}$	$\frac{\left(\sum_{i \in G_p} \lambda_{ip}\right)^2}{\sigma_{X_i \in G_p}^2}$
Second-order factor (ρ_{SOF})	$\frac{\left(\sum_{i \in O_p} \lambda_{ip} \gamma_p\right)^2 + \left(\sum_{i \in O_p} \lambda_{ip} (1 - \gamma_p^2)^{1/2}\right)^2}{\sigma_{X_i \in O_p}^2}$	$\frac{\left(\sum_{i \in O_p} \lambda_{ip} \gamma_p\right)^2}{\sigma_{X_i \in O_p}^2}$	$\frac{\left(\sum_{i \in O_p} \lambda_{ip} (1 - \gamma_p^2)^{1/2}\right)^2}{\sigma_{X_i \in O_p}^2}$
Correlated factors (ρ_{CF})	$\frac{\left(\sum_{i \in D_p} \lambda_{ip}\right)^2}{\sigma_{X_i \in D_p}^2}$	—	—

although little consensus exists about exactly which alternative technique should replace alpha (Bentler, 2009; Green & Yang, 2009b; Hunt & Bentler, 2015; Osburn, 2000; Revelle & Zinbarg, 2009; Schmidt, Le, & Ilies, 2003; Sijtsma, 2009b; van der Ark et al., 2011). The current study does not propose a specific reliability estimator as an alternative; rather, it delineates a system composed of multiple methods.

This study does not agree with the unconditional use and denial of tau-equivalent reliability. Criticizing alpha’s unconditional use is different from advocating that another alternative should take its seat and be universally used instead. To prevent the blind use of tau-equivalent reliability, previous studies have criticized it using rather strong language. For example, Peters (2014) claimed that we should abandon alpha because it is “a fatally flawed estimate of its reliability” (p. 56). This study encourages the use of tau-equivalent reliability if the data meet the condition of being tau-equivalent. What this study disapproves of is the concept of a single best reliability coefficient that is appropriate for all types of data sets, which implicitly assumes that the objective function is one dimensional.

Any reliability coefficient in a system is not superior or inferior to another because they simply assume different measurement models. The evaluation criteria of a scientific model are at least two

STEP 1. Identify the dimensionality of the data.

Dependencies: Chi-square difference test, exploratory factor analysis, confirmatory factor analysis
 Decision: If unidimensional, go to step 2; otherwise, go to step 3.

STEP 2. Identify the statistical similarity of the unidimensional data

Dependencies: Chi-square difference test
 Decision: If parallel, use ρ_p (i.e., standardized alpha). If tau-equivalent, use ρ_T (i.e., coefficient alpha); otherwise, use ρ_C (i.e., so-called composite reliability).

STEP 3. Determine the measurement model of the multidimensional data

Dependencies: Chi-square difference test and theoretical considerations
 Decision: If multidimensional parallel, use ρ_{MP} , ω_{HP} and sub-test reliability coefficients. If multidimensional tau-equivalent, use ρ_{MT} (i.e., stratified alpha), ω_{HT} and sub-test reliability coefficients. If the bi-factor model, use ρ_{BF} , ω_{HB} and sub-test reliability coefficients. If the second-order factor model, use ρ_{SOF} , ω_{HS} and sub-test reliability coefficients. If the correlated factor model, use ρ_{CF} and sub-test reliability coefficients.

Figure 4. A guideline for choosing a reliability coefficient.

dimensional; a good model should explain the maximum amount of data using the fewest elements. SEM measurement models provide a trade-off between goodness of fit (i.e., chi-square) and parsimony (i.e., degree of freedom), and less parsimonious models should have significantly better goodness of fit. Unidimensional models are nested within multidimensional models, and tau-equivalent models are nested within congeneric models. When comparing two competing SEM models where one is nested within another, we typically use the chi-square difference to test significance.

Recommendations for use. Figure 4 shows a guideline for choosing a reliability coefficient. The chi-square difference test represents major statistical criteria at all steps. How to identify the dimensionality (i.e., STEP 1) of data is an important but controversial issue. Numerous methods have been proposed to test unidimensionality (Hattie, 1985). Sometimes, exploratory factor analysis or CFA can be used to examine the dimensionality of data. This study will introduce a gadget that enables users to perform CFA without SEM software.

Organizational researchers should combine theoretical considerations with statistical criteria when deciding which measurement model to use. A researcher's judgment is especially important when choosing between multidimensional models (i.e., STEP 3). A bifactor model or a second-order model have a hierarchy that consists of a common construct and subtest constructs. A correlated factors model is composed of only subtest constructs. A key question is whether the researcher has a theoretical interest in a common construct and whether it has theoretical underpinnings (Brunner et al., 2012). The answer is affirmative in many cases because a common construct is what scale developers originally intended to measure and what typical researchers are most interested in (Cho & Kim, 2015; Reise, 2012). The use of correlated factor reliability is questionable because the use of a total test score of a correlated factors model is not recommended (McDonald, 1999). Brunner et al. (2012) advocated the use of only subtest reliability coefficients when a correlated factors model is selected. Reporting correlated factor reliability may have the merit of providing additional information about the measurement; however, complete reliance on it is not recommended.

Examples and a Calculator

Illustrative Examples

Understanding a formula promotes its frequent use. Organizational researchers rarely use multidimensional reliability coefficients, even though they usually study multifaceted phenomena and analyze multidimensional data. Such lack of use probably originates from their limited awareness

Table 9. A Computation of Tau-equivalent Reliability and Congeneric Reliability.

Observed Covariance Matrix							
	X1	X2	X3	X4	$\sum_{total} = \sum_{diagonal} + \sum_{off-diagonal} = 124$		
X1	10				$\sum_{diagonal} = 10 + 11 + 12 + 13 = 46$		
X2	4	11			$\sum_{off-diagonal} = 2 \cdot \sum_{sub-diagonal} = 2 \cdot (4 + 5 + 6 + 7 + 8 + 9) = 78$		
X3	5	6	12				
X4	7	8	9	13			
Proposed formula expression $\hat{\rho}_T = \hat{\alpha} = 4^2 \cdot (78 / (4 \cdot 3)) / 124 = .8387$							
Conventional formula expression $\hat{\rho}_T = \hat{\alpha} = (4/3) \cdot (1 - 46/124) = .8387$							
Fitted/Implied Covariance Matrix				Factor Loadings and Errors			
	X1	X2	X3	X4	$\hat{\lambda}_i$	$\hat{\sigma}_{e_i}^2$	
X1	10.00				X1	1.96	6.13
X2	4.42	11.00			X2	2.25	5.92
X3	4.98	5.71	12.00		X3	2.53	5.56
X4	6.98	7.99	9.01	13.00	X4	3.55	.37
\sum	124.23				\sum	10.30	18.01
	$= 2 \cdot (\sum_{sub-diagonal} + \sum_{diagonal}) - \sum_{diagonal}$			\sum	\sum^2	106.22	
Proposed formula expression $\hat{\rho}_C = 106.22 / 124.23 = .8550$							
Conventional formula expression $\hat{\rho}_C = 106.22 / (106.22 + 18.01) = .8550$							

and understanding of multidimensional reliability coefficients. The terms *awareness* and *understanding* denote different meanings here. For example, thus far, this study is likely to make readers aware of the names and formulas of multidimensional reliability coefficients but is unlikely to make the readers understand their meaning. A good way to understand a complex formula is to apply it to a simple numerical example. This study presents a unidimensional example and five multidimensional examples to make the formulas more comprehensible to readers.

Let us start from a comparison of tau-equivalent reliability and congeneric reliability. Table 9 omits the upper triangle of the covariance matrices, as done by typical SEM software packages. Readers should be careful to use the sum of the fitted or implied covariance matrix (i.e., $\hat{\sigma}_X^2$) instead of the sum of the observed covariance matrix (i.e., σ_X^2) when they calculate the estimate of congeneric reliability. The fitted covariance matrix is computed from the parameter estimates. For example, $4.42 = 1.96 \cdot 2.25$ and $10.00 = 1.96^2 + 6.13$.

