Author(s): Aguirre-Urreta, Miguel I.; Rönkkö, Mikko; McIntosh, Cameron N.

Title: A cautionary note on the finite sample behavior of maximal reliability

Year: 2019

Version: Accepted version (Final draft)

Please cite the original version:

Aguirre-Urreta, M. I., Rönkkö, M., & McIntosh, C. N. (2019). A cautionary note on the finite sample behavior of maximal reliability. Psychological Methods, 24(2), 236-252. https://doi.org/10.1037/met0000176

A Cautionary Note on the Finite Sample Behavior of Maximal Reliability

Miguel I. Aguirre-Urreta

Texas Tech University

Mikko Rönkkö

University of Jyväskylä

Cameron N. McIntosh

Public Safety Canada

Author Note

Miguel I. Aguirre-Urreta, Information Systems and Quantitative Sciences, Jerry S. Rawls College of Business Administration, Texas Tech University; Mikko Rönkkö, Department of Computer Science and Information Systems, University of Jyväskylä; Cameron N. McIntosh, Community Safety and Countering Crime Branch, Public Safety Canada.

This research was supported in part by a grant from the Academy of Finland (grant 311309). Parts of this research were carried out while Mikko Rönkkö was at the Department of Industrial Engineering and Management, Aalto University. We acknowledge the computational resources provided by the Aalto Science-IT project.

Correspondence concerning this article should be addressed to Miguel I. Aguirre-Urreta, Texas Tech University, Box 42101, Lubbock, TX 79409-2101. E-mail: miguel.aguirre-urreta@ttu.edu

Abstract

Several calls have been made for replacing coefficient α with more contemporary model-based reliability coefficients in psychological research. Under the assumption of unidimensional measurement scales and independent measurement errors, two leading alternatives are composite reliability and maximal reliability. Of these two, the maximal reliability statistic, or equivalently Hancock's *H*, has received a significant amount of attention in recent years. The difference between composite reliability and maximal reliability is that the former is a reliability index for a scale mean (or unweighted sum), whereas the latter estimates the reliability of a scale score where indicators are weighted differently based on their estimated reliabilities. The formula for the maximal reliability weights has been derived using population quantities; however, their finite-sample behavior has not been extensively examined. Particularly, there are two types of bias when the maximal reliability statistic is calculated from sample data: (a) the sample maximal reliability estimator is a positively biased estimator of population maximal reliability; and (b) the true reliability of composites formed with maximal reliability weights calculated from sample data is on average less than the population reliability. Both effects are more pronounced in small-sample scenarios (e.g., < 100). We also demonstrate that the composite reliability estimator for equally-weighted composite exhibits substantially less bias, which makes it a more appropriate choice for the small-sample case.

Keywords: maximal reliability, composite reliability, Hancock's *H*, omega total, weighted composites, simulations, small sample bias

A Cautionary Note on the Finite Sample Behavior of Maximal Reliability

Summary estimates of reliability for psychometric instruments serve two main purposes in primary research. First, they provide an assessment of the quality of measurement in latent variable models (Raykov, 1997). Second, in cases where regression or path analyses are performed using (weighted or unweighted) sums of scale scores, reliability estimates can be used to correct for attenuation of coefficients due to random measurement error (Cole & Preacher, 2014; Nimon, Zientek, & Henson, 2012). Outside of primary research, reliability estimates also feature centrally in correcting correlations for measurement error attenuation in meta-analyses (Schmidt & Hunter, 2015, Chapter 3). As such, obtaining unbiased estimates of reliability is of key interest in both primary and secondary studies.

Although coefficient α[1] continues to be the most popular reliability estimator in psychological research, its use rests on some strict measurement assumptions (e.g., tau-equivalence) that are unlikely to be met in practice (Cho & Kim, 2015; McNeish, 2017; Sijtsma, 2009). Consequently, model-based reliability estimators, such as maximal reliability (ρ_{max}, or equivalently, Hancock's H) and composite reliability (ρ_{CR}, or equivalently ω_{total}), have been recommended as more modern alternatives to alpha (McNeish, 2017). Like coefficient α, both of these estimators are applicable under scale unidimensionality and independence of measurement errors, but relax the assumption of equally reliable indicators required by coefficient α. Because both maximal reliability and composite reliability work with the same set of assumptions, these indices are often recommended as two equally preferable alternatives (McNeish, 2017).

Both maximal reliability and composite reliability estimate the reliability of a scale score or index formed as the sum of the indicators (composite), but differ in how the score is calculated. With composite reliability, the composite score is based on equal weights (i.e., an unweighted sum), whereas maximal reliability involves weighting each indicator differently to maximize the reliability of the composite score. Indeed, giving more weight to more reliable indicators is intuitively appealing, and a number of simulation studies have shown good performance of the maximal reliability estimator (e.g., Penev & Raykov, 2006; Raykov, 2004; Raykov, Gabler, & Dimitrov, 2015). However, these investigations have focused on the large sample context, which may not be representative of the bulk of applied research studies. This is problematic because there are important limitations to the small-sample applications of the maximal reliability index that are currently overlooked in the literature, as is evident from a number of recent articles that recommend its use without discussing these limitations (e.g., Dimitrov, Raykov, & AL-Qataee, 2015; Gabler & Raykov, 2017; McNeish, 2017).

In this research, however, we demonstrate the following two issues: (a) the maximal reliability estimator ($\hat{\rho}_{max}$) is a positively biased estimator of population maximal reliability; and (2) the true sample reliability (r_{max}) of composites formed with maximal reliability weights is less than the population maximal reliability. Moreover, both of these effects are most pronounced in small samples (e.g., < 100). For researchers looking for a successor to coefficient α for reliability estimation in latent variable models, the practical implication of the first result above is that when using maximal reliability, reliability will be over- rather than under-estimated, as would be the case with coefficient α. For those who prefer working with scale scores, these opposing effects are compounded, leading to overestimation of both population and sample reliability. That is, when calculating maximal reliability composites from small samples, applied researchers will form composites which will be, on average, markedly less reliable than what the calculated statistic would indicate. The equally-weighted composite and the associated composite reliability estimator, on the other hand, exhibit substantially less bias, which makes them more appropriate in small-sample scenarios for both of these applications.

We begin by reviewing the literature on the reliability of composites of congeneric measures, with an emphasis on the maximal reliability weights and the associated maximal reliability estimator. Thereafter, we show the equivalence of maximal reliability weights and regression coefficients calculated from factor analysis results, which leads to an equivalence of the maximal reliability estimator and the coefficient of determination. Based on this equivalence, we contend that the maximal reliability estimator is a positively biased estimator of population maximal reliability. We demonstrate this result by means of a simulation study comparing the relative performance of maximal reliability and equal weights composites, in terms of the relative degree of bias present in both estimators under a variety of research conditions (in particular, sample sizes). We conclude by discussing the implications of these results for the use of reliability indices with small samples.

Maximal Reliability

Congeneric Measures and Reliability

Following past research in this area (e.g., Penev & Raykov, 2006; Raykov, 2004), we are interested in a set of k congeneric items, which measure the same underlying latent variable with possibly different units and measurement locations, as well as random error variances (Jöreskog, 1971). For simplicity, we assume throughout this research that $k > 1$ and all location parameters are zero. The following relationship thus holds:

$$y_i = \lambda_i \eta + \varepsilon_i, \tag{1}$$

where each y_i is an indicator (i = 1 through k), η denotes the underlying latent variable, the λ_i are the loadings relating each indicator to the latent variable, and the ε_i are random measurement errors in the y_i, assumed to be uncorrelated with both the latent variable and among themselves.

The reliability of a weighted composite is the squared correlation between the composite and the latent variable that it measures, calculated as a non-linear function of the loadings, weights, and error variances of each indicator; more formally (e.g., H. Li, Rosenthal, & Rubin, 1996, Appendix B):

$$\rho_W = \frac{Var(\eta)\left(\sum_{i=1}^{k} w_i\lambda_i\right)^2}{Var(\eta)\left(\sum_{i=1}^{k} w_i\lambda_i\right)^2 + \sum_{i=1}^{k}[(w_i)^2 Var(\varepsilon_i)]}, \tag{2}$$

where $Var(\eta)$ is the variance of the latent variable, the w_i are the weights assigned to each indicator, the λ_i are the loadings as before, and the $Var(\varepsilon_i)$ are the variances of the measurement errors.

Maximal Reliability

The maximum reliability of a linear composite can be expressed in terms of individual indicator reliabilities (ρ_{i}, cf., Bentler, 2007; Conger, 1980; H. Li, 1997; Raykov, 2004):

$$\rho_{max} = \frac{\sum_{i=1}^{k}\frac{\rho_i}{1-\rho_i}}{1+\sum_{i=1}^{k}\frac{\rho_i}{1-\rho_i}}, \tag{3}$$

which is maximized when the weights of the individual indicators are defined as follows:

$$w_{max_i} = \frac{\rho_i}{\lambda_i(1-\rho_i)}. \tag{4}$$

Furthermore, given that the individual indicator reliabilities ρ_i can be expressed as:

$$\rho_i = \frac{\lambda_i^2}{\lambda_i^2+Var(\varepsilon_i)}, \tag{5}$$

the maximal reliability weights then become:

$$w_{max_i} = \frac{\lambda_i}{Var(\varepsilon_i)}. \quad (6)$$

That is, for a composite of congeneric measures to exhibit maximum reliability, the weight for each indicator should be set to the ratio of its loading to its error variance. The expression shown in Eq. 3 is equivalent to Hancock's coefficient *H* (e.g., Hancock, 2001, pp. 380–382; Hancock & Mueller, 2001), which figures prominently in discussions on power analysis and effect sizes when comparing latent means (Hancock, 2001)[2]. This index has received considerable attention in the past. Under either label, studies have examined the change in this index when the number of indicators of a latent variable changes (Raykov & Hancock, 2005), its use in multilevel designs (Geldhof, Preacher, & Zyphur, 2014; Raykov & Penev, 2009), the assessment of its invariance over time or in distinct populations (Raykov, 2005), and its role in statistical power in covariance structure modeling (Penev & Raykov, 2006), to name just a few. It is clear that this concept, in either of its formulations, has wide applicability in a number of different research scenarios as indicated by a recent call for the replacement of coefficient α by ρ_{max} and other model-based reliability statistics in psychological research (McNeish, 2017).

