UC Berkeley
UC Berkeley Previously Published Works

Title
A Simulation Study on the Performance of Different Reliability Estimation Methods

Permalink
https://escholarship.org/uc/item/44z85306

Journal
Educational and Psychological Measurement, 81(6)

ISSN
0013-1644

Authors
Edwards, Ashley A
Joyner, Keanan J
Schatschneider, Christopher

Publication Date
2021-12-01

DOI
10.1177/0013164421994184

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Article

Educational and Psychological Measurement
2021, Vol. 81(6) 1089–1117
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0013164421994184
journals.sagepub.com/home/epm

A Simulation Study on the Performance of Different Reliability Estimation Methods

Ashley A. Edwards[1], **Keanan J. Joyner**[1] and **Christopher Schatschneider**[1]

Abstract

The accuracy of certain internal consistency estimators have been questioned in recent years. The present study tests the accuracy of six reliability estimators (Cronbach's alpha, omega, omega hierarchical, Revelle's omega, and greatest lower bound) in 140 simulated conditions of unidimensional continuous data with uncorrelated errors with varying sample sizes, number of items, population reliabilities, and factor loadings. Estimators that have been proposed to replace alpha were compared with the performance of alpha as well as to each other. Estimates of reliability were shown to be affected by sample size, degree of violation of tau equivalence, population reliability, and number of items in a scale. Under the conditions simulated here, estimates quantified by alpha and omega yielded the most accurate reflection of population reliability values. A follow-up regression comparing alpha and omega revealed alpha to be more sensitive to degree of violation of tau equivalence, whereas omega was affected greater by sample size and number of items, especially when population reliability was low.

Keywords

reliability, simulation, internal consistency

[1]Florida State University, Tallahassee, FL, USA

Corresponding Author:
Ashley A. Edwards, Department of Psychology, Florida State University, 1107 West Call Street, Tallahassee, FL 32306, USA.
Email: aedwards@psy.fsu.edu

Introduction

Psychological sciences have long been interested in the psychometric properties of tests and scales. While it is certainly true that not enough attention is paid in practice to psychometrics (e.g., Sijtsma, 2012), it is the bedrock of psychological science. One aspect of psychometrics concerns the reliability of scales—that is, the degree to which a scale or composite is consistent, internally or temporally, in measuring variation in a sample of subjects. The main reason reliability is critically important is that it is a necessary condition for validity (Downing, 2003; Feldt & Brennan, 1989). The ceiling on the meaningful associations between two variables is also determined by the reliability of the individual scales. However, there are now myriad estimation methods to choose from to estimate internal consistency, and different situations affect the estimates yielded by each of them (Green & Yang, 2009; McNeish, 2018). Thus, accurate estimation of reliability is an integral part of the foundation of psychological sciences. The current work presents a simulation study to understand the effects different relevant factors (e.g., sample size, number of items, differential loadings) have on accurately estimating internal consistency using Cronbach's alpha, and the composite reliability estimators omega, omega hierarchical, Revelle's omega, and greatest lower bound.

Since Cronbach (1951) popularized alpha, there has been a long-standing reliance on Cronbach's alpha for estimation of internal consistency (Hogan et al., 2000). One reason for this is that alpha is very practical to compute—it does not use factor analytic techniques, which were relatively inaccessible until more modern computing became widely available. Despite early (Jackson & Agunwamba, 1977; McDonald, 1970) and continued (Crutzen & Peters, 2015; Dunn et al., 2014; Geldhof et al., 2014; Green & Yang, 2009; Teo & Fan, 2013) challenges to the use of alpha, its use has remained undeterred in psychological science (McNeish, 2018).

However, when key factors affecting the estimation of alpha are violated, the estimate produced is purported to reflect an underestimate of the true population value. Some of the most commonly encountered factors affecting estimates produced by alpha are outlined by McNeish (2018). First, the alpha estimate can suffer if tau equivalence is not met—that is, if the loadings for all items in the scale are not equal. The impact of tau nonequivalence on alpha was demonstrated mathematically by Raykov (1997). In the article, Raykov had suggested that there may be times where the use of latent variable modeling and factor loadings to estimate reliability might be warranted. Second, items should be continuously distributed, and all have normal distributions; if not, the interitem Pearson correlations that alpha is based on will be biased downward. Third, alpha does not model the error term covariances (e.g., greatest lower bound), and therefore, may be affected if items have correlated errors. This can actually upward bias the alpha estimate. Last, if the scale is not truly unidimensional, the alpha estimate will be downward biased. However, others have argued that these factors do not affect alpha to the degree that it should become obsolete (Raykov & Marcoulides, 2019). These authors mainly argue that the factors cited above are actually testable, and when met, alpha provides a good estimation of the lower bound of reliability of a unidimensional measure (consistent with early work by Novick & Lewis, 1967).

If these factors affecting accuracy of internal consistency are commonly encountered, several alternatives for this purpose exist. While there are more than 30 different methods for estimating internal consistency (Hattie, 1985), we will be focusing on omega, omega hierarchical, Revelle's omega, and greatest lower bound in the current work (see Table 1 for formulas). *Omega* (McDonald, 1970, 1999) is calculated through variance decomposition, that is, by comparing the amount of variance explained by items in a scale versus the total amount of variance in the scale. *Omega hierarchical* (Zinbarg et al., 2006) is an extension of this process that is designed for scales that may include subfactors in addition to a general factor and attempts to capture the reliability of the higher order general factor. *Revelle's omega* (Revelle, 2019) treats the scale as unidimensional, but computes subfactors that explain residual covariance and uses the Schimd–Leiman rotation (Schmid & Leiman, 1957) to force orthogonality between these subfactors with the general factor, and then estimates reliability from the loadings onto the general factor. *Greatest lower bound* (GLB; Jackson & Agunwamba, 1977) is meant to create a lower bound of reliability by computing reliability through computing the covariance of true score estimates and also covariance of the error terms, and finding the maximal values for the error term, thereby showing the lowest reliability estimate.

While these *estimators* (e.g., the mathematical equation that is applied to data to produce a reliability estimate in a given sample) have fewer common factors that have consequences for the accuracy of estimation compared with alpha in their calculation of the same intended *estimand* (e.g., the quantity representing internal consistency), they are more computationally intensive than alpha, requiring factor analytic techniques to compute. Despite the comparatively fewer factors that can produce bias in the calculation of omega and GLB, it is unclear if the *estimates* produced are substantially better than those produced by alpha to offset the computational cost (manifested as lack of widespread availability in commonly used statistics programs in applied settings). The goal of the current work was to explore the degree to which internal consistency estimates produced by these composite reliability estimators are an improvement over alpha under conditions where tau equivalency is not met and how much the difference between alpha and the composite reliabilities is affected by sample size.

Factors That Affect Reliability Estimates

While there appears to be disagreement about the degree of impact on the estimate by degree of violation of the above-referenced factors that affect the accuracy of the estimate produced by alpha, extant simulation work tested some of these empirically (Green & Yang, 2009). The first of these, *tau equivalence*, refers to items in the scale or composite having equivalent loadings. Green and Yang demonstrated this bias through varying the factor loadings for their simulation, and this bias indeed did

Table 1. Estimator Formulas.

Estimator	Formula	Variables
Cronbach's alpha	$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_X^2}\right)$	Where k is the number of items, s_i^2 is the variance of individual item i, where i = 1, . . ., k, and s_X^2 is the variance for all items on the scale
Omega	$\omega = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \sum_{i=1}^{k} \theta_{ii}}$	Where λ_i is the factor loading (not necessarily standardized) for the ith item on the scale, θ_{ii} is the error variance for the ith item, and k is the number of items on the scale
Omega hierarchical	$\omega_H = \frac{\left(\sum_{i=1}^{k} \lambda_{gi}\right)^2}{V_X}$	Where λ_{gi} is the loading of the ith item on the general factor, k is the total number of items, V_X is the total variance after rotation, which is equal to the sum of each element of the sample Pearson (or polychoric) correlation matrix
Revelle's omega	$\omega_R = \frac{\left(\sum_{i=1}^{k} \lambda_{gi}\right)^2 + \left(\sum_{f=1}^{F} \sum_{i=1}^{k_f} \lambda_{fi}\right)^2}{V_X}$	Where λ_{gi} is the loading of the ith item on the general factor, λ_{fi} is the standardized loading of the ith item on the fth group factor, k is the total number of items, F is the total number of group factors, k_f is the number of items that load on the fth group factor, and V_X is the total variance after rotation, which is equal to the sum of each element of the sample Pearson (or polychoric) correlation matrix
GLB	$GLB = 1 - \frac{trace[Cov(E)]}{s_X^2}$	Where s_X^2 is the variance of the observed items, $Cov(E)$ is the covariance matrix of all the errors

Note. Formulas taken from McNeish (2018).

increase based on how unbalanced the loadings were. However, with twice as many items (12), this bias was somewhat reduced compared with the shorter scale (6), though still present overall. Thus, it can also be concluded that the number of items within a scale also affects reliability estimates and protects against downward biasing for coefficient alpha.