Table 10 presents a decomposition of the covariance matrix of multidimensional parallel data. Readers may find the meaning of multidimensional parallel reliability difficult to understand because it is new to them and has a complex form. The denominator of the formula of multidimensional parallel reliability, that is, $\frac{k^2 \bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d)}{k^2 \bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d) + k(1 - \bar{\rho}_s)}$, consists of three parts, corresponding to the sum of all elements of **B**, **C**, and **D** in Table 10. $k^2 \bar{\rho}_d$ is proportional to the variance due to the general factor ($\sigma_{General}^2$). In the example, the number of items (k) is 6, and the average product-moment correlation between items that have different group factors is .3, which leads to the value of $6^2 \cdot .3 = 10.8$. $nk(\bar{\rho}_s - \bar{\rho}_d)$ is proportional to the variance due to group factors (σ_{Group}^2). In Table 10, the average product-moment correlation between items that have the same group factor is .5, and the sum of all elements of **C** is $3 \cdot 6 \cdot (.5 - .3) = 3.6$. $k(1 - \bar{\rho}_s)$ is proportional to the variance due to errors (σ_E^2), and the corresponding value in Table 10 is $6(1 - .5) = 3$. Substitution of these values into the formula leads to $\hat{\rho}_{MP} = .8276$.

Table 10. A Computation Example of Multidimensional Parallel Reliability.

A Observed Test Score Variance ($\sigma_X^2 = 17.4$)							B The Variance Due to General Factor ($\sigma_{General}^2 = 6^2 \cdot .3 = 10.8$)					
	X1	X2	X3	X4	X5	X6	X1	X2	X3	X4	X5	X6
X1	1	.5	.5	.3	.3	.3	.3	.3	.3	.3	.3	.3
X2	.5	1	.5	.3	.3	.3	.3	.3	.3	.3	.3	.3
X3	.5	.5	1	.3	.3	.3	.3	.3	.3	.3	.3	.3
X4	.3	.3	.3	1	.5	.5	.3	.3	.3	.3	.3	.3
X5	.3	.3	.3	.5	1	.5	.3	.3	.3	.3	.3	.3
X6	.3	.3	.3	.5	.5	1	.3	.3	.3	.3	.3	.3

C The Variance Due to Group Factors ($\sigma_{Group}^2 = 3 \cdot 6 \cdot (.5 - .3) = 3.6$)						D The Variance Due to Errors ($\sigma_E^2 = 6 \cdot (1 - .5) = 3$)						
	X1	X2	X3	X4	X5	X6	X1	X2	X3	X4	X5	X6
X1	.2	.2	.2	0	0	0	.5	0	0	0	0	0
X2	.2	.2	.2	0	0	0	0	.5	0	0	0	0
X3	.2	.2	.2	0	0	0	0	0	.5	0	0	0
X4	0	0	0	.2	.2	.2	0	0	0	.5	0	0
X5	0	0	0	.2	.2	.2	0	0	0	0	.5	0
X6	0	0	0	.2	.2	.2	0	0	0	0	0	.5

Note: If tau-equivalent reliability (e.g., alpha) is misapplied to the above data, it underestimates the reliability by .0414 ($\hat{\rho}_T = \hat{\alpha} = (6^2 \cdot ((.3 \cdot 9 + .5 \cdot 6)/15))/17.4 = .7862$).

Matrices **A** and **B** of Table 11 illustrate the commonalities and differences between multidimensional tau-equivalent reliability and other related reliability coefficients. First, the values of multidimensional tau-equivalent reliability equal those of stratified alpha, and the computation of the former is more understandable than that of the latter. Second, tau-equivalent reliability (i.e., alpha) underestimates the reliability when it is applied to multidimensional data. Third, the values of hierarchical omega that were computed from the two data sets differ greatly, even though those of reliability are equal.

Let us further discuss the relationships between reliability, hierarchical omega and unidimensionality based on Table 11. First, unidimensionality should be identified before reliability is calculated. Many users perform the reverse; they calculate reliability to identify unidimensionality, misconceiving that a high value of a unidimensional reliability coefficient indicates unidimensionality (Cortina, 1993; Green, Lissitz, & Mulaik, 1977; Hattie, 1985; Schmitt, 1996). Table 11 disproves this misconception; the tau-equivalent reliability estimates of Matrices **A** and **B** exceed the commonly used cut-off criterion (Lance, Butts, & Michels, 2006), even though both were computed from multidimensional data.

Second, the concept of hierarchical omega and unidimensionality should be distinguished. A glance at Table 11 may give readers a misleading impression that the level of hierarchical omega is related with the degree of unidimensionality or homogeneity. Although hierarchical omega is a matter of degree, dimensionality is a yes-or-no issue (Zinbarg et al., 2006); all data are either unidimensional or multidimensional. Unidimensionality should be distinguished from the degree to which total test scores reflect a common construct (Reise, 2012). For example, Matrix **A** is multidimensional, but most of its variances are due to a general factor. Matrix **D** is unidimensional, but only a small fraction of its variances originate from a general factor.

Third, hierarchical omega is not a substitute of reliability; it is a complement of reliability. Although hierarchical omega originated from another definition of reliability that is derived from

Table 11. A Computation Example of Multidimensional Tau-equivalent Reliability.

A Multidimensional (Strong General Factor and Weak Group Factors)							B Multidimensional (Weak General Factor and Strong Group Factors)					
	X1	X2	X3	X4	X5	X6	X1	X2	X3	X4	X5	X6
X1	8	5	5	4	4	4	10	8	8	1	1	1
X2	5	8	5	4	4	4	8	12	8	1	1	1
X3	5	5	7	4	4	4	8	8	12	1	1	1
X4	4	4	4	8	5	5	1	1	1	12	8	8
X5	4	4	4	5	8	5	1	1	1	8	10	8
X6	4	4	4	5	5	9	1	1	1	8	8	10

$\hat{\rho}_{MT} = (6^2 \cdot 4 + 3 \cdot 6 \cdot (5 - 4))/180 = .9$	$\hat{\rho}_{MT} = (6^2 \cdot 1 + 3 \cdot 6 \cdot (8 - 1))/180 = .9$
$\hat{\omega}_{HT} = (6^2 \cdot 4)/180 = .8$	$\hat{\omega}_{HT} = (6^2 \cdot 1)/180 = .2$
$\hat{\rho}_T = \hat{\alpha} = (6^2 \cdot 4.4)/180 = .88$	$\hat{\rho}_T = \hat{\alpha} = (6^2 \cdot 3.8)/180 = .76$
$\hat{\alpha}_{Stratified} = 1 - (53 \cdot .1510 + 55 \cdot .1819)/180 = .9$	$\hat{\alpha}_{Stratified} = 1 - (82 \cdot .1220 + 80 \cdot .1)/180 = .9$
$1\hat{\rho}_{MT} = (3^2 \cdot 5)/53 = .8490$	$1\hat{\rho}_{MT} = (3^2 \cdot 8)/82 = .8780$
$2\hat{\rho}_{MT} = (3^2 \cdot 5)/55 = .8181$	$2\hat{\rho}_{MT} = (3^2 \cdot 8)/80 = .9$

C Unidimensional (Strong General Factor and No Group Factors)							D Unidimensional (Weak General Factor and No Group Factors)					
	X1	X2	X3	X4	X5	X6	X1	X2	X3	X4	X5	X6
X1	10	4	4	4	4	4	20	1	1	1	1	1
X2	4	10	4	4	4	4	1	25	1	1	1	1
X3	4	4	8	4	4	4	1	1	25	1	1	1
X4	4	4	4	10	4	4	1	1	1	25	1	1
X5	4	4	4	4	12	4	1	1	1	1	25	1
X6	4	4	4	4	4	10	1	1	1	1	1	30

$\hat{\rho}_T = \hat{\alpha} = \hat{\omega}_{HT} = (6^2 \cdot 4)/180 = .8$	$\hat{\rho}_T = \hat{\alpha} = \hat{\omega}_{HT} = (6^2 \cdot 1)/180 = .2$
--	--

multidimensional models, its characteristics are different from other reliability coefficients. For example, matrices **A** and **B** of Table 11 show that the values of hierarchical omega are substantially smaller than those of other reliability coefficients. Various reliability coefficients share the common definition of $1 - \frac{\sigma_E^2}{\sigma_X^2}$, whereas hierarchical omega is defined by $1 - \left(\frac{\sigma_{Group}^2 + \sigma_E^2}{\sigma_X^2} \right)$. Hierarchical omega does not differentiate the variances due to group factors from the variances due to errors. Neither reliability nor hierarchical omega offers complete information about the data. For example, knowing that the value of the multidimensional tau-equivalent reliability is .9 does not allow us to judge whether it was calculated from **A** or **B**. The value of hierarchical omega alone cannot discriminate between **A** and **C** nor between **B** and **D**. Hierarchical omega displays a different aspect of the data that reliability does not show, and using a combination of the two enables us to see other sides of the data.