Although the maximal reliability estimator is most commonly referred to as ρ_{max} in methodological research, it is more commonly used under the alternative label of coefficient *H* in applied research. To gain a better understanding of the sample sizes involved in empirical applications of coefficient *H*, we reviewed all articles citing the work of Hancock and Mueller (2001) in Google Scholar resulting in a collection of 198 unique applied research articles written in the English language in which coefficient *H* was applied to sample data[3]. The median sample size was 344[4], indicating that the coefficient is most commonly applied to large samples. Of the 198 articles reviewed, however, 31 (or 15.6%) employed samples less than 150. An additional 43 (or 21.7%) studies had samples in the 150-250 range. Given that the maximal reliability estimator has been used with widely varying sample sizes in applied research and that the recent publications in premier methodological journals (e.g., Gabler & Raykov, 2017; McNeish, 2017; Raykov et al., 2015) advocate its use, it is important to understand the behavior of the statistic in finite samples, which is neither theoretically explained nor empirically demonstrated in any of the existing studies.

Finite-Sample Behavior of Maximal Reliability: An Explanation of the Bias

The maximal reliability weights are optimal (i.e., maximize reliability) when they are derived from population data where the scales consist of congeneric items with known reliabilities. However, the performance of the maximal reliability weights may be sub-optimal when these conditions do not hold. Particularly, researchers rarely work with full populations, and indicator reliabilities are typically unknown and must be estimated. In these scenarios, both sampling error and estimation errors of individual indicator reliabilities can substantially influence the performance of the maximal reliability estimator. Both sources of error depend on sample size, and we argue that there are small-sample scenarios that are problematic for the maximal reliability estimator, but which have not been thoroughly addressed in the literature.

Our concern is motivated by a large body of evidence demonstrating that, in small samples, several measures of association are biased estimators of their population counterparts, such as the Pearson correlation coefficient (*r*) and its square (Shieh, 2010; Skidmore & Thompson, 2011; Wang & Thompson, 2007; Zumbo, Williams, & Zimmerman, 2003), the coefficient of determination (R^2) in multiple regression (Leach & Henson, 2007; Shieh, 2008; Yin & Fan, 2001), and the canonical correlation coefficient (Leach & Henson, 2014; Thompson, 1990), to name a few. Moreover, this small-sample overestimation tends to be more pronounced when effect sizes are low (Thompson, 1999; Vacha-Haase & Thompson, 2004). Similar overestimation has also been shown to arise in other reliability estimates, such as the greatest lower bound (glb; Bentler, 2009; Ten Berge & Sočan, 2004). Although small-sample bias has neither been formally described nor systematically studied with respect to maximal reliability, it seems reasonable to expect that the sample maximal reliability estimator would not be unbiased, for two reasons. First, estimates of model parameters, such as those of λ_i, $Var(\eta)$, and $Var(\varepsilon_i)$ in Eq. 2, may be biased. In particular, commonly-used maximum likelihood estimates have been proven to be unbiased only asymptotically (Bollen, 1989; see also Cordeiro & Cribari-Neto, 2014; Kosmidis, 2014). Indeed, several authors have warned that reliability indices may not be trustworthy in small samples because factor analysis estimates may be biased (e.g., Raykov, 1997; Sijtsma, 2009). Second, even if both $\hat{\lambda}_i$ and $\widehat{Var(\varepsilon_i})$ were unbiased, estimation error in these statistics will lead to bias in the maximal reliability estimator, as we will explain next.

To see why the maximal reliability estimator is biased even when unbiased estimates are used in its calculation, it is helpful to consider it within the framework of multiple regression. The multiple regression model can be written as follows (where all variables are expressed as deviations from their means, with no loss of generality):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mu}, \tag{7}$$

where **y** is a $N \times 1$ vector of response values, **β** is a $k \times 1$ vector of regression coefficients, **X** is a $N \times k$ matrix of observations for the predictors, and **μ** is a $N \times 1$ vector of errors, which are assumed to be uncorrelated with the predictors. The ordinary least squares (OLS) estimate of **β** is calculated as:

$$\widehat{\boldsymbol{\beta}} = \mathbf{S}_{xx}^{-1}\mathbf{s}_{xy}, \tag{8}$$

where $\mathbf{S}_{xx}^{-1}$ is the inverse of the $k \times k$ sample variance-covariance matrix of the k predictors and $\mathbf{s}_{xy}$ is a $k \times 1$ vector of the sample covariances between each predictor and the response variable. The coefficient of determination for the regression in Eq. 7 is given by:

$$R^2 = \widehat{\boldsymbol{\beta}}^T \mathbf{s}_{xy}. \tag{9}$$

There is a direct correspondence between the maximal reliability estimator and the coefficient of determination in multiple regression shown in Eq. 9. This equivalence derives from the fact that the fitted values of the regression, $\hat{\mathbf{y}} = \widehat{\boldsymbol{\beta}}\mathbf{X}$**,** are composites of the predictors that are maximally correlated with the dependent variable. Now, consider a regression of a latent variable on its indicators:

$$\boldsymbol{\eta} = \mathbf{Y}\boldsymbol{\beta} + \boldsymbol{\mu}, \tag{10}$$

where $\boldsymbol{\eta}$ is a $N \times 1$ vector of values of the common factor, $\boldsymbol{\beta}$ is the $k \times 1$ vector of regression coefficients, $\mathbf{Y}$ is the $N \times k$ matrix of observations for the indicators, and $\boldsymbol{\mu}$ is the $N \times 1$ vector of uncorrelated errors. Although it would seem that estimating this regression would not be possible because no case values of $\boldsymbol{\eta}$, are observed, the regression coefficients $\boldsymbol{\beta}$ can be estimated because the OLS estimator in Eq. 8 requires only information about the sample variances and covariances. Although the sample covariances between a latent and an observed variable are not observed, the OLS estimator can be applied to model-implied covariances. Assuming that the latent variable has unit variance in the sample, the model-implied covariances between the latent variable and its indicators are equivalent to the estimated factor loadings ($\hat{\boldsymbol{\lambda}}$). Further, the model-implied variance-covariance matrix of the indicators is given as:

$$\widehat{\boldsymbol{\Sigma}}_{yy} = \hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\lambda}}^T + \widehat{\boldsymbol{\Theta}}, \tag{11}$$

where $\widehat{\boldsymbol{\Theta}}$ is a diagonal $k \times k$ matrix of estimated error variances ($\widehat{Var(\varepsilon_\iota)}$). The maximal reliability weights can therefore be expressed as[5]:

$$\mathbf{w}_{max} = \widehat{\boldsymbol{\Sigma}}_{yy}^{-1}\widehat{\boldsymbol{\lambda}}. \tag{12}$$

The maximal reliability weights discussed here are equivalent to the weights used to create factor scores proposed by Thurstone (1935) (compare, for example, Eq. 11 above with Eq. 1 in Grice, 2001). The estimated maximal reliability is then given by:

$$\hat{\rho}_{max} = \mathbf{w}_{max}^{T}\widehat{\boldsymbol{\lambda}}, \tag{13}$$

which is equivalent with the coefficient of determination in Eq. 9 because of equivalence between $\widehat{\boldsymbol{\beta}}$ and $\mathbf{w}_{max}$, and $\widehat{\boldsymbol{\lambda}}$ and $\mathbf{s}_{xy}$.

This derivation is novel, as the equivalence of both quantities has not thus far been recognized in the literature on maximal reliability. Having established the equivalence of both formulations of the maximal reliability estimator, we now focus on its behavior when applied to finite samples. In an ideal scenario where the factor model holds in the population and $\hat{\lambda}_i$ and $\widehat{Var(\varepsilon_i)}$ are unbiased and independent, the model–implied moments of $\widehat{\boldsymbol{\sigma}}_{xy}$ and $\widehat{\boldsymbol{\Sigma}}_{xx}$ are unbiased estimators of their population counterparts, thereby yielding unbiased estimates of the coefficients in the regressions of factors on observed indicators (as per Eq. 8 for the OLS estimator), which are equivalent to maximal reliability weights except for a scaling difference. The relevance of the OLS regression framework for the small-sample characteristics of the maximal reliability estimator stems from the fact that even under this ideal scenario where OLS estimates of regression coefficients are unbiased, the R^2 statistic calculated from sample data is positively biased (Cohen, Cohen, West, & Aiken, 2003, Chapter 3.5). Given the important effects of sample size on the degree of bias in the R^2 statistic (Yin & Fan, 2001), we also expect bias in the maximal reliability estimator to diminish as sample size increases.