Additionally, it has long been known that the sample size affects the precision of the reliability estimate (Charter, 1999), such that larger sample sizes allow for more precise estimates of reliability. For methods relying on factor analytic techniques, sample size becomes particularly important. In the same way that alpha is based on correlational techniques, and anything affecting the estimation of the correlation also affects the alpha estimate, conditions that affect factor analytic solutions (e.g., sample size) also are likely to bias reliability estimates based on factor analytic techniques. Extant work on this topic demonstrates that sample size alone does not tend to bias factor analytic solutions (MacCallum et al., 1999; MacCallum et al., 2001); however, smaller samples sizes strongly interact with low variable communality in recovering population factors (de Winter et al., 2009; MacCallum et al., 2001). Thus, reliability estimates produced by omega and GLB may demonstrate bias when sample sizes are small and item loadings are unequal simultaneously (e.g., violation of tau equivalence). There are some suggestions in the extant literature that alternative measures of reliability beyond alpha may show bias under certain conditions. As discussed in Sijtsma (2009a, 2009b), GLB and omega estimates may be biased upward under conditions of small samples sizes due to increased likelihood of high sample variability affecting the covariance matrix used to compute these metrics. Indeed, ten Berge and Sočan (2004) commented that GLB ''has a reputation for overestimating the population value when the sample size is small'' (p. 613) and further suggest that overestimation is most accentuated when the true population reliability is low. Notably, a recent simulation testing a closed-form expression of omega (Hancock & An, 2020) did not test any sample sizes lower than $n = 250$, which unfortunately is substantially larger than those common in many applied areas of psychology and education research. Even among the larger sample sizes examined by these authors, estimation failures for omega were not uncommon, which leads to a question about how prevalence of estimation failures change for omega as the sample size decreases. Thus, the current work will investigate how omega estimators perform under smaller sample sizes that tend to be more widespread in these applied settings.

Current Study

While coefficient alpha has been thoroughly critiqued in the literature and simulation studies have quantified the bias introduced by deviation from optimal conditions (e.g., Green & Yang, 2009), the degree to which these same conditions affect bias in other methods of estimating reliability have yet to be empirically investigated. In the current work, we sought to test the ability of different methods of estimating reliability—using Cronbach's alpha, omega, omega hierarchical, Revelle's omega, and greatest lower bound—under different common conditions affecting reliability via simulation. Our work is an extension of the work of Green and Yang (2009) through extending these simulation conditions to estimators other than alpha; as such, all of their simulation conditions will be repeated here. We varied relevant components of a unidimensional scale, including the number of items in the scale (6 vs. 12), the sample size (from 30 to 1,000), population reliabilities, and varying degrees of violation of tau equivalence. Given the exploratory nature of the current simulation work, we did not make specific a priori hypotheses about the performance of different reliability metrics under these various conditions. However, based on the previous literature reviewed above, it was likely to be observed that (a) coefficient alpha would be downward biased when tau equivalence was not met, (b) larger sample sizes would increase accuracy of all reliability estimates, (c) more items would increase the accuracy of all reliability estimates, and (d) small sample sizes would produce positive bias in factor analytically derived indices of reliability when tau equivalence was not met.

Method

Data were simulated using the simulateData function from the lavaan package in R (Rosseel, 2012). This function generates multivariate normal data based on a given factor structure and sample size. Factor loadings were varied in accordance with those used in Green and Yang (2009) for both a six-item and 12-item scale. These factor loadings allow for the examination of conditions in which factor loadings are not all equal violating tau equivalence. For each set of factor loadings, data were simulated with samples sizes of 30, 50, 100, 500, and 1,000. All conditions simulated here were unidimensional without correlated errors. These conditions represent only a small subset of possible conditions but provide insight for common conditions observed in psychology and education sciences that may affect internal consistency (i.e., sample size, number of items, population reliability, degree of violation of tau equivalence). Reliability measures were calculated using the scaleStructure function from the "userfriendlyscience" package in R (Peters, 2014). Reliability measures investigated here included Cronbach's alpha, omega, greatest lower bound, Revelle's omega, and omega hierarchical. Data were simulated 1,000 times and reliability estimates were averaged across the 1,000 runs. Runs that obtained omega values greater than 1 were removed from the average of all measures (the percentage of simulations that obtained impossible omega values can be found in Tables 2 and 3). The reason for the impossible omega values have to do with the computation of the denominator in the equation for omega: $\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \sum_{i=1}^{k} \theta_{ii}$, where λ_i is the factor loading for the ith item on the scale, θ_{ii} is the error variance for the ith item, and k is the number of items on the scale (see Table 1 for full equation). Thus, if the derived factor loadings produce either (a) a nonpositive-definite implied covariance matrix (affecting the first half of the above-noted part of the denominator) or (b) negative residual variances (i.e., Heywood cases; affecting the second half of the above-noted part of the

Table 2. Results for Six-Item Scales.

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB	Impossible omega (%)
.2 .2 .2 .2 .2 .2	.200	30	.1307	.2882	.3621	.6373	.4813	21.53
.2 .2 .2 .2 .2 .2	.200	50	.1857	.2930	.3132	.5732	.4472	15.97
.2 .2 .2 .2 .2 .2	.200	100	.1906	.2549	.2460	.4882	.3776	15.41
.2 .2 .2 .2 .2 .2	.200	500	.2010	.2144	.1567	.3817	.2822	1.58
.2 .2 .2 .2 .2 .2	.200	1,000	.1991	.2034	.1437	.3639	.2564	0.32
.5 .2 .2 .2 .2 .2	.288	30	.2476	.3728	.3717	.6543	.5573	17.16
.5 .2 .2 .2 .2 .2	.288	50	.2747	.3655	.3192	.5885	.5087	15.86
.5 .2 .2 .2 .2 .2	.288	100	.2786	.3398	.2683	.5192	.4480	8.09
.5 .2 .2 .2 .2 .2	.288	500	.2738	.2999	.2270	.4476	.3670	0.41
.5 .2 .2 .2 .2 .2	.288	1,000	.2774	.2940	.2365	.4285	.3535	0
.5 .5 .2 .2 .2 .2	.378	30	.3472	.4482	.3891	.6784	.6238	14.55
.5 .5 .2 .2 .2 .2	.378	50	.3546	.4283	.3459	.6180	.5712	9.95
.5 .5 .2 .2 .2 .2	.378	100	.3533	.4040	.3089	.5671	.5152	4.84
.5 .5 .2 .2 .2 .2	.378	500	.3593	.3813	.3047	.5172	.4506	0
.5 .5 .2 .2 .2 .2	.378	1,000	.3616	.3805	.3156	.5018	.4355	0
.5 .5 .5 .2 .2 .2	.462	30	.4340	.5178	.4149	.7084	.6753	7.90
.5 .5 .5 .2 .2 .2	.462	50	.4415	.4973	.3866	.6629	.6328	5.38
.5 .5 .5 .2 .2 .2	.462	100	.4363	.4703	.3569	.6211	.5779	0.41
.5 .5 .5 .2 .2 .2	.462	500	.4442	.4640	.3856	.5818	.5117	0
.5 .5 .5 .2 .2 .2	.462	1,000	.4441	.4626	.3996	.5662	.4942	0
.5 .5 .5 .5 .2 .2	.539	30	.5048	.5618	.4351	.7336	.7142	5.51
.5 .5 .5 .5 .2 .2	.539	50	.5158	.5536	.4145	.7035	.6791	1.76
.5 .5 .5 .5 .2 .2	.539	100	.5159	.5390	.4137	.6727	.6336	0.10
.5 .5 .5 .5 .2 .2	.539	500	.5231	.5385	.4577	.6386	.5736	0
.5 .5 .5 .5 .2 .2	.539	1,000	.5249	.5395	.4756	.6283	.5576	0
.5 .5 .5 .5 .5 .2	.608	30	.5803	.6174	.4558	.7641	.7531	2.87
.5 .5 .5 .5 .5 .2	.608	50	.5911	.6127	.4567	.7470	.7246	0.20

(continued)

Table 2. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB	Impossible omega (%)
.5 .5 .5 .5 .5 .2	.608	100	.5894	.6039	.4640	.7208	.6830	0
.5 .5 .5 .5 .5 .2	.608	500	.5993	.6082	.5273	.6926	.6358	0
.5 .5 .5 .5 .5 .2	.608	1,000	.5982	.6065	.5484	.6818	.6211	0
.5 .5 .5 .5 .5 .5	.667	30	.6507	.6705	.4906	.7986	.7915	2.17
.5 .5 .5 .5 .5 .5	.667	50	.6504	.6612	.4864	.7784	.7585	0.10
.5 .5 .5 .5 .5 .5	.667	100	.6611	.6669	.5212	.7646	.7313	0
.5 .5 .5 .5 .5 .5	.667	500	.6662	.6673	.5914	.7388	.6919	0
.5 .5 .5 .5 .5 .5	.667	1,000	.6668	.6674	.6117	.7286	.6832	0
.8 .2 .2 .2 .2 .2	.386	30	.3188	.4443	.3895	.6715	.6096	14.15
.8 .2 .2 .2 .2 .2	.386	50	.3402	.4323	.3553	.6145	.5717	14.12
.8 .2 .2 .2 .2 .2	.386	100	.3523	.4267	.3239	.5707	.5342	7.62
.8 .2 .2 .2 .2 .2	.386	500	.3408	.4015	.3371	.5199	.4978	0.91
.8 .2 .2 .2 .2 .2	.386	1,000	.3408	.3919	.3504	.5084	.4971	0.10
.8 .8 .2 .2 .2 .2	.558	30	.4871	.5897	.4510	.7440	.7421	7.55
.8 .8 .2 .2 .2 .2	.558	50	.4841	.5762	.4293	.7145	.7131	4.18
.8 .8 .2 .2 .2 .2	.558	100	.4945	.5674	.4210	.6883	.6919	0.90
.8 .8 .2 .2 .2 .2	.558	500	.5006	.5589	.4468	.6555	.6724	0
.8 .8 .2 .2 .2 .2	.558	1,000	.5018	.5589	.4619	.6485	.6689	0
.8 .8 .8 .2 .2 .2	.694	30	.6249	.6905	.5273	.8213	.8204	0.20
.8 .8 .8 .2 .2 .2	.694	50	.6324	.6937	.5212	.8005	.8020	0
.8 .8 .8 .2 .2 .2	.694	100	.6386	.6942	.5426	.7887	.7856	0
.8 .8 .8 .2 .2 .2	.694	500	.6435	.6943	.5453	.7642	.7684	0
.8 .8 .8 .2 .2 .2	.694	1,000	.6434	.6940	.5487	.7536	.7639	0
.8 .8 .8 .8 .2 .2	.794	30	.7409	.7878	.6386	.8791	.8752	0
.8 .8 .8 .8 .2 .2	.794	50	.7474	.7894	.6484	.8642	.8574	0
.8 .8 .8 .8 .2 .2	.794	100	.7541	.7922	.6407	.8576	.8417	0
.8 .8 .8 .8 .2 .2	.794	500	.7581	.7941	.6288	.8306	.8229	0