Table 12 shows a computation of bifactor reliability. The parameter estimates are $\hat{\lambda}_{1F} = \hat{\lambda}_{6F} = \hat{\lambda}_{31} = \hat{\lambda}_{42} = 1$, $\hat{\lambda}_{2F} = \hat{\lambda}_{5F} = \hat{\lambda}_{11} = \hat{\lambda}_{21} = \hat{\lambda}_{52} = 2$, $\hat{\lambda}_{3F} = \hat{\lambda}_{62} = 3$, $\hat{\lambda}_{4F} = 4$ and $\hat{\lambda}_{12} = \hat{\lambda}_{22} = \hat{\lambda}_{32} = \hat{\lambda}_{41} = \hat{\lambda}_{51} = \hat{\lambda}_{61} = 0$. A proportionality constraint does not exist (e.g., $1 : 2 : 3 \neq 2 : 2 : 1$). These values lead to $\hat{\rho}_{BF} = (169 + 61)/260 = .8846$, $1\hat{\rho}_{BF} = (36 + 25)/75 = .8133$, $2\hat{\rho}_{BF} = (49 + 36)/101 = .8415$, and $\hat{\omega}_{HBF} = 169/260 = .65$. The value of the tau-equivalent reliability (i.e., alpha) is .7795, which underestimates the reliability by .1051.

Table 12. A Computation Example of Bifactor Reliability.

A Observed Test Score Variance $\sigma_X^2 = 260$							B Variances Due to a General Factor $(1 + 2 + 3 + 4 + 2 + 1)^2 = 169$						
	X1	X2	X3	X4	X5	X6		1	2	3	4	2	1
X1	10	6	5	4	2	1	1	1·1	1·2	1·3	1·4	1·2	1·1
X2	6	12	8	8	4	2	2	2·1	2·2	2·3	2·4	2·2	2·1
X3	5	8	15	12	6	3	3	3·1	3·2	3·3	3·4	3·2	3·1
X4	4	8	12	22	10	7	4	4·1	4·2	4·3	4·4	4·2	4·1
X5	2	4	6	10	14	8	2	2·1	2·2	2·3	2·4	2·2	2·1
X6	1	2	3	7	8	15	1	1·1	1·2	1·3	1·4	1·2	1·1

C Variances Due to Group Factors $(2 + 2 + 1)^2 + (1 + 2 + 3)^2 = 61$							D Variances Due to Errors $\sigma_E^2 = 30$						
	2	2	1	1	2	3		X1	X2	X3	X4	X5	X6
2	2·2	2·2	2·1	0	0	0	X1	5	0	0	0	0	0
2	2·2	2·2	2·1	0	0	0	X2	0	4	0	0	0	0
1	1·2	1·2	1·1	0	0	0	X3	0	0	5	0	0	0
1	0	0	0	1·1	1·1	1·3	X4	0	0	0	5	0	0
2	0	0	0	2·1	2·1	2·3	X5	0	0	0	0	6	0
3	0	0	0	3·1	3·1	3·3	X6	0	0	0	0	0	5

Table 13. A Computation Example of Second-order Factor Reliability.

A Observed Test Score Variance $\sigma_X^2 = 1400$							B Variances Due to a Second-Order Factor $(3 + 6 + 3 + 8 + 4 + 4)^2 = 784$						
	Y1	Y2	Y3	Y4	Y5	Y6		3	6	3	8	4	4
Y1	60	50	25	24	12	12	3	3·3	3·6	3·3	3·8	3·4	3·4
Y2	50	126	50	48	24	24	6	6·3	6·6	6·3	6·8	6·4	6·4
Y3	25	50	60	24	12	12	3	3·3	3·6	3·3	3·8	3·4	3·4
Y4	24	48	24	130	50	50	8	8·3	8·6	8·3	8·8	8·4	8·4
Y5	12	24	12	50	60	25	4	4·3	4·6	4·3	4·8	4·4	4·4
Y6	12	24	12	50	25	80	4	4·3	4·6	4·3	4·8	4·4	4·4

C Variances Due to Disturbances $(4 + 8 + 4)^2 + (6 + 3 + 3)^2 = 400$							D Variances Due to Errors $\sigma_E^2 = 216$						
	4	8	4	6	3	3		Y1	Y2	Y3	Y4	Y5	Y6
4	4·4	4·8	4·4	0	0	0	Y1	35	0	0	0	0	0
8	8·4	8·8	8·4	0	0	0	Y2	0	26	0	0	0	0
4	4·4	4·8	4·4	0	0	0	Y3	0	0	35	0	0	0
6	0	0	0	6·6	6·3	6·3	Y4	0	0	0	30	0	0
3	0	0	0	3·6	3·3	3·3	Y5	0	0	0	0	35	0
3	0	0	0	3·6	3·3	3·3	Y6	0	0	0	0	0	35

Table 13 shows a computation of second-order factor reliability. The parameter estimates are $\hat{\lambda}_{11} = \hat{\lambda}_{31} = \hat{\lambda}_{52} = \hat{\lambda}_{62} = 5$, $\hat{\lambda}_{21} = \hat{\lambda}_{42} = 10$, $\hat{\lambda}_{12} = \hat{\lambda}_{22} = \hat{\lambda}_{32} = \hat{\lambda}_{41} = \hat{\lambda}_{51} = \hat{\lambda}_{61} = 0$, $\hat{\gamma}_1 = .6$, $\hat{\gamma}_2 = .8$, $(1 - \hat{\gamma}_1^2)^{1/2} = .8$, $(1 - \hat{\gamma}_2^2)^{1/2} = .6$, and $\sigma_X^2 = \hat{\sigma}_X^2 = 1400$. The italicized numbers on the outside of the box originate from the parameter estimates. For example, 3 in **B** is $\hat{\lambda}_1 \cdot \hat{\gamma}_1 = 5 \cdot .6 = 3$, and 4 in **C** is $\hat{\lambda}_1 \cdot (1 - \hat{\gamma}_1^2)^{1/2} = 5 \cdot .8 = 4$. A proportionality constraint exists such that

Table 14. A Computation Example of Correlated Factor Reliability.

A Observed Test Score Variance $\sigma_X^2 = 200$							B Variances Only Due to Lambdas $(2 + 2 + 4)^2 + (2 + 3 + 2)^2 = 113$						
	X1	X2	X3	X4	X5	X6		2	2	4	2	3	2
X1	8	4	8	2	3	2	2	2.2	2.2	2.4	0	0	0
X2	4	8	8	2	3	2	2	2.2	2.2	2.4	0	0	0
X3	8	8	20	4	6	4	4	4.2	4.2	4.4	0	0	0
X4	2	2	4	10	6	4	2	0	0	0	2.2	2.3	2.2
X5	3	3	6	6	16	6	3	0	0	0	3.2	3.3	3.2
X6	2	2	4	4	6	10	2	0	0	0	2.2	2.3	2.2

C Variances Due to Correlations and Lambdas $0.5 \cdot 2 \cdot (2 + 2 + 4) \cdot (2 + 3 + 2) = 56$							D Variances Due to Errors $\sigma_E^2 = 31$					
	2	2	4	2	3	2	X1	X2	X3	X4	X5	X6
1	0	0	0	1.2	1.3	1.2	X1	4	0	0	0	0
1	0	0	0	1.2	1.3	1.2	X2	0	4	0	0	0
2	0	0	0	2.2	2.3	2.2	X3	0	0	4	0	0
1	1.2	1.2	1.4	0	0	0	X4	0	0	0	6	0
1.5	1.5.2	1.5.2	1.5.4	0	0	0	X5	0	0	0	0	7
1	1.2	1.2	1.4	0	0	0	X6	0	0	0	0	6

3 : 6 : 3 = 4 : 8 : 4 and 8 : 4 : 4 = 6 : 3 : 3. These values lead to $\hat{\rho}_{SOF} = (784 + 400)/1400 = .8457$, ${}_1\hat{\rho}_{SOF} = (144 + 256)/496 = .8064$, ${}_2\hat{\rho}_{SOF} = (256 + 144)/520 = .7692$, and $\hat{\omega}_{HSOF} = 784/1400 = .56$. The value of the tau-equivalent reliability (i.e., alpha) is .7577, which underestimates the reliability by .088.