True and Estimated Composite Reliability

Although the maximal reliability estimator is typically used as an estimator of population maximal reliability, the composite scores may also be of interest in some research settings. Therefore, two important questions are: (a) how reliable are the maximal reliability composites calculated from sample data; and (b) how does this reliability compare to the population maximal reliability? The sample reliability (r) of a composite is defined as the squared correlation of that composite with the case values of the underlying latent variable in the sample; Table 1 provides the definitions of all reliability values considered in this article. As stated above, the maximal reliability composites are equivalent to the factor score estimators proposed by Thurstone (1935), which have been shown to become more weakly correlated with the true latent variable values as sample size decreases (Grice, 2001). Therefore, the expected value of the sample reliability of a composite formed with maximal reliability weights, r_{max}, is anticipated to be systematically smaller than the population maximal reliability, ρ_{max}.

Given that the two effects described in this section operate in opposite directions, researchers using composites created with maximal reliability weights in small samples will be faced with a dual challenge. On the one hand, their estimate of the reliability of the composite will be, on average, higher than the population value of the same. On the other hand, the sample reliability of said composite can be expected to be lower than the population value:

$$\mathbb{E}[r_{max}] < \rho_{max} < \mathbb{E}[\hat{\rho}_{max}]. \tag{13}$$

Taken together, the net effect is a weighted composite that will appear much more reliable than is truly the case, that is, the total discrepancy between $\mathbb{E}[r_{max}]$ and $\mathbb{E}[\hat{\rho}_{max}]$ widens with decreasing *N*, as we demonstrate later in our simulation experiments.

The two sources of bias can be intuitively attributed to how the reliability estimates affect the calculation of indicator weights. On the one hand, indicators whose reliability estimates are affected by estimation error in a positive direction are given larger weights, leading to positive bias in $\hat{\rho}_{max}$. On the other hand, the calculated weights are unlikely to be exactly optimal, because indicator reliability estimates are not error-free. Therefore, we can expect that the use of equal weights and the associated composite reliability estimator would mitigate these biases such that the sample reliability, estimated reliability, and population reliability would be approximately equal (i.e., $\mathbb{E}[r_{CR}] \cong \rho_{CR} \cong \mathbb{E}[\hat{\rho}_{CR}]$). To be sure, we cannot expect the composite reliability estimator to be completely unbiased, as it is a ratio (i.e., proportion of unit-weighted true score variance over observed score variance), and these are generally biased to some extent even if their input quantities are not. Nevertheless, the composite reliability estimator should still be less biased than the maximal reliability estimator, because it uses equal, fixed weights rather than estimating them from sample data. However, the relative small-sample performance of the two reliability coefficients has not yet been systematically examined. Indeed, Raykov and colleagues (Raykov et al., 2015; Raykov & Marcoulides, 2016) emphasize the large-sample nature of model-based reliability coefficients, noting that factor analysis estimators such as maximum likelihood are only asymptotically unbiased, and call for future work on establishing the sample sizes where researchers could safely rely on asymptotic theory. Although our formal analysis above reveals that the maximal reliability estimator is, in principle, biased within finite samples, the actual extent of the problem in any given modeling situation is unknown. Therefore, we conducted a series of Monte Carlo simulation experiments to address these gaps.

Simulation Design and Results

Our set of Monte Carlo simulations consisted of the following specifications. First, we generated data following a multivariate normal distribution from a model containing a single common factor measured with 3, 5, 7, or 9 indicators[6]. In one set of conditions, the loadings of each indicator on the common factor were all equal at 0.6, 0.7, or 0.8. In a second set of conditions, the loadings were unequal but centered around the same base loadings as before, as follows. First, a vector containing all elements from 1 to the desired number of indicators (i.e., 3, 5, 7, or 9) was constructed. Second, the vector was standardized to zero mean and unit variance. Third, the resulting vector was multiplied by a factor of 0.10. Finally, the base loading for the condition (0.6, 0.7, or 0.8) was added. The full listing of all loadings employed in the unequal loadings condition is included in Table 2. For each scenario, measurement errors were specified to give the indicators unit variance. We also varied the size of the generated sample at 25, 50, 75, 100, 150, 300, 500, and 1000. All conditions were fully-crossed in a factorial design for a total of 192 different simulated conditions: number of indicators (4) x base loadings (3) x loading equality (2) x sample size (8), with 10,000 replications each. Data were generated by first generating a vector of latent variable scores, which was then multiplied by the factor loading vector after which random error was added. All data generation and parameter estimation steps were carried out in the *R Statistical Environment* (R Core Team, 2016); factor models were estimated with the *lavaan* package (Rosseel, 2012), using the maximum likelihood technique and a correctly specified model for all scenarios. Annotated simulation and analysis code is available in Appendix D.

Simulation Results

Convergence was not a major issue, which is to be expected because the model was correctly specified (convergence rates, per unique combination of conditions, ranged from 93% to 100%), and therefore we simply discarded the non-convergent replications. Within those replications that converged, in some conditions we encountered several inadmissible solutions due to Heywood cases. These were much more prevalent in the small sample size conditions, rising to as much as 28% of the converged replications for the case of $N = 25$, 3 indicators, and unequal loadings with a base loading of 0.60; in this condition, only 67% of all replications were both convergent and admissible. Whereas Heywood case can be an indication of model misspecification, they can also be a result of sampling variability of the variance estimates in a correctly specified model (Kolenikov & Bollen, 2012). Thus, after concluding that the Heywood case is not to be interpreted as evidence of model misspecification, researchers often re-estimate the model with a constraint that the error term is positive (Savalei & Kolenikov, 2008). However, as explained by Savalei and Kolenikov, this constraint leads to problems for model test statistics and may therefore be suboptimal. Constraining error terms to be positive is problematic for the maximal reliability statistic as well. Although estimates obtained this way satisfy the assumption that the error variances are positive, as required for the calculation of the maximal reliability weights, the estimates obtained this way converge to values that are arbitrarily close to zero and consequently $\hat{\rho}_{max}$ is arbitrarily close to one[7]. Therefore, following earlier research on factor score estimation (e.g., Grice, 2001), we discarded all replications that contained Heywood cases[8].

We calculated the maximal and composite reliability estimators and also created composites of the observed indicators using either the maximal reliability or equal weights. Taking the squared correlation between these composites and the latent variable values used in data generation gives the true reliability of each of those composites in the sample: r_{max} and r_{CR}. This allowed us to compare maximal and composite reliability estimators to r_{max} and r_{CR}, as well as ρ_{max} and ρ_{CR}.

For the case of unequal population loadings, the overall degree of bias in $\hat{\rho}_{max}$ and r_{max} was, respectively, 1.4%, and -1.5%, when compared to the population maximal reliability; in the case of equal population loadings, these biases were 1.6% and -1.6%, respectively. The performance of CR was identical over the equal and unequal loading conditions, and the biases for $\hat{\rho}_{CR}$ and r_{CR} were 0.0% and -0.1%, respectively. The seemingly small global bias in the maximal reliability case, however, obscures some specific scenarios in which the bias is more prominent. As expected, these are the simulation conditions with the poorest measurement conditions; that is, those with the fewest indicators, the smallest samples and the weakest loadings. Conversely, bias decreases with increasing sample size, more indicators, and stronger loadings.

Tables 3 and 4 show the results of selected conditions for the unequal and equal loading scenarios, respectively (full results for all conditions are presented in Appendixes B and C). The results show that in several scenarios, the maximal reliability estimators are severely affected by bias, whereas the composite reliability estimators are relatively unbiased. It is also critical to realize that the biases demonstrated here for $\hat{\rho}_{max}$ and r_{max} operate in opposite directions. To see how this compounds the problem, consider the case of equal loadings, sample size of 25, 3 indicators, and loadings of 0.6 (the first row of Table 4). In this scenario, the long-run average of $\hat{\rho}_{max}$ is 18.8% *higher* than the population value of the index, whereas r_{max} is, on average, 8.7% *lower* than the population value. This leads to an overall 27.5% (18.8% + 8.7%) discrepancy between the estimated and true sample reliability of a composite. Similar calculations can be made for all other results presented in Tables 3 and 4. Therefore, even if the biases present in $\hat{\rho}_{max}$ and r_{max} may not be individually large, their opposing nature (i.e., $\hat{\rho}_{max}$ is a positively bias estimator of ρ_{max} and r_{max} on average falls below ρ_{max}) leads to composites that are believed to be much more reliable than is truly the case. That is, the estimated reliability that researchers observe from a composite that was created with maximal reliability weights is positively biased, compared to the population value, whereas the true reliability of the same composite is negatively biased, compared to the population value.

To more clearly show the effects of the various design conditions on the maximal reliability estimates, we present next a series of plots of the estimates and population values for equal and unequal loadings under selected conditions. Figure 1 shows maximal reliability estimates for the case of 3 indicators and 0.6 loadings over a range of sample sizes, for the equal [Panel (a)] and unequal [Panel (b)] loadings conditions. Figure 2 shows the same estimates for the case of 3 indicators, a sample size of 25, and over the range of loading values examined here. Finally, Figure 3 shows the estimates for the case of 0.6 loadings, sample size of 25, and indicators from 3 through 9.

As these Figures show, both $\hat{\rho}_{max}$ and r_{max} converge on ρ_{max} as measurement conditions improve. That is, for any given value of the loadings and the number of indicators, results obtained from larger samples will be closer, on average, to the population values than those obtained from smaller samples. The same pattern occurs when the sample size and number of indicators are held fixed, with the results improving as loading strength increases; when sample size and loading strength are fixed, unbiasedness improves as the number of indicators increases. A comparison between the two panels in each figure reveals that these effects operate similarly in both the equal and unequal loading conditions.