(continued)

Table 2. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB	Impossible omega (%)
.8 .8 .8 .8 .2 .2	.794	1,000	.7585	.7942	.6075	.8235	.8193	0
.8 .8 .8 .8 .8 .2	.865	30	.8368	.8618	.7414	.9195	.9140	0
.8 .8 .8 .8 .8 .2	.865	50	.8414	.8624	.7592	.9116	.9002	0
.8 .8 .8 .8 .8 .2	.865	100	.8450	.8643	.7691	.9039	.8865	0
.8 .8 .8 .8 .8 .2	.865	500	.8461	.8642	.7102	.8864	.8617	0
.8 .8 .8 .8 .8 .2	.865	1,000	.8465	.8643	.6677	.8711	.8559	0
.8 .8 .8 .8 .8 .8	.914	30	.9088	.9120	.8199	.9467	.9432	0
.8 .8 .8 .8 .8 .8	.914	50	.9106	.9126	.8408	.9421	.9343	0
.8 .8 .8 .8 .8 .8	.914	100	.9126	.9136	.8604	.9381	.9270	0
.8 .8 .8 .8 .8 .8	.914	500	.9141	.9143	.8671	.9315	.9188	0
.8 .8 .8 .8 .8 .8	.914	1,000	.9140	.9141	.8543	.9260	.9168	0
.8 .8 .5 .5 .2 .2	.685	30	.6277	.6813	.5226	.8098	.8097	1.73
.8 .8 .5 .5 .2 .2	.685	50	.6403	.6841	.5379	.7948	.7918	0.10
.8 .8 .5 .5 .2 .2	.685	100	.6478	.6854	.5628	.7803	.7739	0
.8 .8 .5 .5 .2 .2	.685	500	.6508	.6844	.5883	.7564	.7581	0
.8 .8 .5 .5 .2 .2	.685	1,000	.6511	.6847	.5983	.7469	.7563	0

Note. GLB = greatest lower bound.

denominator), the formula for omega will then produce an imaginary or negative number and cause these failures. A comparison of performance for the estimation of reliability is shown in Tables 2 and 3.

Results and Discussion

Results showed significant variability in estimation accuracy across the five reliability measures investigated here. Surprisingly, alpha and omega performed similarly well, whereas the rest of the measures were often consistently poor estimators. However, omega presents a problem of estimation failure in which the value returned is outside the range of possible values (i.e., greater than 1). An estimation failure can occur due to (a) a nonpositive-definite implied covariance matrix or (b) negative residual variances (e.g., ''Heywood cases''), when the sum of the covariance terms is negative within a square root. This occurred for as much as 21.53% of the simulations in a given condition. Failures occurred most when sample size and population reliability were low. The estimation failures observed here align with the results of Hancock and An (2020) who reported greater failure rates when factor loadings and sample sizes are lower. Hancock and An proposed a closed-form expression of omega that does avoid these reasons for estimation failure; however, estimation failure can occur using this expression when the sum of the covariance terms is negative within a square root. Thus, there is a legitimate practical concern about using omega in smaller samples with uneven loadings, despite the lack of consequences for violation of tau equivalence.

Cronbach's alpha appears to slightly underestimate reliability especially with smaller sample sizes. Replicating the simulation-based findings of Green and Yang (2009), alpha did have a negative bias that is more pronounced when tau equivalence is not met. However, this negative bias was small (−.02 on average) and even when tau equivalence was not met, alpha outperformed all measures except omega in terms of absolute distance from the population reliability value. One of the major criticisms of Cronbach's alpha is the consequences of violating tau equivalence (Green & Yang, 2009; McNeish, 2018; Sijtsma, 2009a). Although more accurate estimations were obtained when all the factor loadings were equal, when tau equivalence was violated, Cronbach's alpha provided an estimate smaller than the population value in all cases, which is consistent with arguments favoring its legitimacy as a lower bound.[1] Raykov and Marcoulides (2015) demonstrate that the underestimation by alpha is practically negligible when the average loading is above .7 and the difference between the individual loading and average loading is less than .2. Here, we expand this significantly by showing that even when these guidelines are not met, when continuous data are unidimensional with uncorrelated errors, alpha provides a reasonable estimate of internal consistency. Here, we present multiple conditions in which these two criteria are both violated and the resulting estimate is relatively accurate. Even when sample size was 30, population reliability was low, and these two criteria were not upheld, the largest underestimation provided by alpha was .07.

Table 3. Results for 12-Item Scales.

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB	Impossible omega (%)
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	30	.3032	.3064	.3706	.6196	.7320	13.34
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	50	.3120	.3005	.3150	.5306	.6549	9.97
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	100	.3251	.3070	.2554	.4431	.5734	6.46
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	500	.3303	.3297	.1774	.3755	.4432	0
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	1,000	.3326	.3338	.1811	.3855	.4112	0
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	30	.4115	.4263	.3811	.6452	.7769	10.36
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	50	.4163	.4254	.3384	.5707	.7097	5.29
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	100	.4344	.4477	.2945	.5212	.6455	1.54
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	500	.4378	.4477	.3122	.4935	.5383	0
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	1,000	.4373	.4467	.3392	.4882	.5056	0
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	30	.5181	.5310	.4069	.6872	.8181	4.70
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	50	.5217	.5367	.3621	.6350	.7650	1.35
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	100	.5277	.5436	.3517	.6022	.7076	0.10
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	500	.5355	.5482	.4071	.5875	.6151	0
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	1,000	.5361	.5480	.4408	.5834	.5770	0
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	30	.6105	.6248	.4295	.7315	.8538	1.75
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	50	.6159	.6341	.4181	.6996	.8124	0.10
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	100	.6214	.6382	.4290	.6855	.7667	0
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	500	.6329	.6457	.5149	.6772	.6822	0
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	1,000	.6331	.6455	.5388	.6729	.6544	0
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	30	.6766	.6876	.4511	.7704	.8801	0.72
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	50	.6801	.6931	.4543	.7521	.8423	0
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	100	.6891	.7009	.4779	.7427	.8053	0
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	500	.6918	.7006	.5746	.7297	.7436	0
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	1,000	.6923	.7008	.6036	.7256	.7230	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	30	.7321	.7414	.4823	.8067	.8987	0.41
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	50	.7373	.7471	.4881	.7965	.8684	0

(continued)

Table 3. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB	Impossible omega (%)
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	100	.7471	.7546	.5380	.7914	.8393	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	500	.7496	.7546	.6418	.7781	.7894	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	1,000	.7512	.7558	.6696	.7760	.7765	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	30	.7858	.7919	.4994	.8424	.9189	0.20
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	50	.7909	.7954	.5236	.8358	.8939	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	100	.7961	.7985	.5842	.8289	.8671	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	500	.7998	.8003	.6990	.8188	.8283	0
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	1,000	.7996	.7999	.7304	.8161	.8178	0
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	30	.5032	.5483	.4132	.6888	.8226	4.97
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	50	.5116	.5471	.3978	.6404	.7762	1.73
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	100	.5190	.5523	.3818	.6023	.7337	0.20
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	500	.5263	.5562	.4236	.5868	.7042	0
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	1,000	.5279	.5571	.4496	.5883	.7018	0
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	30	.6593	.7059	.4757	.7793	.8822	0.10
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	50	.6708	.7110	.4576	.7536	.8497	0
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	100	.6778	.7140	.4675	.7422	.8186	0
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	500	.6831	.7165	.5253	.7380	.7979	0
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	1,000	.6835	.7164	.5495	.7375	.7928	0
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	30	.7788	.8129	.5325	.8536	.9093	0
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	50	.7865	.8178	.5257	.8443	.8964	0
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	100	.7905	.8193	.5456	.8372	.8809	0
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	500	.7918	.8192	.6052	.8355	.8580	0
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	1,000	.7926	.8197	.6023	.8334	.8501	0
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	30	.8573	.8801	.6031	.9063	.9259	0
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	50	.8634	.8841	.6028	.9032	.9147	0
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	100	.8648	.8842	.6025	.8977	.9013	0
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	500	.8666	.8848	.6106	.8902	.8855	0

(continued)

Table 3. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB	Impossible omega (%)
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	1,000	.8668	.8849	.6444	.8967	.8849	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	30	.8885	.9059	.6543	.9281	.9357	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	50	.8905	.9059	.6339	.9260	.9261	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	100	.8928	.9072	.6264	.9210	.9155	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	500	.8950	.9083	.6565	.9175	.9067	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	1,000	.8954	.9087	.6315	.9149	.9065	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	30	.9522	.9538	.8468	.9655	.9806	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	50	.9537	.9547	.8736	.9638	.9755	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	100	.9546	.9550	.8992	.9618	.9694	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	500	.9551	.9552	.8763	.9578	.9606	0
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	1,000	.9552	.9552	.8111	.9546	.9586	0
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	30	.7793	.8058	.5630	.8545	.9156	0.10
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	50	.7872	.8097	.5840	.8433	.8945	0
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	100	.7887	.8089	.5959	.8324	.8646	0
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	500	.7941	.8124	.6506	.8277	.8391	0
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	1,000	.7945	.8124	.6491	.8252	.8343	0

Note. GLB = greatest lower bound.