Table 14 shows a computation of the correlated factor reliability. The parameter estimates are $\hat{\lambda}_{11} = \hat{\lambda}_{21} = \hat{\lambda}_{42} = \hat{\lambda}_{62} = 2$, $\hat{\lambda}_{31} = 4$, $\hat{\lambda}_{52} = 3$, $\hat{\lambda}_{12} = \hat{\lambda}_{22} = \hat{\lambda}_{32} = \hat{\lambda}_{41} = \hat{\lambda}_{51} = \hat{\lambda}_{61} = 0$, $\hat{\phi}_{11} = \hat{\phi}_{22} = 1$, and $\hat{\phi}_{12} = \hat{\phi}_{21} = .5$. The italicized numbers on the outside of the box originate from the parameter estimates. For example, 1.5 in C is $\hat{\phi}_{12} \cdot \hat{\lambda}_{52} = .5 \cdot 3 = 1.5$. These values lead to $\hat{\rho}_{CF} = (113 + 56)/200 = .8450$, ${}_1\hat{\rho}_{CF} = 64/76 = .8421$, and ${}_2\hat{\rho}_{CF} = 49/68 = .7205$. The value of the tau-equivalent reliability (i.e., alpha) is .7680, which underestimates the reliability by .077.

RelCalc: A Calculator That Computes Reliability Coefficients

A key to ending the blind use of tau-equivalent reliability (i.e., alpha) is improving the user convenience of its alternatives. Thus far, this study has provided various solutions to resolve the common misuse of reliability coefficients. Now, it offers a quick fix to the last but not least problem. Whereas popular statistical software packages such as SPSS and SAS offer an automatic calculation of tau-equivalent reliability, commonly used SEM software packages, except EQS (Bentler, 2006), do not produce SEM-based reliability estimates. Users of such programs should personally calculate the value of a reliability coefficient based on its formula. Such computations are inconvenient and susceptible to mistakes. This is a possible reason why organizational researchers who use SEM rely on tau-equivalent reliability instead of SEM-based reliability coefficients. If they could compute congeneric reliability by simply clicking a mouse, they are likely to use it substantially more frequently. I developed a calculator to overcome this obstacle. The history of reliability coefficients provides a lesson that publishing something namelessly is likely to produce an uninformative or confusing name. I call this calculator RelCalc. It is free to use and distribute.²

RelCalc is a Microsoft Excel spreadsheet consisting of two modules. The first module is designed to help its users choose the right unidimensional reliability coefficient that fits their data. Graham (2006) gave an excellent lecture on choosing between tau-equivalent reliability and congeneric reliability. Following his advice requires readers to fully understand the statistical procedures used to examine the tau-equivalency assumption. As the results of the first section show, this teach-them-how-to-fish approach has seen little effect on how typical organizational researchers use reliability coefficients. This study adopts a give-them-a-fish approach and introduces a program that automates the required statistical procedure.

The first module examines whether the data meet the condition of being parallel or tau-equivalent and calculates three unidimensional reliability coefficients based on the user's input of a covariance matrix. This idea originated from Miles's (2005) suggestion that the maximum likelihood estimation, the most commonly used estimation method in SEM, is an optimization technique that finds a solution to minimize the discrepancy function, and for this, Microsoft Excel offers an optimization tool.

The second module of RelCalc can compute all multidimensional reliability coefficients that were included in this study (i.e., Tables 6, 7, and 8). Unlike the first module, the second module does not estimate the parameters of measurement models; users should copy and paste the parameter estimates that are obtained from an SEM software.

Conclusion

Reliability coefficients should be understood as building blocks of a single system rather than as a collection of completely unrelated methods. Although various formulas have been proposed to estimate the reliability of unidimensional data, they all start from a single formula: the ratio of true score variance to test score variance. Because the variance of a sum is equal to the sum of all elements in the covariance matrix of the components, we are actually discussing the covariance matrix. This study demonstrated the decomposition of covariance matrix for estimating reliability based on SEM models. We do not require a dozen names for reliability coefficients that are seemingly unrelated or a dozen formula expressions that are seemingly unrelated; we need only a single principle that connects all these reliability coefficients.

Appendix A

Earlier Studies on Unidimensional Congeneric Reliability Coefficients

Special case ($k = 2$). Raju (1970) introduced a split-half reliability coefficient for when the relative item lengths are known. Let π_1 denote the ratio of the length of the first item (i.e., λ_1) to the total length (i.e., $\lambda_1 + \lambda_2$), and let π_2 denote the ratio of the length of the second item to the total length. Raju's (1970) original formula ($\tilde{\rho}_{R1970}$) is algebraically equivalent to ρ_{SC} , that is,

$$\tilde{\rho}_{R1970} = \frac{\sigma_{12}}{\pi_1 \pi_2 \sigma_X^2} = \frac{(\lambda_1 + \lambda_2)(\lambda_1 + \lambda_2)\lambda_1 \lambda_2}{\lambda_1 \lambda_2 \sigma_X^2} = \frac{(\lambda_1 + \lambda_2)^2}{\sigma_X^2} = \rho_{SC}.$$

Second, Angoff (1953) and Feldt (1975) independently introduced a split-half reliability coefficient for when the relative item lengths are unknown by restricting $\pi_1 = (\sigma_1^2 + \sigma_{12})/\sigma_X^2$ and $\pi_2 = (\sigma_2^2 + \sigma_{12})/\sigma_X^2 = 1 - \pi_1$. The Angoff-Feldt coefficient ($\tilde{\rho}_{AF}$) is a special case of ρ_{SC} despite their different appearances, that is,

$$\tilde{\rho}_{AF} = \frac{4\sigma_{12}}{\sigma_X^2 - \left(\frac{\sigma_1^2 - \sigma_2^2}{\sigma_X}\right)^2} = \frac{4\sigma_{12}\sigma_X^2}{\sigma_X^4 - \left(\sigma_X^2(\pi_1 - \pi_2)\right)^2} = \frac{4\sigma_{12}}{\left(1 + (\pi_1 - \pi_2)\right)\left(1 - (\pi_1 - \pi_2)\right)\sigma_X^2} = \tilde{\rho}_{R1970}$$

General case. In addition to Jöreskog’s (1971) work, two approaches deserve mention. Let π_i denote the ratio of the length of item i (i.e., λ_i) to the total length (i.e., $\sum_{i=1}^k \lambda_i$), and let σ_{iX} denote $\sum_{j=1}^k \sigma_{ij}$. Raju’s (1977) generalization of his previous formula ($\tilde{\rho}_{R1977}$) is algebraically equivalent to ρ_C . Feldt’s generalization of the Angoff–Feldt coefficient, or classical congeneric coefficient ($\tilde{\rho}_{CC}$; Feldt & Brennan, 1989), is a special case of ρ_C when $\pi_i = \sigma_{iX} / \sigma_X^2$ (Haertel, 2006), that is,

$$\tilde{\rho}_{R1977} = \frac{\sigma_X^2 - \sum_{i=1}^k \sigma_i^2}{\left(1 - \sum_{i=1}^k \pi_i^2\right) \sigma_X^2} = \frac{\sum_{i \neq j} \sum \sigma_{ij}}{\sum_{i \neq j} \sum \pi_i \pi_j \sigma_X^2} = \frac{\left(\sum_{i=1}^k \lambda_i\right) \left(\sum_{j=1}^k \lambda_j\right) \left(\sum_{i \neq j} \sum \lambda_i \lambda_j\right)}{\left(\sum_{i \neq j} \sum \lambda_i \lambda_j\right) \sigma_X^2} = \rho_C, \text{ and}$$

$$\tilde{\rho}_{CC} = \frac{\sigma_X^2 \left(\sigma_X^2 - \sum_{i=1}^k \sigma_i^2\right)}{\sigma_X^4 - \sum_{i=1}^k \left(\sigma_{iX}\right)^2} = \frac{\sigma_X^2 - \sum_{i=1}^k \sigma_i^2}{\left(1 - \sum_{i=1}^k \pi_i^2\right) \sigma_X^2} = \tilde{\rho}_{R1977} = \rho_C$$