Figures 2 and 3 showed results for $N = 25$, to showcase the maximal reliability bias in its most extreme. To address more realistic sample sizes, Figures 4 and 5 display the same results as in Figures 2 and 3 but with $N = 150$. As these show, the bias in maximal reliability estimates, though improving as measurement conditions improve (e.g., more indicators and stronger loadings), is still noticeable. That the performance of maximal reliability increases raises the question of how many indicators are necessary for the bias to become negligible. Although our focus on 3- to 9-indicator scales covers the majority of empirical research, scales that are much longer are still fairly common in psychological assessment and personality research. To address this scenario, we ran a small additional simulation with only the poorest measurement conditions of those included in our original design (sample size of 25 or 50, base loadings of .6 or .7) but increased the number of items to 20. Results from this simulation run (not reported, but available from the first author) revealed minimal bias (e.g., in the 3% range), though still in the expected direction.

Turning now to the case of equally-weighted composites, we note that $\hat{\rho}_{CR}$, as well as r_{CR} of equal-weighted composites, are largely unbiased with respect to ρ_{CR}. In contrast with the results presented above for maximal reliability, the largest discrepancy between $\hat{\rho}_{CR}$ and r_{CR} is only 3.7% (also in the first row of Table 4). Moreover, Figures 6, 7, and 8 parallel those for maximal reliability results shown above in Figures 1, 2, and 3. As these clearly show, the composite reliability estimates do not suffer from the same issues as the maximal reliability procedure. In particular, the $\hat{\rho}_{CR}$ are quite close to both the population ρ_{CR} as well as the sample r_{CR}. Tables 3 and 4 (for selected conditions), as well as Appendixes B and C (full results), provide further evidence of the lack of bias exhibited by both $\hat{\rho}_{CR}$ and r_{CR}. That is, on average, the estimated composite reliability that researchers observe from their results is very close to the true reliability of the composite, thus making this an appealing alternative for applied scenarios.

Finally, we also include, in Appendixes B and C, a calculation of the relative bias of each estimator, by taking the difference between the estimated value of the reliability of a composite and its true sample reliability, which can be calculated from the case values of the latent variables that were generated as part of the simulation (but which would not be available to applied researchers). These results show that bias in the maximal reliability estimator is substantial under suboptimal measurement conditions and always exceeds that of the composite reliability estimates, which were largely unaffected.

Separating the Sources of Bias in the Maximal reliability estimator

Bias in the maximal reliability estimator stems from both: (a) the small-sample bias in the maximum likelihood estimators of the factor model parameters used as input (i.e., estimated loadings and error variances); and (b) bias arising from the reliability maximization formula itself. Particularly, in our simulation, the estimated error variances were negatively biased. Therefore, it is fair to ask how large the relative contributions of the two sources of bias might be in a given application. In finite samples, does the bias in the maximum likelihood estimates dominate over bias engendered by the maximal reliability formula, or vice versa? To address this question, we extracted all loading and error variance estimates from the simulations and centered them at their population values – thereby artificially removing bias from the estimates – followed by recalculating all reliability estimates and composites. The results for these additional analyses are shown in Tables 5 and 6

The maximal reliability estimates obtained from this alternative approach therefore reflected the bias specifically due to the maximal reliability weights rather than that due to maximum likelihood estimation of the factor model parameters. Comparing the results in Table 5 and Table 6 against the corresponding results in Table 3 and Table 4 reveals that on average, about two thirds of the bias of maximal reliability estimates remain even when the bias of the loading and error variance estimates was eliminated. The sample reliabilities of the maximal reliability composites also increased slightly, which is explained by the fact that our centering procedure increased the average error variance estimates, thereby mostly eliminating scenarios where one indicator was substantially overweighted at the expense of other indicators, due to having an error variance estimate close to zero.

As expected, eliminating bias from the factor model estimates also affected the composite reliability estimates, leading to a negative bias. However, even under the poorest modeling conditions examined here – $N = 25$, 3 indicators, and all loadings at 0.60 – the bias of the composite reliability estimator was only -1.7%, as compared with 10.2% for the maximal reliability estimator. The sample reliabilities of the unit-weighted composites are unaffected because they do not depend on the loading estimates.

Empirical Illustration

We now provide an empirical illustration and comparison of the two techniques. For this purpose, we employ the large personality dataset from Johnson (2014)[9]. These data were collected as part of a large-scale effort aimed at the further development and validation of the IPIP-NEO personality scales. The full inventory consists of 300 items, of which we focus on the Assertiveness subscale of the Extraversion construct (the E3 scale in the original research), comprised of four items (I12: “Take charge”, I42: “Try to lead others”, I72: “Take control of things”, and I102: “Wait for others to lead the way”; the latter item is reverse-coded), measured on a 5-point scale. This scale is a part of a reduced IPIP-NEO-120 inventory for which Johnson has collected a sample of *N* = 619,150. Because some of the observations contained missing data for the variables of interest, we use only complete cases, reducing the sample size to *N* = 605,607. Out of this reduced sample, we took random samples of *N* = 25, *N* = 50, *N* = 100, *N* = 500, and *N* = 1000, to showcase the performance of both reliability indices. Table 7 presents the standardized loading and error variance estimates for each of these analyses. These can be plugged into Eq. 2 using either maximal reliability weights (e.g., Eq. 6) or equal weights to calculate the maximal and composite reliability indices, respectively. As expected, the estimate of maximal reliability is always higher than that of composite reliability and this difference is larger in smaller samples, consistent with our claim that $\hat{\rho}_{max}$ is positively biased in small samples.

Discussion and Conclusions

In this research, we examined the finite-sample performance of the maximal reliability estimator as well as that of composites formed with maximal reliability weights. Although prior literature has shown that the use of weights equal to the ratio of the loading to the measurement error variance for each indicator produces composites with maximal reliability, these expressions were derived in the context of population values. However, our analysis of the finite sample behavior of the maximal reliability estimator reveals a positive bias, and that the actual sample reliability of the maximal reliability composites falls short of population maximal reliability. These biases are much more pronounced with small sample sizes (e.g., < 100), tending to diminish as sample size increases and being minimally noticeable with samples larger than 1,000. In comparison, the reliability of equally-weighted composites was unaffected by sample size and the associated composite reliability coefficient was a virtually unbiased estimator of the population composite reliability. Moreover, we showed that the finite-sample bias in the maximal reliability estimator is not simply attributable to bias in the maximum likelihood estimators of the input quantities (i.e., estimates of loadings and error variances), but can be mostly attributed to the maximal reliability formula itself.

This investigation contributes to the existing literature in two ways. First, while previous studies have noted that reliability estimates can be biased in small samples because the factor analysis estimates used as input for their calculation can be biased (e.g., Raykov et al., 2015; Raykov & Marcoulides, 2016; Sijtsma, 2009), how estimation error of the indicator reliabilities impacts maximal reliability has not been addressed. Using a multiple regression framework, we provide the first statistical explanation of small-sample bias in the maximal reliability estimator. Second, our results extend the recent research by Raykov et al. (2015) on choosing between maximal and composite reliability. Whereas Raykov and colleagues presented a technique for testing the equivalence of the two indices in a given application, they do not present any explicit recommendations on which reliability estimator and set of indicator weights should be used. Similarly, McNeish (2017), while recommending both $\hat{\rho}_{MR}$ and $\hat{\rho}_{CR}$ (under labels Coefficient H and ω_{total}) as "successors" to coefficient α in psychological research, refrains from making recommendations or providing any criteria for choosing between the two indices in a given application.

To answer the question of which of the two indices is more appropriate for applied research, we need to consider how they differ and the two different purposes of these reliability estimates: (a) as an overall indicator of measurement quality in latent variable models; and (b) as a measure of reliability when working with actual composites (e.g., aggregate test scores). The only difference between the two reliability indices is in how the indicators are weighted in their calculation. When indicator reliability is homogeneous – that is, all indicators of a latent variable are equally reliable – both indices will lead to identical values in the population. However, the indices are not equivalent when calculated from sample estimates because estimation error makes the individual reliability estimates, and consequently the weights used in the maximal reliability estimator, uneven. On the other hand, when indicator reliability is heterogeneous – that is, when some indicators are more reliable than others – this should lead to performance advantages for the maximal reliability approach, as it is designed to exploit heterogeneous reliabilities by differentially weighting the indicators in its calculation (cf., McNeish, 2017).

When used as a model quality index in latent variable modeling, it may not matter so much how reliable the composites are compared to what could be constructed in an ideal case. Rather, the more important issue is that that the estimator performs in a predictable manner, independently of sample size (i.e., $\hat{\rho}_{max}$ should be an unbiased estimator of ρ_{max}). This is particularly relevant for the equal loadings scenarios, where the maximal and composite reliability estimators converge asymptotically, and therefore the choice between the techniques is made solely based on their finite-sample performance. As our results show, in this case the composite reliability estimator should be preferred throughout all examined conditions, as the estimated composite reliability is uniformly closer to its population value. This suggests that the current practice of relying mostly on the composite reliability estimator (Raykov et al., 2015) may in fact be superior to the maximal reliability estimator for this purpose.

When working with actual composites, the choice of the reliability estimator is a part of the more general choice of how the indicators should be weighted, which is a more general question preceding any summary assessment of reliability. Although weighting the indicators to maximize reliability is in principle appealing, a large body of previous research provides strong support for the robustness and utility of equal weights and unit weights in particular (e.g., Bobko, Roth, & Buster, 2007; Cohen, 1990; Cohen et al., 2003, pp. 97–98; Raju, Bilgic, Edwards, & Fleer, 1999). The same conclusion can be drawn from our results as well: in the population with unequal indicator reliabilities, the discrepancy in reliability between maximal reliability and equally-weighted composites was very small; and in small samples, equal weights produced more reliable composites than the maximal reliability weights calculated from sample data. Therefore, the fact that the maximal reliability statistic is larger than composite reliability calculated from the same sample cannot always be interpreted as evidence that the maximal reliability composites provide a meaningful advantage over equally-weighted sums (cf., McNeish, 2017). Indeed, although these simpler composites may not be optimal in the population, their robustness to sampling variability may make them more appropriate for use with small samples.