Alpha outperformed omega when population reliability, sample size, and number of items were low. Whereas, conditions under which omega provided more accurate estimations included when there was a large difference between factor loadings and multiple loadings that were different from the others (e.g., .8 .8 .8 .2 .2 .2). When tau equivalence was met, alpha and omega preformed equally well, with a slight favoring toward alpha when the population reliability was low. Overall, both alpha and omega provided largely accurate estimates of reliability under the conditions simulated here; however, the estimation failure presented by omega suggests additional challenges for the use of omega as a complete replacement for alpha.

Accuracy in estimation of reliability was influenced by not only sample size and degree of violation of tau equivalence but also the magnitude of the population reliability itself. Across metrics, more accurate predictions were estimated for larger population reliabilities. The lack of precision in estimation can also be observed in the 95% confidence intervals shown in Tables 4 and 5, in which greater variability in estimates are observed for lower population reliabilities, particularly, accentuated by smaller sample sizes. Practically, this quality of omega is difficult to assess, as one is unable to determine the observed reliability without using an estimator, but the estimate produced by that estimator may be producing a biased estimate when the true population reliability is lower. This presents another reason why it may be immature to completely abandon the common use of alpha for omega in applied psychological sciences; one may need to use alpha to first know if observed reliability is low enough to cause concern about the accuracy of the omega estimate. As such, the ideal scenario may not be to replace one with the other, but rather use alpha and omega as *complementary* estimators in tandem.

The results of the present study are consistent with those of ten Berge and Sočan (2004) who found greatest lower bound to overestimate reliability when sample size and reliability are low. Ironically, greatest lower bound consistently overestimated reliability despite the idea that it is supposed to be a lower bound as suggested by its name. On average, GLB overestimated reliability on all but seven of the 140 conditions simulated here. Although the probability of GLB to overestimate reliability in small samples has been previously discussed, no study to date has defined what a ''small sample'' means for this bias. Here, we show that GLB overestimates reliability even when the sample size is 1,000. Similarly, Shapiro and ten Berge (2000) show GLB to overestimate reliability in samples of 100, 500, and 2,000 on five- and 10-item scales. Consistent with what was observed here, greater overestimations were observed at smaller samples sizes, but even at larger sample sizes, GLB rarely resulted in underestimation of the true population reliability.

Revelle's omega performed differentially well between the six-item and 12-item scales, such that Revelle's omega overestimated reliability more on the six-item scales than the 12-item scales. This may be due to a tendency of this metric to more often overestimate smaller reliabilities than larger ones, and this intersected with the fact that population reliabilities in the six-item scales were smaller on average than the 12-item scales.

Table 4. Ninety-Five Percent Confidence Intervals for Six-Item Scales.

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB
.2 .2 .2 .2 .2 .2	.200	30	[−.2409, .4689]	[−.0804, .5875]	[.0021, .639]	[.3596, .8113]	[.1465, .7172]
.2 .2 .2 .2 .2 .2	.200	50	[−.0977, .4412]	[.016, .5283]	[.0382, .5441]	[.3509, .7344]	[.1929, .6453]
.2 .2 .2 .2 .2 .2	.200	100	[−.006, .3731]	[.0615, .4298]	[.0521, .422]	[.3228, .6248]	[.1957, .5344]
.2 .2 .2 .2 .2 .2	.200	500	[.1154, .2837]	[.1291, .2965]	[.07, .2411]	[.3042, .4542]	[.1995, .361]
.2 .2 .2 .2 .2 .2	.200	1,000	[.1389, .258]	[.1432, .2621]	[.0824, .2038]	[.3088, .4165]	[.1975, .3134]
.5 .2 .2 .2 .2 .2	.288	30	[−.1237, .5581]	[.0144, .6462]	[.0132, .6455]	[.3847, .821]	[.2465, .7641]
.5 .2 .2 .2 .2 .2	.288	50	[−.004, .5137]	[.097, .5844]	[.0449, .5488]	[.3709, .7449]	[.2683, .6894]
.5 .2 .2 .2 .2 .2	.288	100	[.0869, .4503]	[.1537, .5027]	[.0758, .4414]	[.3594, .6494]	[.2758, .5923]
.5 .2 .2 .2 .2 .2	.288	500	[.1907, .3531]	[.218, .3777]	[.1421, .3085]	[.3746, .5151]	[.2886, .4405]
.5 .2 .2 .2 .2 .2	.288	1,000	[.2192, .3337]	[.2363, .3496]	[.1771, .2942]	[.3765, .4778]	[.2981, .4066]
.5 .5 .2 .2 .2 .2	.378	30	[−.015, .6288]	[.1049, .6961]	[.0336, .6573]	[.421, .8347]	[.3399, .8035]
.5 .5 .2 .2 .2 .2	.378	50	[.0846, .5761]	[.1702, .6314]	[.0748, .5694]	[.4102, .7648]	[.3482, .733]
.5 .5 .2 .2 .2 .2	.378	100	[.1686, .5141]	[.2255, .5563]	[.1197, .4764]	[.4171, .687]	[.3547, .6462]
.5 .5 .2 .2 .2 .2	.378	500	[.2805, .4333]	[.3038, .4538]	[.223, .3822]	[.4499, .5786]	[.3779, .5179]
.5 .5 .2 .2 .2 .2	.378	1,000	[.3064, .4143]	[.3262, .4323]	[.2587, .3704]	[.4539, .5468]	[.3838, .4844]
.5 .5 .5 .2 .2 .2	.462	30	[.0874, .6869]	[.1937, .7401]	[.0642, .6744]	[.4675, .8514]	[.4164, .833]
.5 .5 .5 .2 .2 .2	.462	50	[.186, .6411]	[.2542, .6814]	[.1213, .6003]	[.4716, .7947]	[.4302, .7747]
.5 .5 .5 .2 .2 .2	.462	100	[.2623, .5827]	[.3018, .6104]	[.1725, .517]	[.4837, .7286]	[.4304, .6954]
.5 .5 .5 .2 .2 .2	.462	500	[.3709, .5119]	[.3923, .5301]	[.3084, .4579]	[.5207, .637]	[.4439, .5736]
.5 .5 .5 .2 .2 .2	.462	1,000	[.3929, .4925]	[.4125, .51]	[.3462, .4505]	[.5225, .6069]	[.4459, .5396]
.5 .5 .5 .5 .2 .2	.539	30	[.1766, .7319]	[.2526, .7668]	[.0887, .6876]	[.5074, .8652]	[.4766, .8546]
.5 .5 .5 .5 .2 .2	.539	50	[.2772, .6944]	[.3254, .7209]	[.1539, .6212]	[.5287, .8211]	[.4941, .8052]
.5 .5 .5 .5 .2 .2	.539	100	[.3555, .6468]	[.3832, .665]	[.2365, .5643]	[.5488, .7677]	[.4993, .7381]
.5 .5 .5 .5 .2 .2	.539	500	[.4564, .584]	[.4732, .598]	[.3855, .5243]	[.5835, .6877]	[.5116, .6296]
.5 .5 .5 .5 .2 .2	.539	1,000	[.4785, .5684]	[.494, .582]	[.4261, .5222]	[.5893, .6644]	[.5133, .5989]
.5 .5 .5 .5 .5 .2	.608	30	[.2782, .7779]	[.3306, .7998]	[.1143, .701]	[.5572, .8817]	[.5392, .8758]
.5 .5 .5 .5 .5 .2	.608	50	[.3743, .7466]	[.4031, .7612]	[.2043, .6522]	[.5916, .8488]	[.559, .8346]

(continued)

Table 4. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB
.5 .5 .5 .5 .5 .2	.608	100	[.4444, .7043]	[.4623, .7155]	[.2944, .6052]	[.6108, .8034]	[.562, .7754]
.5 .5 .5 .5 .5 .2	.608	500	[.5399, .6527]	[.5498, .6607]	[.4609, .5878]	[.6441, .7357]	[.5805, .6853]
.5 .5 .5 .5 .5 .2	.608	1,000	[.5568, .6366]	[.5658, .6443]	[.5036, .5904]	[.6471, .7136]	[.5815, .6577]
.5 .5 .5 .5 .5 .5	.667	30	[.3793, .819]	[.4091, .8302]	[.1583, .7231]	[.6153, .8999]	[.6033, .8962]
.5 .5 .5 .5 .5 .5	.667	50	[.4543, .7864]	[.4692, .7935]	[.2406, .6735]	[.6384, .8686]	[.6087, .8561]
.5 .5 .5 .5 .5 .5	.667	100	[.534, .759]	[.5413, .7633]	[.3618, .6509]	[.6686, .8355]	[.6246, .8112]
.5 .5 .5 .5 .5 .5	.667	500	[.6144, .7123]	[.6157, .7133]	[.5313, .6456]	[.6962, .7762]	[.6433, .735]
.5 .5 .5 .5 .5 .5	.667	1,000	[.6309, .6999]	[.6315, .7004]	[.5714, .6491]	[.6982, .7565]	[.6486, .7149]
.8 .2 .2 .2 .2 .2	.386	30	[−.0469, .6091]	[.1001, .6936]	[.034, .6575]	[.4105, .8308]	[.3195, .7952]
.8 .2 .2 .2 .2 .2	.386	50	[.0683, .565]	[.1751, .6344]	[.0853, .5766]	[.4056, .7625]	[.3489, .7334]
.8 .2 .2 .2 .2 .2	.386	100	[.1674, .5132]	[.2514, .575]	[.1361, .4892]	[.4215, .6898]	[.3774, .6612]
.8 .2 .2 .2 .2 .2	.386	500	[.2609, .416]	[.3253, .4726]	[.2571, .4126]	[.4528, .5811]	[.4288, .561]
.8 .2 .2 .2 .2 .2	.386	1,000	[.2848, .3944]	[.3381, .4431]	[.2948, .4036]	[.4609, .5529]	[.4489, .5424]
.8 .8 .2 .2 .2 .2	.558	30	[.1538, .7209]	[.2913, .7835]	[.1083, .6979]	[.5242, .8708]	[.5211, .8698]
.8 .8 .2 .2 .2 .2	.558	50	[.2378, .6719]	[.3548, .7365]	[.1714, .6321]	[.5445, .8282]	[.5424, .8273]
.8 .8 .2 .2 .2 .2	.558	100	[.3301, .6297]	[.4175, .6872]	[.2448, .5703]	[.5688, .7794]	[.5735, .7821]
.8 .8 .2 .2 .2 .2	.558	500	[.4318, .5635]	[.4955, .6164]	[.3737, .5143]	[.6024, .7028]	[.6214, .7178]
.8 .8 .2 .2 .2 .2	.558	1,000	[.4539, .5468]	[.5148, .6001]	[.4117, .5093]	[.6111, .6831]	[.6331, .7018]
.8 .8 .8 .2 .2 .2	.694	30	[.3415, .8041]	[.4396, .8414]	[.2062, .7459]	[.6548, .9118]	[.6532, .9113]
.8 .8 .8 .2 .2 .2	.694	50	[.4297, .7744]	[.5147, .8147]	[.284, .6982]	[.6719, .8823]	[.6742, .8832]
.8 .8 .8 .2 .2 .2	.694	100	[.5056, .742]	[.5763, .7837]	[.3875, .6679]	[.7009, .853]	[.6966, .8507]
.8 .8 .8 .2 .2 .2	.694	500	[.589, .6921]	[.6459, .7371]	[.4806, .6041]	[.7251, .7984]	[.7299, .8021]
.8 .8 .8 .2 .2 .2	.694	1,000	[.6056, .6784]	[.6605, .7248]	[.5039, .5906]	[.7255, .7792]	[.7368, .7886]
.8 .8 .8 .8 .2 .2	.794	30	[.5192, .8692]	[.5969, .8942]	[.3615, .812]	[.7593, .9413]	[.7521, .9393]
.8 .8 .8 .8 .2 .2	.794	50	[.5922, .8491]	[.655, .8754]	[.4516, .7851]	[.7715, .921]	[.7606, .9169]
.8 .8 .8 .8 .2 .2	.794	100	[.6547, .8279]	[.7056, .8555]	[.5082, .7435]	[.7952, .9021]	[.7731, .8908]
.8 .8 .8 .8 .2 .2	.794	500	[.7181, .7931]	[.7592, .8243]	[.5727, .679]	[.8013, .856]	[.7924, .8493]