Appendix B

Derivations of Multidimensional Reliability Coefficients

Bifactor Model

The definition of reliability. The bifactor model assumes that the observed score of item i consists of four components such that $X_i = \mu_i + \lambda_{iF}F + \lambda_{ip}G_p + e_i$. This study assumes that no specific factors exist to circumvent the controversial issue of specific factors and errors. The general factor F is common to all items. This study fixes the variance of the general factor to unity (i.e., $Var(F) = 1$) without loss of generality to determine the scale. The group factor G_p is common to items that have the same group p . This study also fixes the variances of group factors to unity (i.e., $Var(G_p) = 1$ for all p). When k denotes the number of all items, m denotes the number of groups, and n_p denotes the number of items that have group factor G_p , the relationships $1 < n_p < k$ and $\sum_{p=1}^m n_p = k$ hold. The general factor F , group factor G_p and random error e_i are defined to be independent of each other. This study further assumes that group factors are independent of each other (i.e., $Cov(G_p, G_q) = 0$ for all p, q) and that errors are independent of each other (i.e., $Cov(e_i, e_j) = 0$ for all i, j).

McDonald (1985, 1999) introduced two reliability definitions. Let $\sigma_{General}^2$ denote the sum of the variances that are due to the general factor, and let σ_{Group}^2 denote the sum of the variances that are due to the group factors: $\sigma_{General}^2 = \left(\sum_{i=1}^k \lambda_{iF}\right)^2$, $\sigma_{Group}^2 = \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip}\right)^2$. The unidimensional model is nested within the bifactor model. Reliability is defined as the ratio of the sum of $\sigma_{General}^2$ and σ_{Group}^2 to the test score variance, and hierarchical omega is defined as the ratio $\sigma_{General}^2$ to the test score variance: $\rho_{XX'} = \frac{\sigma_{General}^2 + \sigma_{Group}^2}{\sigma_X^2}$, and $\omega_H = \frac{\sigma_{General}^2}{\sigma_X^2}$.

Notations. Before proceeding with the derivation, let us define matrix notations used in this study. A matrix may consist of partitioned matrices or submatrices. For example, the matrix \mathbf{A} in Table 10 can be partitioned into 4 3×3 submatrices such that $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, where

$$\mathbf{A}_{11} = \mathbf{A}_{22} = \begin{bmatrix} 1 & .5 & .5 \\ .5 & 1 & .5 \\ .5 & .5 & 1 \end{bmatrix}, \mathbf{A}_{12} = \mathbf{A}_{21} = \begin{bmatrix} .3 & .3 & .3 \\ .3 & .3 & .3 \\ .3 & .3 & .3 \end{bmatrix}. \text{ When a matrix consists of } m \text{ row}$$

submatrices and n column submatrices, this study says that the matrix is an $m * n$ block matrix. For example, the matrix \mathbf{A} is a $2 * 2$ block matrix. A square matrix $\| a_{ii}, a_{ij} \|$ is a matrix whose (i, i) entry (i.e., diagonal element) is a_{ii} and whose (i, j) entry ($i \neq j$, off-diagonal element) is a_{ij} . For example, \mathbf{A}_{11} is a $3 \times 3 \| 1, .5 \|$ matrix. A square matrix $\| a_{ij} \|$ is a matrix whose (i, j) entry is a_{ij} . For example, \mathbf{A}_{12} is a $3 \times 3 \| .3 \|$ matrix. A square matrix $\| \mathbf{A}_{ii}, \mathbf{A}_{ij} \|$ is a matrix whose (i, i) submatrix (i.e., diagonal submatrix) is \mathbf{A}_{ii} and whose (i, j) submatrix ($i \neq j$, off-diagonal submatrix) is \mathbf{A}_{ij} . For example, Table 6 shows a matrix decomposition of $\| \| 1, .5 \|, \| .3 \| \| = \| .3 \| + \| \| .2 \|, \| 0 \| \| + \| .5, 0 \|$, where $\| \| 1, .5 \|, \| .3 \| \|$ and $\| \| .2 \|, \| 0 \| \|$ are $2 * 2$ block matrices and all submatrices are 3×3 matrices.

Bifactor parallel model. The bifactor parallel model requires four restrictions: the general factor loadings of all items are restricted to be equal to each other (i.e., $\lambda_{iF} = \lambda_F$ for all i), the group factor loadings of all items are restricted to be equal to each other (i.e., $\lambda_{iG_p} = \lambda_G$ for all i, p), the error variances of all items are restricted to be equal to each other (i.e., $\sigma_{e_i}^2 = \sigma_e^2$ for all i), and all group factors are restricted to have equal numbers of items (i.e., $n_p = n$ for all $p, k = mn$). The variance of item i is $\lambda_F^2 + \lambda_G^2 + \sigma_e^2$, and the covariance between item i and item j is $\lambda_F^2 + \lambda_G^2$ for $i \neq j$ if i and j have the same group factor and is λ_F^2 for $i \neq j$ if i and j have different group factors. Let c denote $\lambda_F^2 + \lambda_G^2 + \sigma_e^2$, let $\bar{\rho}_s$ denote $(\lambda_F^2 + \lambda_G^2)/c$, and let $\bar{\rho}_d$ denote λ_F^2/c . $\bar{\rho}_s$ equals the average product-moment correlation between items that have the same group factor, and $\bar{\rho}_d$ equals the average product-moment correlation between items that have different group factors. The covariance matrix of inter-item scores will be $\Sigma_{mp} = c \| \| 1, \bar{\rho}_s \|, \| \bar{\rho}_d \| \| = c \| \bar{\rho}_d \| + c \| \| \bar{\rho}_s - \bar{\rho}_d \|, \| 0 \| \| + c \| 1 - \bar{\rho}_s, 0 \|$.

The sum of the $k \times k$ matrix $c \| \bar{\rho}_d \|$ represents $\sigma_{General}^2$ and equals $ck^2\bar{\rho}_d$. The matrix $c \| \| \bar{\rho}_s - \bar{\rho}_d \|, \| 0 \| \|$ is an $m * m$ block matrix, which consists of $n \times n$ diagonal submatrices $\| \bar{\rho}_s - \bar{\rho}_d \|$ and $n \times n$ off-diagonal submatrices $\| 0 \|$. Its sum represents σ_{Group}^2 and equals $cnk(\bar{\rho}_s - \bar{\rho}_d)$. The sum of the $k \times k$ matrix $c \| 1 - \bar{\rho}_s, 0 \|$ represents σ_E^2 and equals $ck(1 - \bar{\rho}_s)$. The bifactor parallel reliability coefficient and parallel form of hierarchical omega are calculated as follows:

$$\rho_{BP} = \frac{\sigma_{General}^2 + \sigma_{Group}^2}{\sigma_X^2} = \frac{k^2\bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d)}{k^2\bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d) + k(1 - \bar{\rho}_s)}, \text{ and}$$

$$\omega_{HP} = \frac{\sigma_{General}^2}{\sigma_X^2} = \frac{k^2\bar{\rho}_d}{k^2\bar{\rho}_d + nk(\bar{\rho}_s - \bar{\rho}_d) + k(1 - \bar{\rho}_s)}$$

Bifactor tau-equivalent model. The bifactor parallel model requires three restrictions: the general factor loadings of all items are restricted to be equal to each other (i.e., $\lambda_{iF} = \lambda_F$ for all i), the group factor loadings of all items are restricted to be equal to each other (i.e., $\lambda_{iG_p} = \lambda_G$ for all i, p), and all group factors are restricted to have an equal number of items (i.e., $n_p = n$ for all p , and $k = mn$). The bifactor parallel model is nested within the bifactor tau-equivalent model, and the only difference between the models is the equality restriction of error variances. The variance of item i is $\lambda_F^2 + \lambda_G^2 + \sigma_{e_i}^2$, and the covariance between item i and item j is $\lambda_F^2 + \lambda_G^2$ for $i \neq j$ if i and j have the same group factor and is λ_F^2 for $i \neq j$ if i and j have different group factors. Let $\bar{\sigma}_s$ denote $\lambda_F^2 + \lambda_G^2$, and let $\bar{\sigma}_d$ denote λ_F^2 . $\bar{\sigma}_s$ equals the average covariance between items that have the same group factor, and $\bar{\sigma}_d$ equals the average covariance between items that have different group factors. The covariance matrix of the interitem scores will be $\Sigma_{mq} = \| \| \bar{\sigma}_s + \sigma_{e_i}^2, \bar{\sigma}_s \|, \| \bar{\sigma}_d \| \| = \| \bar{\sigma}_d \| + \| \| \bar{\sigma}_s - \bar{\sigma}_d \|, \| 0 \| \| + \| \sigma_{e_i}^2, 0 \|$.