When calculating composites, the primary concern is not whether the reliability statistics are unbiased estimators of their population counter parts, but how accurately they estimate the sample reliability of the composites (i.e., $\hat{\rho}_{max}$ should be an unbiased estimator of r_{max}). A potential bias has important implications not only for interpreting research findings in the presence of measurement error, but also for those instances where reliability information is used to correct for attenuation both in primary studies (Cole & Preacher, 2014) and in meta-analyses (Nimon et al., 2012; Schmidt & Hunter, 2015, Chapter 3). In this situation, the maximal reliability estimator is potentially more problematic because the results are affected by two sources of bias, $\hat{\rho}_{max} > \rho_{max}$ and $r_{max} < \rho_{max}$, and may make composites appear much more reliable than is truly the case, leading to undercorrection for attenuation. Whereas this bias may not be very large in many of the conditions examined here, it can be avoided altogether by using the composite reliability estimator, which is essentially unbiased.

Given that our simulation considered a variety of other factors in addition to sample size (e.g., loading strength, number of indicators, and loading heterogeneity), we can offer some practical counsel to applied researchers on the situations in which the bias in maximal reliability estimates would likely manifest in sufficient strength to make its use unadvisable. The full simulation results, presented in Appendix B, provide applied researchers with the ability to look up the scenario that most closely resembles their particular research conditions in order to make an informed decision on which reliability estimator to use. When working with composites, a sum of the two biases needs to be considered and compared against the amount of bias one is willing to tolerate. As our results show, there are conditions where the use of maximal reliability estimates and composites is not advisable, and some of these include what would traditionally be considered large samples. For example, scenarios with unequal loadings, 3 indicators, base loading strength of 0.6 and up to (and including) samples of $N = 150$ are problematic if one requires less than 5% bias. With a more conservative 1% level, even 1000 observations are insufficient. Although unbiasedness improves as measurement conditions improve, our results show that the maximal reliability estimates and weights can hardly be recommended with poor measurement conditions and sample sizes under $N = 100$.

Although our results are supportive of the use of the composite reliability statistic as a general reliability estimate, it is important to point out that both composite reliability and maximal reliability assume a congeneric measurement model, both these statistics, which in practice means that the indicators form a unidimensional scale and measurement errors are independent (McNeish, 2017). In practice, this assumption means that after partialling out the common factor, the items are independent and there are no subdimensions in the scale. Such scale homogeneity would require that item similarity or positioning in the instrument has no effect on their correlations (Kelley & Pornprasertmanit, 2016), and would imply that arbitrary reweighting of the scale items or even item elimination would only affect the reliability but not the validity of the composite score. In other words, adding or removing items can alter the accuracy of the scale, but not its nature. Clearly, although not as stringent as the tau-equivalence assumptions of coefficient α, the congeneric model is still relatively rigid and may not always hold in real applications. For instance, items may load on more than one factor (Marsh, Morin, Parker, & Kaur, 2014), and measurement errors may intercorrelate for methodological and/or substantive reasons (Cole, Ciesla, & Steiger, 2007; Reddy, 1992).

From a practical perspective, the question is therefore not whether the scale is exactly unidimensional, but whether it is sufficiently unidimensional for the congeneric measurement model to be a useful approximation (Rodriguez, Reise, & Haviland, 2016). If strict unidimensionality does not hold, the weaker assumption of essential unidimensionality (e.g., Yang & Green, 2011) may be applied if the items measure one main dimension but are contaminated with other sources of nuisance variation, such as item context effects. In this case, bifactor models, where each indicator loads on both a general and a group factor, accompanied by the hierarchical omega reliability statistics can be useful (see McNeish, 2017; Rodriguez et al., 2016). Rodriquez et al. explain several tests that can be applied to evaluate whether this alternative approach is necessary.

Finally, although our results can be interpreted as indicating that the usefulness of the maximal reliability estimator can be limited in applied research settings, particularly when working with small samples and a limited number of indicators, the source of bias presented in this article also opens up an avenue for the development of bias corrections for maximal reliability. Indeed, a number of studies have examined finite-sample bias in various model-free composite reliability estimators (e.g., Benton, 2015; Osburn, 2000; Shapiro & Ten Berge, 2000; Ten Berge & Sočan, 2004), which has led to the development of successful corrective procedures (e.g., Hunt & Bentler, 2013; J. Li & Bentler, 2011; Shapiro & Ten Berge, 2000). Given the equivalencies shown here between the maximal reliability estimator and R^2, the numerous techniques that have been developed in the multiple regression context for obtaining adjusted R^2 values could potentially be adapted to the maximal reliability case (Shieh, 2008; Walker, 2007). Our preliminary examination of the formulas included in Shieh (2008) and bootstrapped versions of these corrections (J. Li & Bentler, 2011), reported in Appendix G, indicates that although there is some improvement in correcting the positive bias shown by $\hat{\rho}_{max}$ (as the correction formulas are intended to reduce the estimates), the improvements are mostly of limited magnitude, but in some simulation conditions result in negative bias. Moreover, even with the corrections, the bias of $\hat{\rho}_{max}$ remains generally larger than the bias of $\hat{\rho}_{CR}$. In addition, these corrections do not address the other issue discussed here, namely that r_{max} is negatively biased, because they only focus on the correction of the estimated maximal reliability. Therefore, future work should examine the potential for other bias correction approaches to the estimation of model-dependent composite reliability based on maximal reliability weights. One such clear example is the work of Penev and Raykov (2010), who developed a correction formula that can be used with composites of dichotomous indicators and their maximal reliability estimate, with an additional application for the calculation of confidence intervals for the same. Clearly, much more work is needed in this area in order to develop, and further validate, correction procedures that take into account the dual issues of: (a) an estimate of maximal reliability ($\hat{\rho}_{max}$) that is positively biased compared to its population value; and (b) the negative bias in the true reliability of the composites compared to population maximal reliability. However, until such corrective procedures are developed and extensively tested, researchers are advised to be cautious about the use of the maximal reliability approach with small samples.

References

Bentler, P. M. (2007). Covariance structure models for maximal reliability of unit-weighted composites. In S. Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 1–19). New York, NY: Elsevier.

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, *74*(1), 137–143. https://doi.org/10.1007/s11336-008-9100-1

Benton, T. (2015). An empirical assessment of Guttman’s lambda 4 reliability coefficient. In R. E. Millsap, D. M. Bolt, L. A. van der Ark, & W.-C. Wang (Eds.), *Quantitative psychology research* (Vol. 89, pp. 301–310). Heidelberg, Germany: Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-07503-7_19

Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: A literature review, application to content validity, and meta-analysis. *Organizational Research Methods*, *10*(4), 689–709. https://doi.org/10.1177/1094428106294734

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

Cho, E., & Kim, S. (2015). Cronbach’s coefficient alpha well known but poorly understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994

Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, *45*(12), 1304–1312. https://doi.org/10.1037/0003-066X.45.12.1304

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. London, UK: Lawrence Erlbaum Associates.

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The Insidious Effects of Failing to Include Design-Driven Correlated Residuals in Latent-Variable Covariance Structure Analysis. *Psychological Methods*, *12*(4), 381–398. https://doi.org/10.1037/1082-989X.12.4.381

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–315.

Conger, A. J. (1980). Maximally reliable composites for unidimensional measures. *Educational and Psychological Measurement*, *40*(2), 367–375. https://doi.org/10.1177/001316448004000213

Cordeiro, G. M., & Cribari-Neto, F. (2014). *An introduction to Bartlett correction and bias reduction*. Heidelberg, Germany: Springer.

Cronbach, L. J., & Shavelson, R. J. (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, *64*(3), 391–418. https://doi.org/10.1177/0013164404266386

Dimitrov, D. M., Raykov, T., & AL-Qataee, A. A. (2015). Developing a Measure of General Academic Ability: An Application of Maximal Reliability and Optimal Linear Combination to High School Students' Scores. *Educational and Psychological Measurement*, *75*(3), 475–490. https://doi.org/10.1177/0013164414543179

Gabler, S., & Raykov, T. (2017). Evaluation of Maximal Reliability for Unidimensional Measuring Instruments With Correlated Errors. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(1), 104–111.

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

Grice, J. W. (2001). A comparison of factor scores under conditions of factor obliquity. *Psychological Methods*, *6*(1), 67–83. https://doi.org/10.1037//1082-989X.6.1.67

Hancock, G. R. (2001). Effect size, power, and sample size determination for structured means modeling and mimic approaches to between-groups hypothesis testing of means on a single latent construct. *Psychometrika*, *66*(3), 373–388. https://doi.org/10.1007/BF02294440

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural Equation Modeling: Present and Future — A Festschrift in honor of Karl Jöreskog* (pp. 195–216). Lincolnwood, IL: Scientific Software International, Inc.

Hunt, T. D., & Bentler, P. M. (2013). Quantile lower bounds to reliability based on locally optimal splits. *Psychometrika*, *80*(1), 182–195. https://doi.org/10.1007/s11336-013-9393-6

Johnson, J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality*, *51*, 78–89. https://doi.org/10.1016/j.jrp.2014.05.003

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, *36*(2), 109–133. https://doi.org/10.1007/BF02291393

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

Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances is a heywood case a symptom of misspecification? *Sociological Methods & Research*, *41*(1), 124–167. https://doi.org/10.1177/0049124112442138

Kosmidis, I. (2014). Bias in parametric estimation: reduction and useful side-effects. *Wiley Interdisciplinary Reviews: Computational Statistics*, *6*(3), 185–196. https://doi.org/10.1002/wics.1296

Leach, L. F., & Henson, R. K. (2007). The use and impact of adjusted R^2 effects in published regression research. *Multiple Linear Regression Viewpoints*, *33*(1), 1–11.