(continued)

Table 4. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB
.8 .8 .8 .8 .2 .2	.794	1,000	[.7309, .7837]	[.7701, .816]	[.5669, .6452]	[.8024, .8424]	[.7978, .8387]
.8 .8 .8 .8 .8 .2	.865	30	[.6823, .9198]	[.7274, .9325]	[.52, .8694]	[.8362, .9613]	[.8256, .9586]
.8 .8 .8 .8 .8 .2	.865	50	[.7352, .9072]	[.7686, .9199]	[.6097, .8566]	[.8486, .9491]	[.8298, .9424]
.8 .8 .8 .8 .8 .2	.865	100	[.7777, .8932]	[.8045, .9068]	[.6745, .8388]	[.8602, .9344]	[.8356, .9223]
.8 .8 .8 .8 .8 .2	.865	500	[.8192, .8693]	[.8402, .8849]	[.6638, .7511]	[.866, .9038]	[.8373, .8827]
.8 .8 .8 .8 .8 .2	.865	1,000	[.8279, .8632]	[.8477, .8792]	[.6319, .7007]	[.8553, .8853]	[.8384, .8717]
.8 .8 .8 .8 .8 .8	.914	30	[.8155, .956]	[.8218, .9576]	[.6523, .911]	[.89, .9746]	[.8829, .9729]
.8 .8 .8 .8 .8 .8	.914	50	[.8469, .9485]	[.8502, .9497]	[.7342, .9069]	[.8997, .9669]	[.8864, .9624]
.8 .8 .8 .8 .8 .8	.914	100	[.8726, .9404]	[.874, .9411]	[.7991, .9041]	[.9093, .958]	[.8932, .9503]
.8 .8 .8 .8 .8 .8	.914	500	[.8984, .9275]	[.8987, .9276]	[.8435, .8873]	[.9189, .9422]	[.9039, .9314]
.8 .8 .8 .8 .8 .8	.914	1,000	[.9032, .9237]	[.9033, .9238]	[.8366, .8702]	[.9166, .9343]	[.9063, .9262]
.8 .8 .5 .5 .2 .2	.685	30	[.3455, .8057]	[.4255, .8363]	[.1999, .743]	[.6346, .9058]	[.6345, .9058]
.8 .8 .5 .5 .2 .2	.685	50	[.4404, .7797]	[.5012, .8085]	[.3053, .71]	[.6631, .8787]	[.6586, .8769]
.8 .8 .5 .5 .2 .2	.685	100	[.5172, .7489]	[.565, .7772]	[.4119, .6836]	[.6896, .8469]	[.681, .8423]
.8 .8 .5 .5 .2 .2	.685	500	[.5972, .6986]	[.6348, .7283]	[.5279, .6428]	[.7162, .7916]	[.7182, .7931]
.8 .8 .5 .5 .2 .2	.685	1,000	[.6139, .6854]	[.6503, .7162]	[.5569, .6366]	[.7182, .7731]	[.7285, .7817]

Note. Confidence intervals were calculated using the Fisher's (1950) transformational approach (Kelley & Pornprasertmanit, 2016). GLB = greatest lower bound.

Table 5. Ninety-Five Percent Confidence Intervals for 12-Item Scales.

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	30	[−.064, .5982]	[−.0606, .6004]	[.0119, .6448]	[.3339, .8011]	[.5049, .8643]
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	50	[.0368, .5432]	[.0242, .5342]	[.0402, .5456]	[.2959, .7048]	[.4605, .7894]
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	100	[.1375, .4902]	[.1177, .4748]	[.0621, .4302]	[.2702, .5883]	[.4248, .6919]
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	500	[.2499, .4063]	[.2492, .4056]	[.0912, .2611]	[.2976, .4484]	[.3699, .5111]
.2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.333	1,000	[.2763, .3866]	[.2775, .3878]	[.1205, .2404]	[.3314, .437]	[.3583, .4614]
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	30	[.0602, .6721]	[.078, .6819]	[.0241, .6519]	[.3711, .8158]	[.5786, .8885]
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	50	[.156, .6225]	[.1668, .6292]	[.0663, .5637]	[.3475, .7326]	[0.5375, 0.825]
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	100	[.2602, .5812]	[.2755, .592]	[.1041, .4641]	[.3618, .651]	[0.5143, 0.7472]
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	500	[.3641, .506]	[.3747, .5152]	[.2308, .3892]	[.4242, .5571]	[.473, .5978]
.5 .5 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.448	1,000	[.3858, .4861]	[.3956, .495]	[.2832, .3929]	[.4395, .534]	[.458, .5503]
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	30	[.194, .7402]	[.2111, .7481]	[.0546, .6691]	[.4345, .8396]	[.6491, .9101]
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	50	[.2846, .6986]	[.3037, .7091]	[.0932, .5818]	[.4333, .7762]	[.6183, .8602]
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	100	[.3696, .6561]	[.3887, .6687]	[.1668, .5127]	[.4602,.7141]	[.5937, .7937]
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	500	[.4699, .5952]	[.4838, .6067]	[.3312, .4777]	[.5269, .6421]	[.5575, .6668]
.5 .5 .5 .5 .2 .2 .2 .2 .2 .2 .2 .2	.548	1,000	[.4905, .5789]	[.5031, .59]	[.3894, .4894]	[.541, .6229]	[.5341, .6169]
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	30	[.3208, .7958]	[.3414, .8041]	[.082, .684]	[.5041, .8641]	[.7129, .9285]
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	50	[.4074, .7634]	[.432, .7756]	[.1581, .6238]	[.5232, .8186]	[.69, .8896]
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	100	[.4841, .7289]	[.5051, .7417]	[.2539, .5768]	[.5651, .7772]	[.6714, .8371]
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	500	[.5773, .6827]	[.5915, .6941]	[.4475, .5766]	[.6267, .722]	[.6324, .7265]
.5 .5 .5 .5 .5 .5 .2 .2 .2 .2 .2 .2	.632	1,000	[.5944, .6688]	[.6079, .6803]	[.4933, .5814]	[.6375, .7054]	[.6174, .6884]
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	30	[.4183, .8337]	[.4351, .8398]	[.1084, .6979]	[.5677, .885]	[.7611, .9417]
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	50	[.4955, .8059]	[.5138, .8143]	[.2014, .6504]	[.5992, .8521]	[.7367, .9078]
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	100	[.5698, .7799]	[.585, .7887]	[.3107, .6165]	[.6396, .8196]	[.7232, .8649]
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	500	[.6432, .7349]	[.6531, .7427]	[.5128, .6306]	[.6859, .7682]	[.7016, .7804]
.5 .5 .5 .5 .5 .5 .5 .5 .2 .2 .2 .2	.701	1,000	[.6586, .7232]	[.6678, .731]	[.5626, .6415]	[.6948, .7537]	[.692, .7513]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	30	[.505, .8644]	[.52, .8694]	[.1477, .7179]	[.6293, .9042]	[.7963, .9511]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	50	[.5775, .8427]	[.5919, .8489]	[.2428, .6748]	[.6657, .8798]	[.7782, .9235]

(continued)