The sum of the $k \times k$ matrix $\| \bar{\sigma}_d \|$ represents $\sigma_{General}^2$ and equals $k^2\bar{\sigma}_d$. The matrix $\| \| \bar{\sigma}_s - \bar{\sigma}_d \|, \| 0 \| \|$ is an $m * m$ block matrix, which consists of $n \times n$ diagonal submatrices $\| \bar{\sigma}_s - \bar{\sigma}_d \|$ and $n \times n$ off-diagonal submatrices $\| 0 \|$. Its sum represents σ_{Group}^2 and equals

$nk(\bar{\sigma}_s - \bar{\sigma}_d)$. The derived formulas of bifactor tau-equivalent reliability (ρ_{BT}) and the tau-equivalent form of hierarchical omega are as follows:

$$\rho_{BT} = \frac{\sigma_{General}^2 + \sigma_{Group}^2}{\sigma_X^2} = \frac{k^2\bar{\sigma}_d + nk(\bar{\sigma}_s - \bar{\sigma}_d)}{\sigma_X^2}, \text{ and}$$

$$\omega_{HT} = \frac{\sigma_{General}^2}{\sigma_X^2} = \frac{k^2\bar{\sigma}_d}{\sigma_X^2}.$$

Rajaratnam, Cronbach, and Gleser (1965) applied generalizability theory to derive stratified alpha, where the term *stratified* represents multidimensional. Let σ_p^2 denote the sum of all elements of the covariance matrix of group (or strata in the original article) p , and let α_p denote the α that is obtained from group p . The formula for stratified alpha ($\tilde{\rho}_{BT}$) is algebraically equivalent to ρ_{BT} (Rae, 2008):

$$\tilde{\rho}_{BT} = 1 - \frac{\sum_{p=1}^m \sigma_p^2(1 - \alpha_p)}{\sigma_X^2} = \frac{\sigma_X^2 - \sum_{p=1}^m \sigma_p^2 + \sum_{p=1}^m \sigma_p^2 \alpha_p}{\sigma_X^2} = \frac{m(m-1)n^2\bar{\sigma}_d + \sum_{p=1}^m \sigma_p^2 \left(\frac{n^2\bar{\sigma}_s}{\sigma_p^2} \right)}{\sigma_X^2}$$

$$= \frac{m(m-1)n^2\bar{\sigma}_d + mn^2\bar{\sigma}_s}{\sigma_X^2} = \frac{k^2\bar{\sigma}_d + nk(\bar{\sigma}_s - \bar{\sigma}_d)}{\sigma_X^2} = \rho_{BT}$$

Bifactor congeneric model. The bifactor congeneric model is an unrestricted base model. The bifactor tau-equivalent model is nested within the bifactor congeneric model. The variance of item i is $\lambda_{iF}^2 + \lambda_{ip}^2 + \sigma_{ei}^2$, and the covariance between item i and item j is $\lambda_{iF}\lambda_{jF} + \lambda_{ip}\lambda_{jp}$ for $i \neq j$ if i and j have the same group factor, and $\lambda_{iF}\lambda_{jF}$ for $i \neq j$ if i and j have different group factors. The covariance matrix of interitem scores is $\Sigma_{mq} = \|\lambda_{iF}\lambda_{jF}\| + \|\|\lambda_{ip}\lambda_{jp}\|, \|0\|\| + \|\sigma_{ei}^2, 0\|$.

The sum of the $k \times k$ matrix $\|\lambda_{iF}\lambda_{jF}\|$ represents $\sigma_{General}^2$ and equals $\sum_{i=1}^k \sum_{j=1}^k \lambda_{iF}\lambda_{jF} = \left(\sum_{i=1}^k \lambda_{iF}\right)^2$. The sum of the $m * m$ block matrix $\|\|\lambda_{ip}\lambda_{jp}\|, \|0\|\|$ represents σ_{Group}^2 and equals $\sum_{p=1}^m \left(\sum_{i=1}^k \sum_{j=1}^k \lambda_{ip}\lambda_{jp}\right) = \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip}\right)^2$. The derived formula of bifactor congeneric reliability (ρ_{BC}) is algebraically equivalent to McDonald’s original version ($\tilde{\rho}_{BC}$), and the derived formula of hierarchical omega (ω_{HBF}) is equivalent to the conventional version ($\tilde{\omega}_H$) (Cho & Kim, 2015):

$$\rho_{BC} = \frac{\sigma_{General}^2 + \sigma_{Group}^2}{\sigma_X^2} = \frac{\left(\sum_{i=1}^k \lambda_{iF}\right)^2 + \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip}\right)^2}{\sigma_X^2}, \tilde{\rho}_{BC} = \frac{\mathbf{1}'\mathbf{c}\mathbf{c}'\mathbf{1} + \mathbf{1}'\mathbf{A}\mathbf{A}'\mathbf{1}}{\sigma_X^2},$$

$$\omega_{HBF} = \frac{\sigma_{General}^2}{\sigma_X^2} = \frac{\left(\sum_{i=1}^k \lambda_{iF}\right)^2}{\sigma_X^2}, \text{ and } \tilde{\omega}_H = \frac{\mathbf{1}'\mathbf{c}\mathbf{c}'\mathbf{1}}{\sigma_X^2}.$$

Second-order factor model

The definition of reliability. The second-order factor model assumes that the observed score of item i consists of three components such that $Y_i = \mu_i + \lambda_{ip}O_p + e_i$. The first-order factor O_p is determined by the second-order factor Q and the disturbance ζ_p such that $O_p = \gamma_p Q + \zeta_p$. Specifically, $Y_i = \mu_i + \lambda_{ip}(\gamma_p Q + \zeta_p) + e_i$. The variances of a second-order factor and first-order factors are fixed to unity (i.e., $Var(Q) = 1$, $Var(O_p) = 1$ for all p) to determine the scale. The variance of the disturbance ζ_p is $1 - \gamma_p^2$ ($\because Var(O_p) = Var(\gamma_p Q + \zeta_p) = \gamma_p^2 + Var(\zeta_p) = 1$). The second-order

factor Q , disturbance ζ_p , and errors e_i are defined to be independent of each other. This study further assumes that disturbances are independent of each other (i.e., $Cov(\zeta_p, \zeta_q) = 0$ for all p, q), and error variances are independent of each other (i.e., $Cov(e_i, e_j) = 0$ for all i, j). Let k denote the number of all items, let m denote the number of first-order factors, and let n_p denote the number of items that have first-order factor O_p .

Let $\sigma_{Second-order}^2$ denote the sum of the variances that are due to the second-order factor, and let $\sigma_{Disturbance}^2$ denote the sum of the variances that are due to disturbances. Reliability is defined as the ratio of the sum of $\sigma_{Second-order}^2$ and $\sigma_{Disturbance}^2$ to the test score variance, and hierarchical omega is defined as the ratio $\sigma_{Second-order}^2$ to the test score variance: $\rho_{XX'} = \frac{\sigma_{Second-order}^2 + \sigma_{Disturbance}^2}{\sigma_X^2}$, and $\omega_H = \frac{\sigma_{Second-order}^2}{\sigma_X^2}$.