Leach, L. F., & Henson, R. K. (2014). Bias and Precision of the Squared Canonical Correlation Coefficient Under Nonnormal Data Condition. *Journal of Modern Applied Statistical Methods*, *13*(1), 110–139.

Li, H. (1997). A unifying expression for the maximal reliability of a linear composite. *Psychometrika*, *62*(2), 245–249. https://doi.org/10.1007/BF02295278

Li, H., Rosenthal, R., & Rubin, D. B. (1996). Reliability of measurement in psychology: From Spearman-Brown to maximal reliability. *Psychological Methods*, *1*(1), 98. https://doi.org/10.1037/1082-989X.1.1.98

Li, J., & Bentler, P. M. (2011). The greatest lower bound to reliability: corrected and resampling estimators. *Modeling and Data Analysis*, *1*, 87–104. https://doi.org/10.17759/mda

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor

Analysis. *Annual Review of Clinical Psychology*, *10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

McNeish, D. (2017). Thanks Coefficient Alpha, We'll Take It From Here. *Psychological Methods*. https://doi.org/10.1037/met0000144

Nimon, K., Zientek, L. R., & Henson, R. K. (2012). The assumption of a reliable instrument and other pitfalls to avoid when considering the reliability of data. *Quantitative Psychology and Measurement*, *3*(102). https://doi.org/10.3389/fpsyg.2012.00102

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, *5*(3), 343–355. https://doi.org/10.1037/1082-989X.5.3.343

Penev, S., & Raykov, T. (2006). On the relationship between maximal reliability and maximal validity of linear composites. *Multivariate Behavioral Research*, *41*(2), 105–126. https://doi.org/10.1207/s15327906mbr4102_1

Penev, S., & Raykov, T. (2010). A method of bias correction for maximal reliability with dichotomous measures. *British Journal of Mathematical and Statistical Psychology*, *63*(1), 163–175. https://doi.org/10.1348/000711009X429494

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *The Journal of Applied Psychology*, *98*(1), 194–198. https://doi.org/10.1037/a0030767

R Core Team. (2016). R: A Language and Environment for Statistical Computing (Version 3.3.1). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Raju, N. S., Bilgic, R., Edwards, J. E., & Fleer, P. F. (1999). Accuracy of population validity and cross-validity estimation: An empirical comparison of formula-based, traditional empirical, and equal weights procedures. *Applied Psychological Measurement*, *23*(2), 99–115. https://doi.org/10.1177/01466219922031220

Raykov, T. (1997). Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement*, *21*(2), 173–184. https://doi.org/10.1177/01466216970212006

Raykov, T. (2004). Estimation of maximal reliability: A note on a covariance structure modelling approach. *British Journal of Mathematical and Statistical Psychology*, *57*(1), 21–27. https://doi.org/10.1348/000711004849295

Raykov, T. (2005). Studying group and time invariance in maximal reliability for multiple-component measuring instruments via covariance structure modelling. *British Journal of Mathematical and Statistical Psychology*, *58*(2), 301–317. https://doi.org/10.1348/000711005X38591

Raykov, T., Gabler, S., & Dimitrov, D. M. (2015). Maximal Reliability and Composite Reliability: Examining Their Difference for Multicomponent Measuring Instruments Using Latent Variable Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(3), 384–391. https://doi.org/10.1080/10705511.2014.966369

Raykov, T., & Hancock, G. R. (2005). Examining change in maximal reliability for multiple-component measuring instruments. *British Journal of Mathematical & Statistical Psychology*, *58*(1), 65–82. https://doi.org/10.1348/000711005X38753

Raykov, T., & Marcoulides, G. A. (2016). Scale Reliability Evaluation Under Multiple Assumption Violations. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(2), 302–313. https://doi.org/10.1080/10705511.2014.938597

Raykov, T., & Penev, S. (2009). Estimation of maximal reliability for multiple-component instruments in multilevel designs. *British Journal of Mathematical and Statistical Psychology*, *62*(1), 129–142. https://doi.org/10.1348/000711007X255345

Reddy, S. K. (1992). Effects of Ignoring Correlated Measurement Error in Structural Equation Models. *Educational and Psychological Measurement*, *52*(3), 549–570. https://doi.org/10.1177/0013164492052003005

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Savalei, V., & Kolenikov, S. (2008). Constrained versus unconstrained estimation in structural equation modeling. *Psychological Methods*, *13*(2), 150–170. https://doi.org/10.1037/1082-989X.13.2.150

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (Third edition). Thousand Oaks, California: SAGE.

Shapiro, A., & Ten Berge, J. M. (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

Shieh, G. (2008). Improved shrinkage estimation of squared multiple correlation coefficient and squared cross-validity coefficient. *Organizational Research Methods*, *11*(2), 387–407. https://doi.org/10.1177/1094428106292901

Shieh, G. (2010). Estimation of the simple correlation coefficient. *Behavior Research Methods*, *42*(4), 906–917. https://doi.org/10.3758/BRM.42.4.906

Sijtsma, K. (2009). Reliability Beyond Theory and Into Practice. *Psychometrika*, *74*(1), 169. https://doi.org/10.1007/s11336-008-9103-y

Skidmore, S. T., & Thompson, B. (2011). Choosing the Best Correction Formula for the Pearson r^2 Effect Size. *The Journal of Experimental Education*, *79*(3), 257–278. https://doi.org/10.1080/00220973.2010.484437

Ten Berge, J. M., & 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

Thompson, B. (1990). Finding a correction for the sampling error in multivariate measures of relationship: A Monte Carlo study. *Educational and Psychological Measurement*, *50*(1), 15–31. https://doi.org/10.1177/0013164490501003

Thompson, B. (1999). If statistical significance tests are broken/misused, what practices should supplement or replace them? *Theory & Psychology*, *9*(2), 165–181. https://doi.org/10.1177/095935439992006

Thurstone, L. (1935). *The Vectors of Mind*. Chicago, IL: University of Chicago Press.

Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, *51*(4), 473–481. https://doi.org/10.1037/0022-0167.51.4.473

Walker, D. A. (2007). A comparison of eight shrinkage formulas under extreme conditions. *Journal of Modern Applied Statistical Methods*, *6*(1), 16.

Wang, Z., & Thompson, B. (2007). Is the Pearson r^2 biased, and if so, what is the best correction formula? *The Journal of Experimental Education*, *75*(2), 109–125. https://doi.org/10.3200/JEXE.75.2.109-125

Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, *29*(4), 377–392. https://doi.org/10.1177/0734282911406668

Yin, P., & Fan, X. (2001). Estimating R^2 Shrinkage in Multiple Regression: A Comparison of Different Analytical Methods. *The Journal of Experimental Education*, *69*(2), 203–224. https://doi.org/10.1080/00220970109600656

Zumbo, B. D., Williams, R. H., & Zimmerman, D. W. (2003). Bias in estimation and hypothesis testing of correlation. *Psicológica: Revista de Metodología y Psicología Experimental*, *24*(1), 133–158.

Footnotes

[1] This reliability index is more commonly known as “Cronbach’s alpha” in the literature. However, the index was not originally invented by Cronbach and he himself regrets that it carries his name (Cronbach & Shavelson, 2004).

[2] Consider that the formula for Hancock’s coefficient H is $H = \frac{\sum_{i=1}^{k}\frac{\ell_i^2}{1-\ell_i^2}}{1+\sum_{i=1}^{k}\frac{\ell_i^2}{1-\ell_i^2}}$, where ℓ_i^2 (the squared standardized factor loading of indicator i) is the individual reliability of the indicator i, assuming a correctly specified single-factor model, as is the case here, and equivalent to ρ_i in Eq. 3.

[3] The complete list of all studies reviewed is available in Appendix F.

[4] In the few cases where a publication reported multiple studies, we noted the smallest of the samples, as that is an indicator of the lower threshold at which researchers still feel comfortable applying coefficient H.

[5] The weights calculated this way are equivalent to those calculated from the ratio of the loading to the error variance for each individual indicator, up to a scaling constant that is not relevant for obtaining maximal reliability estimates. See Appendix A for an empirical demonstration using the two examples from Raykov et al. (2015).

[6] This range from 3 to 9 indicators covers the majority of scales used in empirical psychological research (McNeish, 2017). For example, Peterson and Kim (2013, p. 196, Table 2) report the results of a large-scale review of psychological, educational, and management research employing composite reliability coefficients and note that, out of 2,448 scales where the number of items was reported, only 24 (or 1.7%) had more than 9 items.

[7] This can be presented formally as a function of reliability: $\underset{\hat{\rho}_j \to 1}{plim}(\hat{\rho}_{max}) = 1$ because both the numerator and denominator of Equation 3 grow infinitely large: $\underset{\hat{\rho}_j \to 1}{plim}\left(\sum_{i=1}^{k}\frac{\hat{\rho}_i}{1-\hat{\rho}_i}\right), \underset{\hat{\rho}_j \to 1}{plim}\left(1+\sum_{i=1}^{k}\frac{\hat{\rho}_i}{1-\hat{\rho}_i}\right) = \infty$, for all $j \in \lceil 1, k \rceil$

[8] As a robustness check, we re-estimated the inadmissible replications with a model that included inequality constraints forcing all variance estimates to be greater than zero (Savalei & Kolenikov, 2008).These results showed even greater bias for $\hat{\rho}_{max}$. The results based on all runs (both unconstrained and constrained) are shown in Appendix E.