Table 5. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	100	[.6454, .8228]	[.6553, .8282]	[.3819, .6642]	[.7045, .8549]	[.7698, .8891]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	500	[.7085, .7857]	[.7142, .79]	[.5871, .6906]	[.7409, .8105]	[.7539, .8203]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .2 .2	.756	1,000	[.7229, .777]	[.7279, .7811]	[.634, .7025]	[.75, .7995]	[.7506, .8]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	30	[.5936, .8932]	[.604, .8964]	[.1696, .7286]	[.6921, .9226]	[.8351, .961]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	50	[.6573, .8764]	[.6641, .8791]	[.2871, .7]	[.7265, .9039]	[.8195, .9387]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	100	[.7108, .8583]	[.7141, .8601]	[.438, .7003]	[.7554, .8817]	[.8083, .9087]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	500	[.7658, .8293]	[.7664, .8298]	[.6513, .7413]	[.7877, .8458]	[.7986, .854]
.5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5 .5	.800	1,000	[.7761, .8209]	[.7764, .8211]	[.7001, .7581]	[.7943, .8358]	[.7962, .8374]
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	30	[.1746, .731]	[.2343, .7587]	[.0621, .6732]	[.437, .8405]	[.6571, .9125]
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	50	[.272, .6915]	[.3171, .7164]	[.1343, .6087]	[.4405, .7797]	[.635, .8672]
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	100	[.3592, .6492]	[.3992, .6754]	[.2004, .5379]	[.4603, .7142]	[.6277, .8129]
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	500	[.4598, .5869]	[.4925, .6139]	[.3489, .493]	[.5262, .6415]	[.6571, .7459]
.8 .8 .2 .2 .2 .2 .2 .2 .2 .2 .2 .2	.557	1,000	[.4817, .5712]	[.5128, .5984]	[.3987, .4977]	[.5463, .6274]	[.669, .732]
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	30	[.3922, .8239]	[.4636, .85]	[.1393, .7137]	[.5826, .8898]	[.7651, .9428]
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	50	[.4826, .7998]	[.5394, .8259]	[.2054, .6528]	[.6013, .853]	[.7483, .9123]
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	100	[.5553, .7715]	[.6021, .7985]	[.2984, .6081]	[.6389, .8192]	[.7413, .8744]
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	500	[.6334, .7272]	[.671, .7567]	[.4587, .586]	[.6954, .7755]	[.7636, .8277]
.8 .8 .8 .8 .2 .2 .2 .2 .2 .2 .2 .2	.716	1,000	[.6491, .7152]	[.6848, .7453]	[.5047, .5913]	[.7079, .7646]	[.7686, .8148]
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	30	[.5818, .8895]	[.64, .9074]	[.2131, .7491]	[.7124, .9284]	[.8165, .9563]
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	50	[.6506, .8736]	[.6984, .8929]	[.2898, .7014]	[.7398, .909]	[.8236, .9402]
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	100	[.7032, .8543]	[.7423, .8749]	[.3911, .6702]	[.7669, .8877]	[.8277, .9184]
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	500	[.7566, .8224]	[.7881, .8461]	[.5465, .658]	[.807, .8602]	[.833, .8795]
.8 .8 .8 .8 .8 .8 .2 .2 .2 .2 .2 .2	.820	1,000	[.7683, .8145]	[.7982, .839]	[.5612, .6404]	[.8134, .8514]	[.832, .8664]
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	30	[.7192, .9303]	[.7611, .9417]	[.3103, .7914]	[.8107, .9548]	[.8487, .9644]
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	50	[.7702, .9205]	[.8035, .9329]	[.3898, .7545]	[.8347, .9442]	[.8537, .9509]
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	100	[.8052, .9072]	[.8324, .9207]	[.4606, .7144]	[.8514, .9301]	[.8566, .9326]
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	500	[.843, .8869]	[.8642, .9025]	[.5525, .6628]	[.8704, .9071]	[.8649, .903]

(continued)

Table 5. Continued

Factor loadings	Population reliability	Sample size	Alpha	Omega	Omega hierarchical	Revelle's omega	GLB
.8 .8 .8 .8 .8 .8 .8 .8 .2 .2 .2 .2	.885	1,000	[.8505, .8815]	[.8707, .8977]	[.6066, .6793]	[.8838, .9082]	[.8707, .8977]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	30	[.777, .946]	[.8099, .9546]	[.3847, .821]	[.853, .9655]	[.8681, .9693]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	50	[.8138, .9367]	[.8391, .9457]	[.4317, .7754]	[.8726, .9576]	[.8727, .9576]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	100	[.8444, .9267]	[.8649, .9367]	[.4904, .7327]	[.8847, .9463]	[.8767, .9424]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	500	[.876, .9111]	[.8917, .9225]	[.6035, .7037]	[.9024, .9303]	[.8897, .9211]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .2 .2	.927	1,000	[.8824, .9071]	[.8973, .9189]	[.5927, .6674]	[.9042, .9245]	[0.8947, 0.9169]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	30	[.9009, .9772]	[.9043, .978]	[.7002, .9249]	[.9281, .9836]	[.9592, .9908]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	50	[.9194, .9736]	[.9211, .9742]	[.7866, .9266]	[.9367, .9794]	[.9571, .9861]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	100	[.9331, .9693]	[.9338, .9696]	[.8535, .9312]	[.9436, .9742]	[.9547, .9793]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	500	[.9467, .9622]	[.9468, .9623]	[.8543, .8952]	[.9499, .9645]	[.9532, .9669]
.8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8 .8	.955	1,000	[.9494, .9603]	[.9495, .9604]	[.7888, .8313]	[.9487, .9598]	[.9533, .9634]
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	30	[.5826, .8898]	[.6277, .9037]	[.2543, .7676]	[.714, .9288]	[.8286, .9594]
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	50	[.6517, .874]	[.6859, .8879]	[.3649, .7417]	[.7383, .9084]	[.8204, .939]
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	100	[.7008, .853]	[.7281, .8675]	[.4525, .7093]	[.7603, .8842]	[.8048,.907]
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	500	[.7593, .8244]	[.7803, .8402]	[.597, .6985]	[.798, .8535]	[.8111, .8633]
.8 .8 .8 .8 .5 .5 .5 .5 .2 .2 .2 .2	.813	1,000	[.7704, .8163]	[.7903, .8325]	[.6117, .6835]	[.8044, .844]	[.8144, .8522]

Note. Confidence intervals were calculated using the Fisher's (1950) transformational approach (Kelley & Pornprasertmanit, 2016). GLB = greatest lower bound.

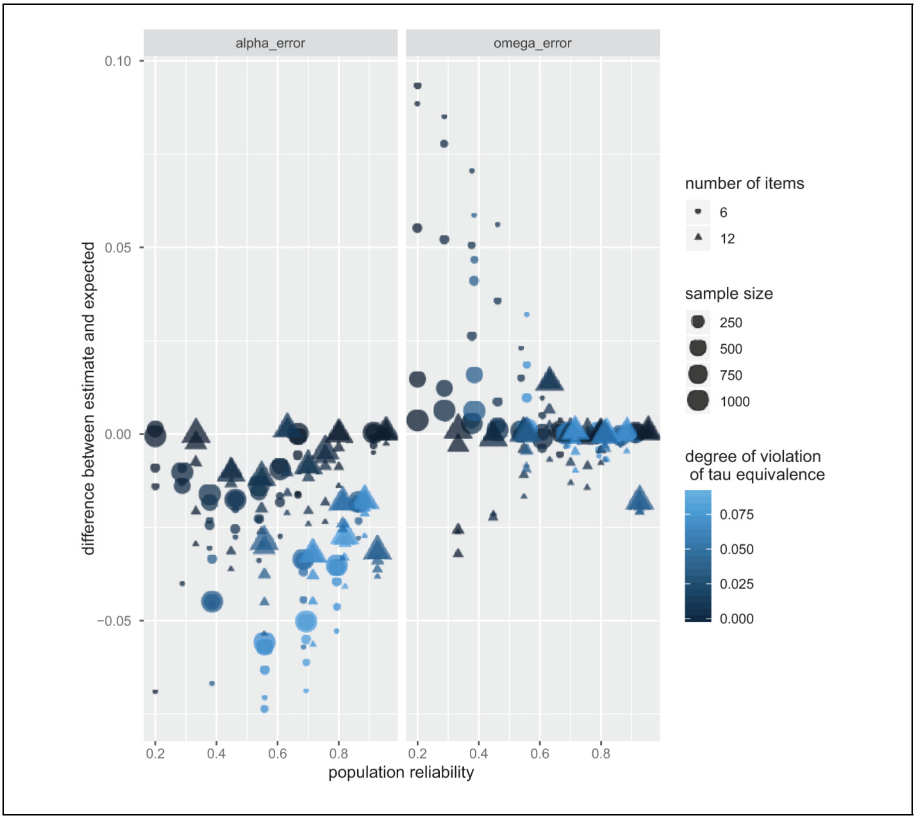

Figure 1. Observed estimation error (difference between estimate and expected reliability) for alpha and omega under the 140 conditions of varying number of items, sample size, and degree of violation of tau equivalence.

Zinbarg et al. (2005) suggest that when data are unidimensional (as simulated here), whether tau equivalence is met or not, omega and omega hierarchical will be equal. However, that is not what was observed here. Under the conditions simulated here, omega hierarchical was often smaller than omega. Furthermore, according to Zinbarg et al. (2005), when the general factor loadings are unequal, omega hierarchical should be larger than alpha, however, this pattern was not observed under the conditions simulated here. In fact, in many cases simulated here omega hierarchical was actually smaller than alpha. Future work needs to critically evaluate the reason for the divergence between omega and omega hierarchical for unidimensional data, and furthermore, why a formulaic account of omega hierarchical does not hold under simulated conditions.

Results showed alpha and omega to be the most accurate estimators. Performance under the 140 conditions for these two estimators is visualized in Figure 1. To further investigate the impact of various conditions on their estimation, a regression was

Table 6. Results of Regression Comparing Distance Between Estimated and Expected Reliability for Alpha and Omega.