Second-order factor parallel model. The second-order factor parallel model requires four restrictions: the path coefficients of the second-order factor on all first-order factors are restricted to be equal to each other (i.e., $\gamma_p = \gamma$ for all p), the first-order factor loadings of all items are restricted to be equal to each other (i.e., $\lambda_i = \lambda$ for all i), all first-order factors are restricted to have equal numbers of items (i.e., $n_p = n$ for all $p, k = mn$), and the errors of all items are restricted to be equal to each other (i.e., $\sigma_{e_i}^2 = \sigma_e^2$ for all i).

The variance of item i is $\lambda^2\gamma^2 + \lambda^2(1 - \gamma^2) + \sigma_e^2 = \lambda^2 + \sigma_e^2$, and the covariance between item i and item j is $\lambda^2\gamma^2 + \lambda^2(1 - \gamma^2) = \lambda^2$ for $i \neq j$ if i and j have the same first-order factor, and $\lambda^2\gamma^2$ for $i \neq j$ if i and j have different first-order factors. Let c denote $\lambda^2 + \sigma_e^2$, let $\bar{\rho}_s$ denote λ^2/c , and let $\bar{\rho}_d$ denote $\lambda^2\gamma^2/c$. The covariance matrix of interitem scores will be $\Sigma_{sofp} = c \begin{bmatrix} 1 & \bar{\rho}_s & \dots & \bar{\rho}_d \\ \bar{\rho}_s & 1 & \dots & \bar{\rho}_d \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\rho}_d & \bar{\rho}_d & \dots & 1 \end{bmatrix} = c \begin{bmatrix} \bar{\rho}_d & \bar{\rho}_s - \bar{\rho}_d & \dots & 0 \\ \bar{\rho}_s - \bar{\rho}_d & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 - \bar{\rho}_s \end{bmatrix}$, which is identical to that of the bifactor parallel model.

Second-order factor tau-equivalent model. The second-order factor tau-equivalent model requires three restrictions that are identical to those of the second-order factor parallel model, except for the equality constraint of the error variances (i.e., $\sigma_{e_i}^2 \neq \sigma_{e_j}^2$). The variance of item i is $\lambda^2 + \sigma_{e_i}^2$, and the covariance between item i and item j is λ^2 for $i \neq j$ if i and j have the same first-order factor and is $\lambda^2\gamma^2$ for $i \neq j$ if i and j have different first-order factors. Let $\bar{\sigma}_s$ denote λ^2 , and let $\bar{\sigma}_d$ denote $\lambda^2\gamma^2$. The covariance matrix of the interitem scores will be $\Sigma_{mq} = \begin{bmatrix} \bar{\sigma}_s + \sigma_{e_i}^2 & \bar{\sigma}_s & \dots & \bar{\sigma}_d \\ \bar{\sigma}_s & \bar{\sigma}_s + \sigma_{e_j}^2 & \dots & \bar{\sigma}_d \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\sigma}_d & \bar{\sigma}_d & \dots & \bar{\sigma}_s + \sigma_{e_k}^2 \end{bmatrix} = \begin{bmatrix} \bar{\sigma}_d & \bar{\sigma}_s - \bar{\sigma}_d & \dots & 0 \\ \bar{\sigma}_s - \bar{\sigma}_d & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 - \bar{\rho}_s \end{bmatrix}$, which is identical to that of the bifactor tau-equivalent model.

Second-order factor congeneric model. The bifactor congeneric model is an unrestricted base model. The variance of item i is $\lambda_{ip}^2 + \sigma_{e_i}^2$, and the covariance between item i and item j is $\lambda_{ip}\lambda_{jp}$ for $i \neq j$ if i and j have the same first-order factor and is $\lambda_{ip}\lambda_{jq}\gamma_p\gamma_q$ for $i \neq j$ if i and j have different first-order factors. The covariance matrix of the interitem scores will be $\Sigma_{sofc} = \begin{bmatrix} \lambda_{i1}^2 + \sigma_{e_i}^2 & \lambda_{i1}\lambda_{j1} & \dots & \lambda_{i1}\lambda_{jq}\gamma_p\gamma_q \\ \lambda_{i1}\lambda_{j1} & \lambda_{j1}^2 + \sigma_{e_j}^2 & \dots & \lambda_{j1}\lambda_{jq}\gamma_p\gamma_q \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{i1}\lambda_{jq}\gamma_p\gamma_q & \lambda_{j1}\lambda_{jq}\gamma_p\gamma_q & \dots & \lambda_{jq}^2 + \sigma_{e_k}^2 \end{bmatrix} = \begin{bmatrix} \lambda_{i1}^2 + \sigma_{e_i}^2 & \lambda_{i1}\lambda_{j1} & \dots & \lambda_{i1}\lambda_{jq}\gamma_p\gamma_q \\ \lambda_{i1}\lambda_{j1} & \lambda_{j1}^2 + \sigma_{e_j}^2 & \dots & \lambda_{j1}\lambda_{jq}\gamma_p\gamma_q \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{i1}\lambda_{jq}\gamma_p\gamma_q & \lambda_{j1}\lambda_{jq}\gamma_p\gamma_q & \dots & \lambda_{jq}^2 + \sigma_{e_k}^2 \end{bmatrix}$. The sum of the $k \times k$ matrix $\begin{bmatrix} \lambda_{i1}\lambda_{jq}\gamma_p\gamma_q & \dots & \lambda_{jq}^2 + \sigma_{e_k}^2 \end{bmatrix}$ represents $\sigma_{Second-order}^2$ and equals $\sum_{p=1}^m \sum_{q=1}^k \sum_{i=1}^k \sum_{j=1}^k \lambda_{ip}\lambda_{jq}\gamma_p\gamma_q = \left(\sum_{p=1}^m \sum_{j=1}^k \lambda_{ip}\gamma_p\right)^2$. The sum of the $m \times m$ block matrix $\begin{bmatrix} (1 - \gamma_p^2)\lambda_{ip}\lambda_{jp} & \dots & 0 \end{bmatrix}$ represents $\sigma_{Disturbance}^2$ and equals $\sum_{p=1}^m \sum_{i=1}^k \sum_{j=1}^k (1 - \gamma_p^2)\lambda_{ip}\lambda_{jp} = \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip}(1 - \gamma_p^2)^{1/2}\right)^2$. The derived formulas of second-order factor reliability (ρ_{SOF}) and second-order factor hierarchical omega (ω_{HSOF}) are as follows:

$$\rho_{SOF} = \frac{\sigma_{Second-order}^2 + \sigma_{Disturbance}^2}{\sigma_X^2} = \frac{\left(\sum_{p=1}^m \sum_{j=1}^k \lambda_{ip}\gamma_p\right)^2 + \sum_{p=1}^m \left(\sum_{i=1}^k \lambda_{ip}(1 - \gamma_p^2)^{1/2}\right)^2}{\sigma_X^2}, \text{ and}$$

$$\omega_{HSOF} = \frac{\sigma_{Second-order}^2}{\sigma_X^2} = \frac{\left(\sum_{p=1}^m \sum_{j=1}^k \lambda_{ip} \gamma_p\right)^2}{\sigma_X^2}.$$