[9] Retrieved from http://osf.io/tbmh5 on January 12, 2017.

Table 1

Notation and Description for Reliability Coefficients

Reliability Coefficients	Description
ρ_{max}	Population maximal reliability Proportion of weighted population true score variance over population observed score variance. Based on composites formed using weights equal to the ratios of population factor loadings and measurement error variances. Estimated as $\hat{\rho}_{max}$ in finite samples.
r_{max}	True sample maximal reliability Squared sample correlation between the true factor scores retained from the data generation process and composites created using maximal reliability weights (equal to the ratio of estimated loadings and error variances).
ρ_{CR}	Population composite reliability Proportion of unit-weighted population true score variance over population observed score variance. Based on composites formed using unit weights (i.e., weights of 1). Estimated as $\hat{\rho}_{CR}$ in finite samples.
r_{CR}	True sample composite reliability Squared sample correlation between true factor scores and composites created using unit weights.

Table 2

Loadings for the Unequal Loadings Simulation Condition

Base Loading	Number of Indicators	Loadings Vector
0.6	3	[0.500, 0.600, 0.700]
	5	[0.474, 0.537, 0.600, 0.663, 0.726]
	7	[0.461, 0.507, 0.554, 0.600, 0.646, 0.693, 0.739]
	9	[0.454, 0.490, 0.527, 0.563, 0.600, 0.637, 0.673, 0.710, 0.746]
0.7	3	[0.600, 0.700, 0.800]
	5	[0.574, 0.637, 0.700, 0.763, 0.826]
	7	[0.561, 0.607, 0.654, 0.700, 0.746, 0.793, 0.839]
	9	[0.554, 0.590, 0.627, 0.663, 0.700, 0.737, 0.773, 0.810, 0.846]
0.8	3	[0.700, 0.800, 0.900]
	5	[0.674, 0.737, 0.800, 0.863, 0.926]
	7	[0.661, 0.707, 0.754, 0.800, 0.846, 0.893, 0.939]
	9	[0.654, 0.690, 0.727, 0.763, 0.800, 0.837, 0.873, 0.910, 0.946]

Note: loadings only presented to three decimals here for ease of exposition.

Table 3

Results for Unequal Loadings (Selected Conditions)

Sample Size	Indicators	Loadings	Difference between population and sample statistics							
			Maximal reliability				Composite reliability			
			Population $(\hat{\rho}_{max} - \rho_{max})$		Sample $(r_{max} - \rho_{max})$		Population $(\hat{\rho}_{CR} - \rho_{CR})$		Sample $(r_{CR} - \rho_{CR})$	
			Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
25	3	0.6	15.2 %	24.8%	-9.3 %	23.8%	4.6 %	19.0%	1.3 %	18.6%
50	3	0.6	10.7 %	19.7%	-7.6 %	17.7%	2.5 %	13.8%	0.2 %	13.2%
75	3	0.6	8.1 %	16.9%	-6.3 %	14.3%	1.3 %	11.4%	-0.2 %	10.8%
25	5	0.6	8.4 %	12.6%	-8.4 %	17.9%	0.3 %	10.9%	-0.2 %	12.4%
50	5	0.6	4.2 %	8.7%	-4.6 %	11.2%	-0.4 %	8.1%	-0.4 %	8.8%
75	5	0.6	2.6 %	6.7%	-2.9 %	8.2%	-0.2 %	6.6%	-0.2 %	7.2%
25	3	0.7	6.8 %	13.9%	-6.8 %	16.3%	1.0 %	11.8%	-0.1 %	12.4%
50	3	0.7	4.8 %	10.9%	-4.7 %	11.2%	0.5 %	8.5%	-0.2 %	8.6%
75	3	0.7	3.7 %	9.3%	-3.6 %	8.7%	0.2 %	6.8%	-0.3 %	7.0%
25	5	0.7	3.4 %	6.8%	-4.7 %	10.8%	-0.5 %	7.1%	-0.6 %	8.1%
50	5	0.7	1.7 %	4.7%	-2.2 %	6.2%	-0.4 %	4.9%	-0.3 %	5.5%
75	5	0.7	1.1 %	3.6%	-1.4 %	4.6%	-0.2 %	4.0%	-0.2 %	4.5%
25	3	0.8	2.0 %	6.9%	-3.9 %	9.1%	-0.1 %	6.7%	-0.6 %	7.3%
50	3	0.8	1.6 %	5.4%	-2.4 %	5.7%	0.0 %	4.6%	-0.3 %	5.1%
75	3	0.8	1.4 %	4.8%	-1.8 %	4.3%	-0.1 %	3.8%	-0.2 %	4.0%
25	5	0.8	1.1 %	3.2%	-2.0 %	5.0%	-0.4 %	4.0%	-0.5 %	4.8%
50	5	0.8	0.7 %	2.4%	-1.0 %	2.8%	-0.2 %	2.7%	-0.2 %	3.1%
75	5	0.8	0.4 %	1.8%	-0.6 %	2.2%	-0.1 %	2.1%	-0.1 %	2.6%

Note: The Mean values shown are calculated as (average estimate for a condition – population value) / population value. Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table 4

Results for Equal Loadings (Selected Conditions)

Sample Size	Indicators	Loadings	Difference between population and sample statistics							
			Maximal reliability				Composite reliability			
			Population ($\hat{\rho}_{max} - \rho_{max}$)		Sample ($r_{max} - \rho_{max}$)		Population ($\hat{\rho}_{CR} - \rho_{CR}$)		Sample ($r_{CR} - \rho_{CR}$)	
			Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
25	3	0.6	18.2 %	27.2%	-8.9 %	24.4%	4.9 %	19.1%	1.5 %	18.6%
50	3	0.6	12.3 %	20.7%	-7.8 %	18.5%	2.4 %	14.0%	0.3 %	13.2%
75	3	0.6	8.6 %	16.6%	-6.5 %	15.3%	1.1 %	11.8%	-0.2 %	10.8%
25	5	0.6	10.2 %	14.0%	-9.4 %	19.5%	0.3 %	11.0%	-0.1 %	12.5%
50	5	0.6	4.4 %	8.8%	-5.0 %	12.2%	-0.4 %	8.3%	-0.4 %	8.9%
75	5	0.6	2.6 %	6.9%	-3.1 %	8.9%	-0.2 %	6.8%	-0.2 %	7.2%
25	3	0.7	8.9 %	15.5%	-6.9 %	17.1%	0.9 %	12.1%	0.0 %	12.4%
50	3	0.7	5.6 %	11.2%	-4.8 %	12.0%	0.4 %	8.6%	-0.2 %	8.8%
75	3	0.7	3.5 %	8.5%	-3.5 %	9.3%	0.0 %	7.1%	-0.3 %	7.0%
25	5	0.7	4.2 %	7.6%	-5.5 %	12.4%	-0.6 %	7.2%	-0.6 %	8.2%
50	5	0.7	1.7 %	5.0%	-2.5 %	7.1%	-0.4 %	5.0%	-0.3 %	5.7%
75	5	0.7	1.1 %	4.0%	-1.5 %	5.2%	-0.2 %	4.0%	-0.2 %	4.5%
25	3	0.8	4.0 %	8.2%	-4.2 %	10.3%	-0.2 %	7.0%	-0.4 %	7.5%
50	3	0.8	2.3 %	5.6%	-2.5 %	6.8%	-0.1 %	4.9%	-0.2 %	5.1%
75	3	0.8	1.4 %	4.3%	-1.7 %	5.0%	-0.1 %	3.9%	-0.2 %	4.0%
25	5	0.8	1.6 %	4.0%	-2.7 %	6.8%	-0.5 %	4.1%	-0.5 %	4.9%
50	5	0.8	0.7 %	2.7%	-1.1 %	3.8%	-0.2 %	2.7%	-0.2 %	3.2%
75	5	0.8	0.4 %	2.1%	-0.7 %	2.9%	-0.1 %	2.1%	-0.1 %	2.6%

Note: The Mean values shown are calculated as (average estimate for a condition – population value) / population value. Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table 5

Results for Unequal Loadings Based on Unbiased Estimates (Selected Conditions)

Sample Size	Indicators	Loadings	Difference between population and sample statistics							
			Maximal reliability				Composite reliability			
			Population $(\hat{\rho}_{max} - \rho_{max})$		Sample $(r_{max} - \rho_{max})$		Population $(\hat{\rho}_{CR} - \rho_{CR})$		Sample $(r_{CR} - \rho_{CR})$	
			Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
25	3	0.6	9.3 %	24.6%	-7.4 %	25.5%	-1.7 %	18.1%	1.3 %	18.8%
50	3	0.6	7.6 %	19.9%	-6.8 %	18.5%	-1.0 %	13.4%	0.2 %	13.3%
75	3	0.6	6.1 %	16.7%	-5.9 %	14.7%	-0.8 %	11.3%	-0.2 %	10.8%
25	5	0.6	6.1 %	12.2%	-7.3 %	18.8%	-1.6 %	11.1%	-0.2 %	12.7%
50	5	0.6	3.3 %	8.2%	-4.3 %	11.1%	-0.9 %	8.1%	-0.4 %	8.9%
75	5	0.6	2.0 %	6.5%	-2.8 %	8.1%	-0.6 %	6.6%	-0.2 %	7.2%
25	3	0.7	4.7 %	13.7%	-5.5 %	16.5%	-1.6 %	11.8%	-0.1 %	12.5%
50	3	0.7	3.7 %	10.8%	-4.2 %	11.0%	-0.9 %	8.4%	-0.2 %	8.6%
75	3	0.7	2.9 %	9.0%	-3.4 %	8.6%	-0.6 %	6.8%	-0.3 %	6.9%
25	5	0.7	2.5 %	6.5%	-4.1 %	10.6%	-1.1 %	7.1%	-0.6 %	8.2%
50	5	0.7	1.3 %	4.5%	-2.1 %	6.1%	-0.6 %	4.9%	-0.3 %	5.6%
75	5	0.7	0.8 %	3.6%	-1.3 %	4.5%	-0.4 %	3.9%	-0.2 %	4.5%
25	3	0.8	2.2 %	7.0%	-3.1 %	8.7%	-1.0 %	6.7%	-0.6 %	7.4%
50	3	0.8	1.6 %	5.5%	-2.1 %	5.5%	-0.5 %	4.7%	-0.3 %	5.0%
75	3	0.8	1.3 %	4.7%	-1.7 %	4.2%	-0.3 %	3.7%	-0.2 %	4.0%
25	5	0.8	0.9 %	3.1%	-1.8 %	4.7%	-0.6 %	3.9%	-0.5 %	4.7%
50	5	0.8	0.5 %	2.3%	-0.9 %	2.8%	-0.3 %	2.6%	-0.2 %	3.1%
75	5	0.8	0.3 %	1.8%	-0.6 %	2.1%	-0.2 %	2.1%	-0.1 %	2.5%