Coefficient	Estimate	*t* Value	*p* Value
Intercept	0.02798	23.62	<.001*
Omega vs. alpha	−0.01794	−10.71	<.001*
12 vs. 6 items	−0.00271	−1.57	.118
Population reliability	−0.04028	−6.27	<.001*
Sample size	−0.00001	−3.96	<.001*
Degree of violation of tau equivalence	0.54387	17.82	<.001*
Omega vs. alpha × 12 vs. 6 items	−0.00137	−0.56	.574
Omega vs. alpha × population reliability	−0.02683	−2.95	.003*
12 vs. 6 items × population reliability	−0.00346	−0.41	.680
Omega vs. alpha × sample size	−0.00001	−1.24	.218
12 vs. 6 items × sample size	0.00000	−0.73	.466
Population reliability × sample size	0.00002	1.32	.188
Omega vs. alpha × degree of violation of tau equivalence	−0.60916	−14.11	<.001*
Population reliability × degree of violation of tau equivalence	−0.85579	−5.15	<.001*
Omega vs. alpha × 12 vs. 6 items × population reliability	0.06582	5.56	<.001*
Omega vs. alpha × 12 vs. 6 items × sample size	0.00001	1.74	.083
Omega vs. alpha × population reliability × sample size	0.00011	5.48	<.001*
12 vs. 6 items × population reliability × sample size	0.00001	0.60	.549
Omega vs. alpha × population reliability × degree of violation of tau equivalence	1.26106	5.37	<.001*
Omega vs. alpha × 12 vs. 6 items × population reliability × sample size	−0.00011	−3.43	<.001*

Note. The *p* values presented here are a factor of the specific 140 conditions simulated.
*$p < .05$.

performed predicting the estimation error (absolute distance between estimate and population reliability) from the estimator (alpha or omega), number of items, sample size, population reliability, and degree of violation of tau equivalence (operationalized as the mean squared deviation from the mean loading). All continuous measures were grand mean centered to aid in interpretation such that the intercept can be interpreted as the absolute distance between the estimated and expected value at the mean of all continuous measures for alpha when the number of items is six. Results of the regression are reported in Table 6. Results showed alpha to be affected greater by

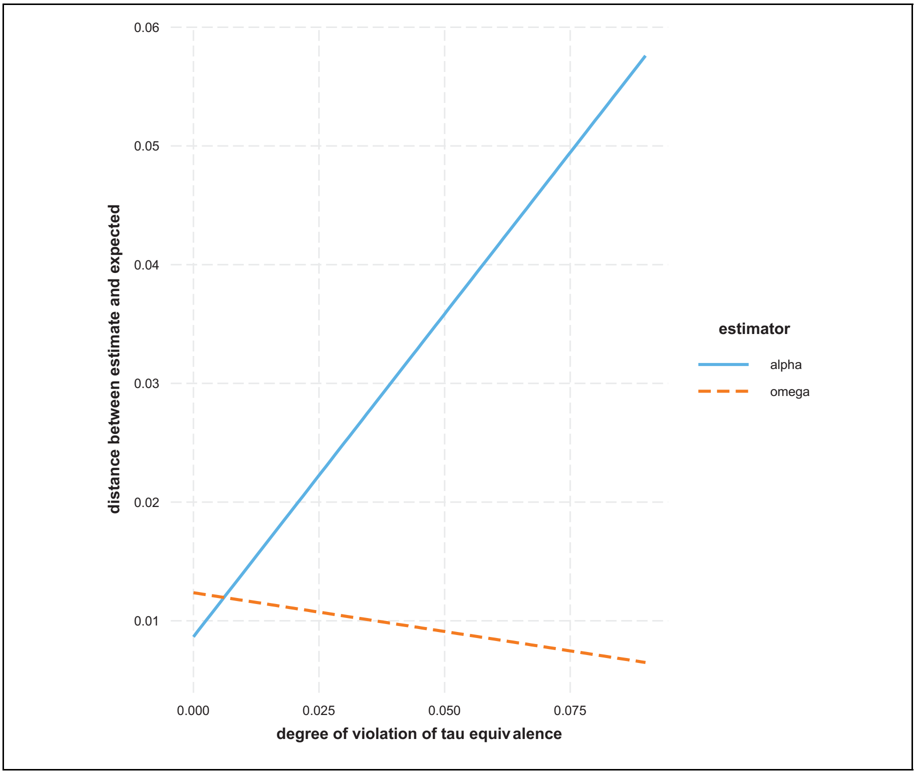

Figure 2. Regression estimated interaction between estimator, population reliability, and degree of violation of tau equivalence predicting the distance between estimate and expected reliability.

degree of violation of tau equivalence (see Figure 2) especially when population reliability was low (see Figure 3). Omega, however, showed a greater impact of sample size especially when population reliability was low (see Figure 4). Additionally, omega was affected greater by the number of items when reliability was low (see Figure 5).

Regression estimated interaction between estimator and degree of violation of tau equivalence predicting the distance between estimate and expected reliability.

Limitations and Conclusions

While the current work simulated 140 different conditions that may affect reliability estimation, there were some conditions the current work did not seek to examine, and thus, the current results almost certainly will not generalize to those conditions. Specifically, lack of unidimensionality, correlated errors/residuals, and deviations

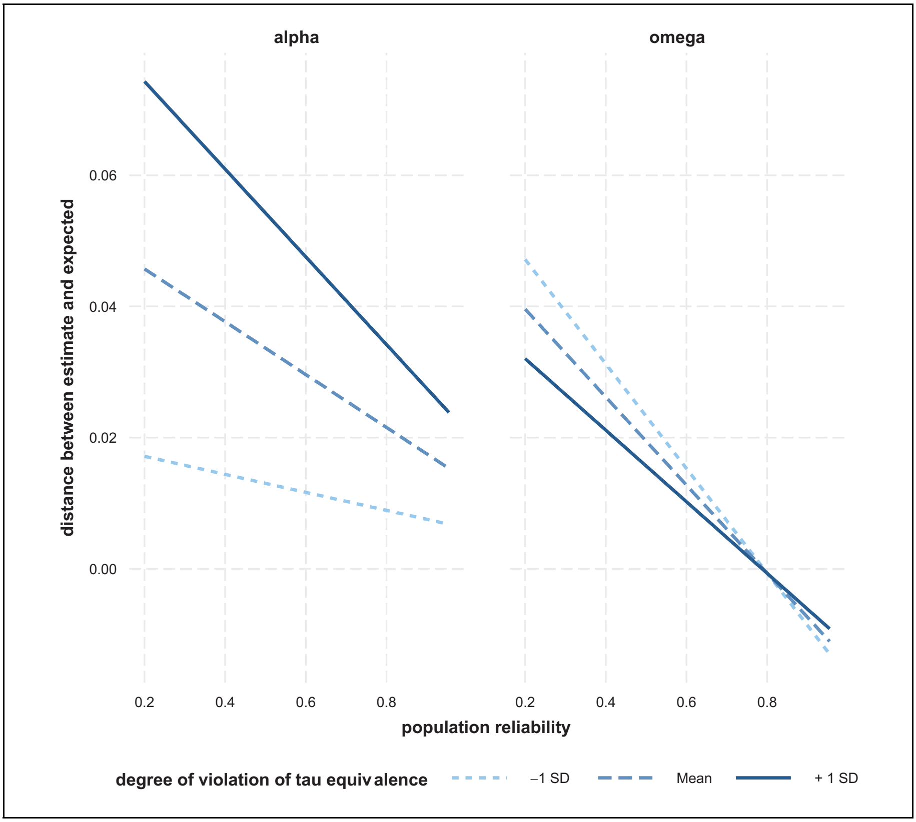

Figure 3. Regression estimated interaction between estimator, population reliability, and degree of violation of tau equivalence predicting the distance between estimate and expected reliability.

from normal distributions are all conditions that have been argued to affect alpha's reliability estimate (e.g., McNeish, 2018) that we did not include in the current work. Thus, our results cannot speak to how various reliability estimators will perform under those conditions. Furthermore, it may be the case that those conditions interact with the conditions tested here. For example, the effects of deviations from tau equivalence on the alpha estimate may be most pronounced when items also are not normally distributed. However, our current results tested several conditions under which alpha purportedly suffers most, and while the alpha estimate did show a negative bias in each case, it was actually a smaller bias than shown by omega hierarchical, Revelle's omega, and GLB estimators. Thus, future work should also test the performance of the internal consistency metrics estimated here under conditions of multidimensionality, correlated errors/residuals, and deviations from normal distributions. Additionally, while there are many scales constructed using continuous

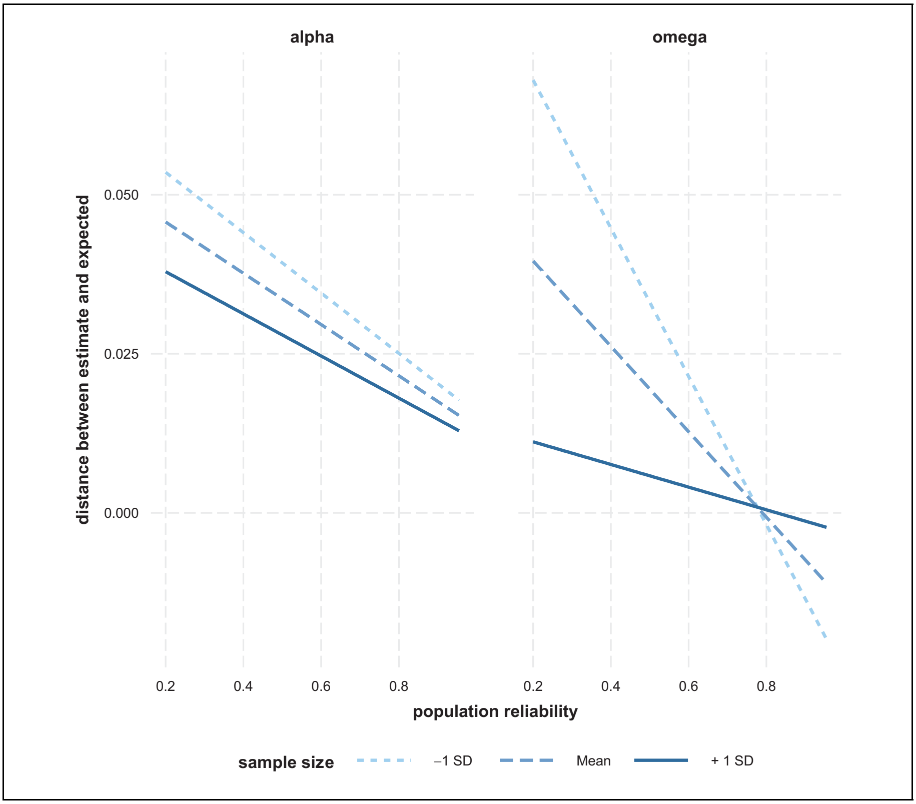

Figure 4. Regression estimated interaction between estimator, population reliability, and sample size predicting the distance between estimate and expected reliability.

indicators, it is also common in the psychological sciences to construct scales using ordinal items (e.g., Likert-type scales). Future work should thoroughly examine the effect that the ordinal nature of indicators has on reliability estimation.