Correlated factor model. The correlated factor model assumes that the observed score of item i consists of three components such that $X_i = \mu_i + \lambda_{ip}D_p + e_i$. Factors are free to be correlated with each other. Let ϕ_{pq} denote $Cor(D_p, D_q)$. This study fixes the variance of the factors to unity (i.e., $Var(D_p) = 1$) without loss of generality to determine the scale. The factor D_p and error e_i are defined to be independent of each other. This study further assumes that errors are independent of each other (i.e., $Cov(e_i, e_j) = 0$ for all i, j). The variance of item i is $\lambda_{ip}^2 + \sigma_{e_i}^2 = \lambda_{ip}^2 \phi_{pp} + \sigma_{e_i}^2$, and the covariance between item i and item j is $\lambda_{ip}\lambda_{jp} = \lambda_{ip}\lambda_{jp}\phi_{pp}$ for $i \neq j$ if i and j have the same factor D_p and is $\lambda_{ip}\lambda_{jq}\phi_{pq}$ for $i \neq j$ if i and j have different group factors. The covariance matrix of the interitem scores will be $\Sigma_{cf} = \|\lambda_{ip}^2 \phi_{pp} + \sigma_{e_i}^2, \lambda_{ip}\lambda_{jq}\phi_{pq}\| = \|\lambda_{ip}\lambda_{jq}\phi_{pq}\| + \|\sigma_{e_i}^2, 0\|$. The sum of the $k \times k$ matrix $\|\lambda_{ip}\lambda_{jq}\phi_{pq}\|$ equals $\sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^k \sum_{j=1}^k \phi_{pq}\lambda_{ip}\lambda_{jq} = \sum_{p=1}^m \sum_{q=1}^m \phi_{pq} \left(\sum_{i=1}^k \lambda_{ip}\right) \left(\sum_{i=1}^k \lambda_{iq}\right)$. The reliability is defined as one minus the ratio of the sum of error variances to the sum of test score variances: $\rho_{XX'} = 1 - \frac{\sigma_{e_i}^2}{\sigma_X^2}$. The derived formula of the correlated factor reliability is as follows:

$$\rho_{CF} = \frac{\sum_{p=1}^m \sum_{q=1}^m \phi_{pq} \left(\sum_{i=1}^k \lambda_{ip}\right) \left(\sum_{i=1}^k \lambda_{iq}\right)}{\sigma_X^2}.$$

Acknowledgments

I am deeply grateful to Adam Meade, associate editor, and three anonymous reviewers for their invaluable guidance and constructive comments. I also appreciate helpful support from Kyung Su Liu, Seonghoon Kim, Yanyun Yang, Richard Zinbarg, Peter M. Bentler, and Jiheon Kim. All errors are my sole responsibility.

Declaration of Conflicting Interests

The author 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 present research has been conducted by the Research Grant of Kwangwoon University in 2015.

Notes

1. I am indebted to anonymous reviewers for this idea.
2. Additional details are available at <http://relcalc.blogspot.com/2016/05/how-to-obtain-and-use-relcalc.html>.

References

Angoff, W. H. (1953). Test reliability and effective test length. *Psychometrika*, 18(1), 1-14.

Armor, D. J. (1974). Theta reliability and factor scaling. In H. L. Costner (Ed.), *Sociological methodology* (pp. 17-50). San Francisco, CA: Jossey-Bass.

Bentler, P. M. (2006). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software.

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, 74(1), 137-143.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: John Wiley.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3(3), 296-322.

- Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*(4), 796-846.
- Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well known but poorly understood. *Organizational Research Methods, 18*(2), 207-230.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.
- Cronbach, L. J. (1978). Citation classics. *Current Contents, 13*, 263.
- Cronbach, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*(3), 391-418.
- Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399-412.
- Falk, C. F., & Savalei, V. (2011). The relationship between unstandardized and standardized alpha, true reliability, and the underlying measurement model. *Journal of Personality Assessment, 93*(5), 445-453.
- Feldt, L. S. (1975). Estimation of the reliability of a test divided into two parts of unequal length. *Psychometrika, 40*(4), 557-561.
- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). New York, NY: American Council on Education and Macmillan.
- Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd Ltd.
- Flanagan, J. C. (1937). A proposed procedure for increasing the efficiency of objective tests. *Journal of Educational Psychology, 28*(1), 17-21.
- Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*(6), 930-944.
- 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*(4), 827-838.
- Green, S. B., & Yang, Y. (2009a). Commentary on coefficient alpha: A cautionary tale. *Psychometrika, 74*(1), 121-135.
- Green, S. B., & Yang, Y. (2009b). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika, 74*(1), 155-167.
- Green, S. B., & Yang, Y. (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficients. *Educational Measurement: Issues and Practice, 34*(4), 14-20.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10*(4), 255-282.
- Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65-110). Westport, CT: American Council on Education and Praeger.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. Upper Saddle River, NJ: Pearson.
- Hattie, J. (1985). Assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9*(2), 139-164.
- Hoyt, C. (1941). Test reliability estimated by analysis of variance. *Psychometrika, 6*(3), 153-160.
- Hunt, T. D., & Bentler, P. M. (2015). Quantile lower bounds to reliability based on locally optimal splits. *Psychometrika, 80*(1), 182-195.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 36*(2), 109-133.
- Kamata, A., Turhan, A., & Darandari, E. (2003, April). *Estimating reliability for multidimensional composite scale scores*. Paper presented at the annual meeting of American Educational Research Association, Chicago, IL.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151-160.

- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria. *Organizational Research Methods, 9*(2), 202-220.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Lucke, J. F. (2005). "Rassling the hog": The influence of correlated item error on internal consistency, classical reliability and congeneric reliability. *Applied Psychological Measurement, 29*(2), 106-125.
- McDonald, R. P. (1978). Generalizability in factorable domains: "Domain validity and generalizability": 1. *Educational and Psychological Measurement, 38*(1), 75-79.
- McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Lawrence Erlbaum.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- Miles, J. N. V. (2005). Confirmatory factor analysis using Microsoft Excel. *Behavior Research Methods, 37*(4), 672-676.
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling, 2*(3), 255-273.
- Mosier, C. I. (1941). A short cut in the estimation of split-halves coefficients. *Educational and Psychological Measurement, 1*(1), 407-408.
- Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*(1), 1-13.
- Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*(3), 343-355.
- Padilla, M. A., & Divers, J. (2013). Coefficient omega bootstrap confidence intervals: Nonnormal distributions. *Educational and Psychological Measurement, 73*(6), 956-972.
- Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity. *European Health Psychologist, 16*(2), 56-69.
- Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology, 98*(1), 194-198.
- Rae, G. (2008). A note on using alpha and stratified alpha to estimate the reliability of a test composed of item parcels. *British Journal of Mathematical and Statistical Psychology, 61*(2), 515-525.
- Rajaratnam, N., Cronbach, L. J., & Gleser, G. C. (1965). Generalizability of stratified-parallel tests. *Psychometrika, 30*(1), 39-56.
- Raju, N. S. (1970). New formula for estimating total test reliability from parts of unequal length. *Proceedings of the 78th Annual Convention of APA, 5*, 143-144.
- Raju, N. S. (1977). A generalization of coefficient alpha. *Psychometrika, 42*(4), 549-565.
- Raykov, T. (2012). Scale construction and development using structural equation modeling. In R. H. Hoyle (Eds.), *Handbook of structural equation modeling* (pp. 472-492). New York, NY: Guilford.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696.
- Revelle, W. (1979). Hierarchical cluster-analysis and the internal structure of tests. *Multivariate Behavioral Research, 14*(1), 57-74.
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega and the glb: Comments on Sijtsma. *Psychometrika, 74*(1), 145-154.
- Rulon, P. J. (1939). A simplified procedure for determining the reliability of a test by split-halves. *Harvard Educational Review, 9*, 99-103.
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*(1), 53-61.
- Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods, 8*(2), 206-224.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350-353.

- Sijtsma, K. (2009a). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107-120.
- Sijtsma, K. (2009b). Reliability beyond theory and into practice. *Psychometrika*, *74*(1), 169-173.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal Psychology*, *3*(3), 271-295.
- Tang, W., & Cui, Y. (2012, April). *A simulation study for comparing three lower bounds to reliability*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, Canada.
- Ten Berge, J. M. F., & Zegers, F. E. (1978). A series of lower bounds to the reliability of a test. *Psychometrika*, *43*, 575-579.
- van der Ark, L. A., van der Palm, D. W., & Sijtsma, K. (2011). A latent class approach to estimating test-score reliability. *Applied Psychological Measurement*, *35*(5), 380-392.
- Werts, C. E., Rock, D. R., Linn, R. L., & Jöreskog, K. G. (1978). A general method of estimating the reliability of a composite. *Educational and Psychological Measurement*, *38*(4), 933-938.
- Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, *29*(4), 377-392.
- Yung, Y. F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*(2), 113-128.
- Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating ω_h for structures containing two group factors: Perils and prospects. *Applied Psychological Measurement*, *31*(2), 135-157.
- 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 alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123-133.
- Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ω_h . *Applied Psychological Measurement*, *30*(2), 121-144.

Author Biography

Eunseong Cho is a professor of marketing in the College of Business Administration, Kwangwoon University, Republic of Korea. He received his PhD from the Korea Advanced Institute of Science and Technology. His current methodological research focuses on reliability and validity.