Despite these limitations, the results of these simulations suggest that Cronbach's alpha estimates reliability fairly accurately under the condition of unidimensionality investigated here, in contrast with the significant criticism toward Cronbach's alpha in recent years (Crutzen & Peters, 2015; Dunn et al., 2014; Geldhof et al., 2014; Green & Yang, 2009; McNeish, 2018; Teo & Fan, 2013). While there are some conditions that may affect reliability estimation that were not examined here, the current results showed that under the conditions examined, alpha and omega deliver more accurate estimations of reliability than GLB, Revelle's omega, and omega hierarchical. Based on the results of this simulation, it is our recommendation that these alternatives to alpha and omega be limited until empirical work demonstrates conditions under which they perform adequately, and potentially discourage their widespread

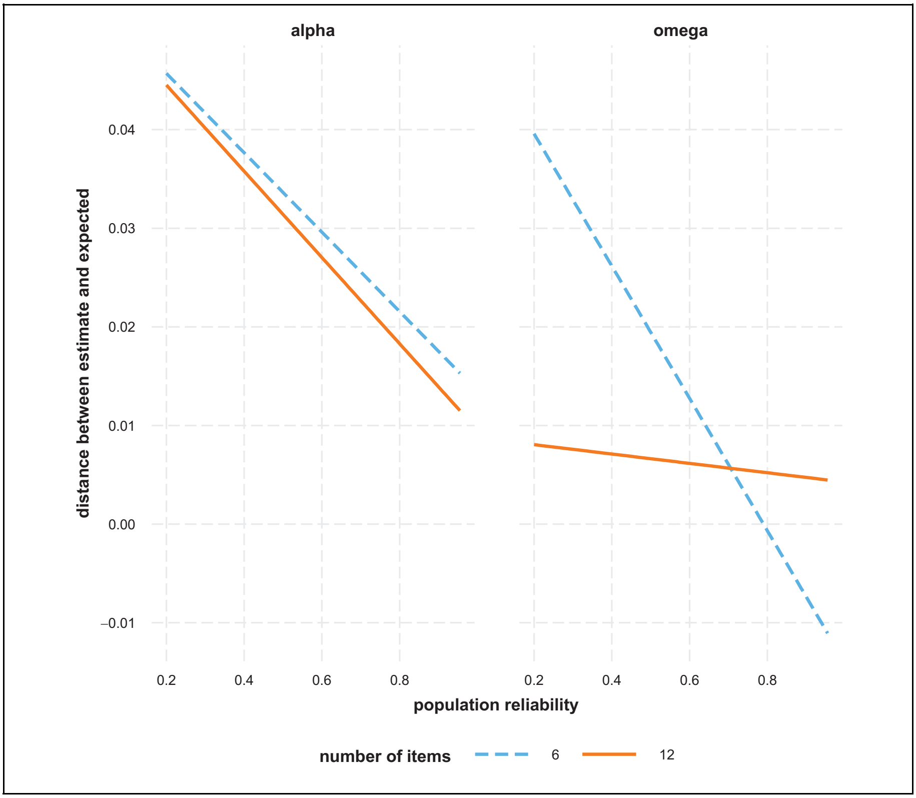

Figure 5. Regression estimated interaction between estimator, population reliability, and number of items predicting the distance between estimate and expected reliability.

adoption. The follow-up regression comparing alpha and omega revealed alpha to be sensitive to the degree of violation of tau equivalence, consistent with previous literature (Green & Yang, 2009). Novel to this investigation, however, was the discovery of the extent to which population reliability affects omega particularly when sample size and number of items are low.

Given that the computational demands of omega are higher, omega showed more estimation failures and greater distance between estimated and expected reliability when sample size, number of items, and population reliability were low. Thus, omega is not quite fit to completely replace alpha, contrary to recent arguments. In addition to conditions in which alpha provided a more accurate estimate than omega, alpha also provided a consistent underestimate in all conditions ensuring alpha to provide a lower bound estimate of internal consistency, whereas omega did not consistently have error in the same direction (sometimes overestimating and sometimes underestimating). Alpha still provides a useful estimate of internal consistency particularly when omega fails to provide an estimate or when sample size, number of items, and reliability is low under conditional of unidimensionality and uncorrelated errors.

Declaration of Conflicting Interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: KJJ was supported by a Ford Foundation Predoctoral Fellowship administered by the National Academies of Sciences, Engineering, and Medicine. CS was supported by Grant P50 HD052120 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

ORCID ID

Ashley A. Edwards https://orcid.org/0000-0002-5268-4235

Note

1. In two out of 140 conditions, alpha provided an overestimate less than .001.

References

Charter, R. A. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of Clinical and Experimental Neuropsychology*, *21*(4), 559-566. https://doi.org/10.1076/jcen.21.4.559.889

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334. https://doi.org/10.1007/BF02310555

Crutzen, R., & Peters, G. J. Y. (2015). Scale quality: Alpha is an inadequate estimate and factor-analytic evidence is needed first of all. *Health Psychology Review*, *11*(3), 242-247. https://doi.org/10.1080/17437199.2015.1124240

de Winter, J. D., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research*, *44*(2), 147-181. https://doi.org/10.1080/00273170902794206

Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, *37*(9), 830-837. https://doi.org/10.1046/j.1365-2923.2003.01594.x

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399-412. https://doi.org/10.1111/bjop.12046

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). American Council on Education and the National Council on Measurement in Education.

Fisher, R. A. (1950). *Statistical methods for research workers*. Oliver & Boyd.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72-91. https://doi.org/10.1037/a0032138

Green, S. B., & Yang, Y. (2009). Commentary on coefficient alpha: A cautionary tale. *Psychometrika*, *74*(1), 121-135. https://doi.org/10.1007/s11336-008-9098-4

Hancock, G. R., & An, J. (2020). A closed-form alternative for estimating ω reliability under unidimensionality. *Measurement: Interdisciplinary Research and Perspectives*, *18*(1), 1-14. https://doi.org/10.1080/15366367.2019.1656049

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9*(2), 139-164. https://doi.org/10.1177/014662168500900204

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement*, *60*(4), 523-531. https://doi.org/10.1177/00131640021970691

Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: I: Algebraic lower bounds. *Psychometrika*, *42*(4), 567-578. https://doi.org/10.1007/BF02295979

Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological methods*, *21*(1), 69-92. https://doi.org/10.1037/a0040086

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, *36*(4), 611-637. https://doi.org/10.1207/S15327906MBR3604_06

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84-99. https://doi.org/10.1037/1082-989X.4.1.84

McDonald, R. P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical & Statistical Psychology*, *23*, 1-21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x

McDonald, R. P. (1999). *Test theory: A unified approach*. Lawrence Erlbaum.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412-433. https://doi.org/10.1037/met0000144

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, *32*(1), 1-13. https://doi.org/10.1007/BF02289400

Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, *16*(2), 56-69. https://www.ehps.net/ehp/index.php/contents/article/view/ehp.v16.i2.p56

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*(4), 329-353. https://doi.org/10.1207/s15327906mbr3204_2

Raykov, T., & Marcoulides, G. A. (2015). A direct latent variable modeling based method for point and interval estimation of coefficient alpha. *Educational and Psychological Measurement*, *75*(1), 146-156.

Raykov, T., & Marcoulides, G. A. (2019). Thanks coefficient alpha, we still need you! *Educational and Psychological Measurement*, *79*(1), 200-210. https://doi.org/10.1177/0013164417725127

Revelle, W. (2019, June). *Using R and the psych package to find ω*. https://personality-project.org/r/psych/HowTo/omega.pdf

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*(1), 53-61. https://doi.org/10.1007/BF02289209

Shapiro, A., & ten Berge, J. M. F. (2000). The asymptotic bias of minimum trace factor analysis, with applications to the greatest lower bound to reliability. *Psychometrika*, *65*(3), 413-425. https://doi.org/10.1007/BF02296154

Sijtsma, K. (2009a). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107-120. https://doi.org/10.1007/s11336-008-9101-0

Sijtsma, K. (2009b). Reliability beyond theory and into practice. *Psychometrika*, *74*(1), 169-173. https://doi.org/10.1007/s11336-008-9103-y

Sijtsma, K. (2012). Future of psychometrics: Ask what psychometrics can do for psychology. *Psychometrika*, *77*(1), 4-20. https://doi.org/10.1007/s11336-011-9242-4

ten Berge, J. M. F., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*(4), 613-625. https://doi.org/10.1007/BF02289858

Teo, T., & Fan, X. (2013). Coefficient alpha and beyond: Issues and alternatives for educational research. *Asia-Pacific Education Researcher*, *22*(2), 209-213. https://doi.org/10.1007/s40299-013-0075-z

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. https://doi.org/10.1007/s11336-003-0974-7

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. https://doi.org/10.1177/0146621605278814