Note: The Mean values shown are calculated as (average estimate for a condition – population value) / population value. Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table 6

Results for Equal Loadings Based on Unbiased Estimates (Selected Conditions)

Sample Size	Indicators	Loadings	Difference between population and sample statistics							
			Maximal reliability				Composite reliability			
			Population ($\hat{\rho}_{max} - \rho_{max}$)		Sample ($r_{max} - \rho_{max}$)		Population ($\hat{\rho}_{CR} - \rho_{CR}$)		Sample ($r_{CR} - \rho_{CR}$)	
			Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
25	3	0.6	10.2 %	26.7%	-7.5 %	26.6%	-1.7 %	18.3%	1.5 %	18.9%
50	3	0.6	8.2 %	21.3%	-7.1 %	19.9%	-1.0 %	13.7%	0.3 %	13.4%
75	3	0.6	6.4 %	16.9%	-6.1 %	15.9%	-0.8 %	11.6%	-0.2 %	10.9%
25	5	0.6	7.1 %	13.3%	-8.2 %	20.7%	-1.6 %	11.1%	-0.1 %	12.8%
50	5	0.6	3.6 %	8.4%	-4.7 %	12.2%	-0.9 %	8.2%	-0.4 %	8.9%
75	5	0.6	2.1 %	6.6%	-3.0 %	8.9%	-0.6 %	6.8%	-0.2 %	7.2%
25	3	0.7	5.3 %	15.1%	-5.8 %	18.2%	-1.6 %	12.0%	0.0 %	12.6%
50	3	0.7	3.9 %	11.0%	-4.4 %	12.3%	-0.9 %	8.7%	-0.2 %	8.7%
75	3	0.7	2.7 %	8.3%	-3.3 %	9.3%	-0.6 %	7.1%	-0.3 %	7.0%
25	5	0.7	3.0 %	7.0%	-4.8 %	12.3%	-1.1 %	7.2%	-0.6 %	8.3%
50	5	0.7	1.4 %	4.7%	-2.3 %	6.9%	-0.6 %	5.0%	-0.3 %	5.6%
75	5	0.7	0.8 %	3.8%	-1.4 %	5.2%	-0.4 %	4.0%	-0.2 %	4.5%
25	3	0.8	2.5 %	7.8%	-3.6 %	10.7%	-1.0 %	7.0%	-0.4 %	7.6%
50	3	0.8	1.6 %	5.3%	-2.3 %	6.7%	-0.5 %	4.8%	-0.2 %	5.1%
75	3	0.8	1.1 %	4.1%	-1.6 %	5.0%	-0.4 %	3.9%	-0.2 %	4.1%
25	5	0.8	1.1 %	3.7%	-2.4 %	6.4%	-0.7 %	4.0%	-0.5 %	4.9%
50	5	0.8	0.5 %	2.6%	-1.1 %	3.7%	-0.3 %	2.7%	-0.2 %	3.2%
75	5	0.8	0.3 %	2.1%	-0.7 %	2.8%	-0.2 %	2.1%	-0.1 %	2.6%

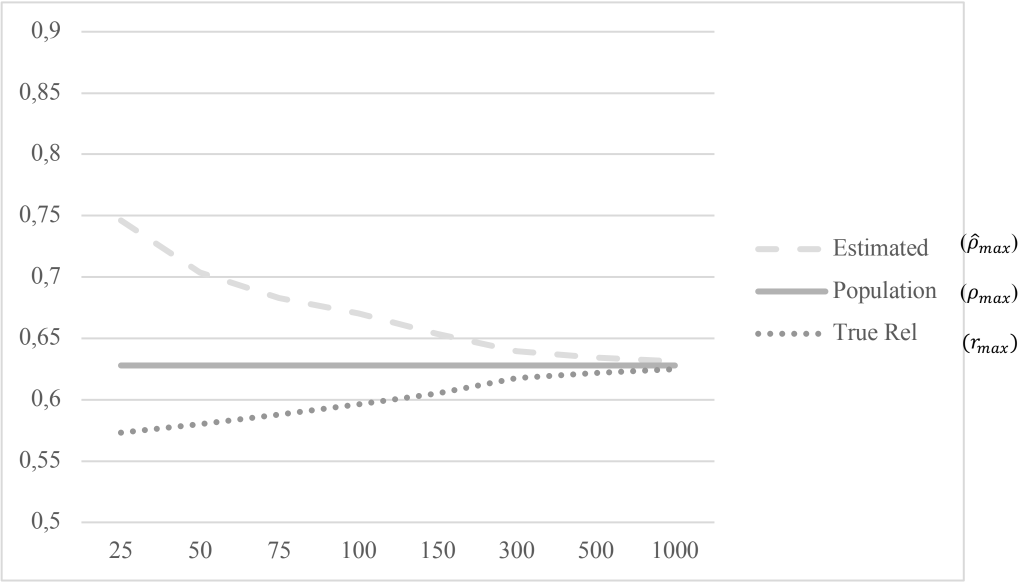
Note: The Mean values shown are calculated as (average estimate for a condition – population value) / population value. Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table 7

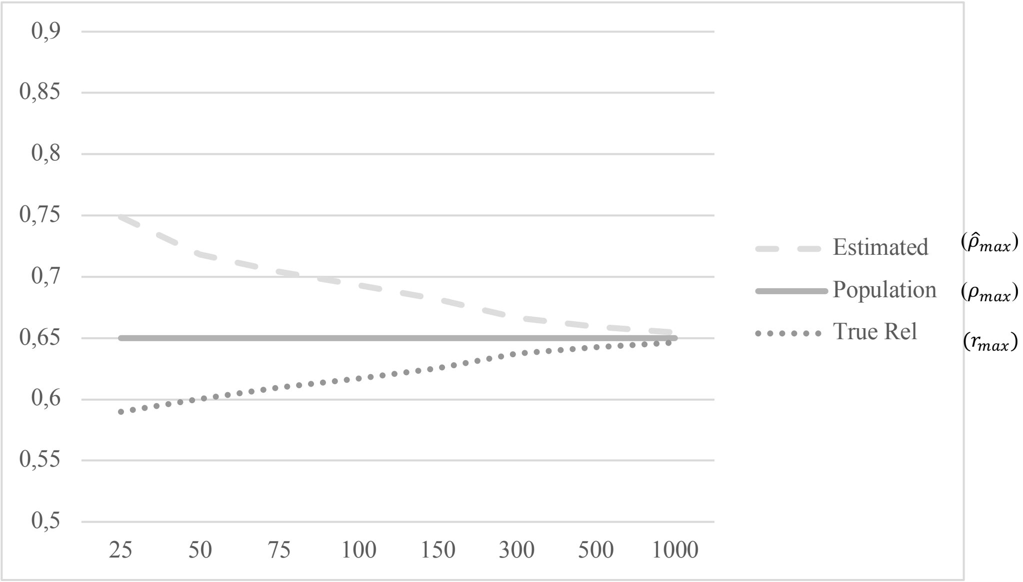
Empirical Illustration Results

Estimate	Sample					
	N = 25	N = 50	N = 100	N = 500	N = 1000	Full data
$\hat{\lambda}_{I12}$	0.795	0.846	0.720	0.855	0.855	0.867
$\hat{\lambda}_{I42}$	0.715	0.731	0.790	0.738	0.740	0.821
$\hat{\lambda}_{I72}$	0.948	0.838	0.844	0.848	0.782	0.778
$\hat{\lambda}_{I102}$	0.585	0.526	0.584	0.695	0.640	0.747
$\widehat{Var(\varepsilon_{I12})},$	0.368	0.285	0.481	0.269	0.269	0.289
$\widehat{Var(\varepsilon_{I42})},$	0.489	0.465	0.376	0.455	0.453	0.525
$\widehat{Var(\varepsilon_{I72})},$	0.102	0.298	0.287	0.282	0.389	0.329
$\widehat{Var(\varepsilon_{I102})},$	0.658	0.723	0.659	0.517	0.591	0.665
$\hat{\rho}_{max}$	0.924	0.865	0.852	0.881	0.861	0.868
$\hat{\rho}_{CR}$	0.851	0.830	0.827	0.866	0.842	0.855

Note: The full data had an N = 605,607 (only cases with no missing data were included). All loading and error variance estimates are standardized.

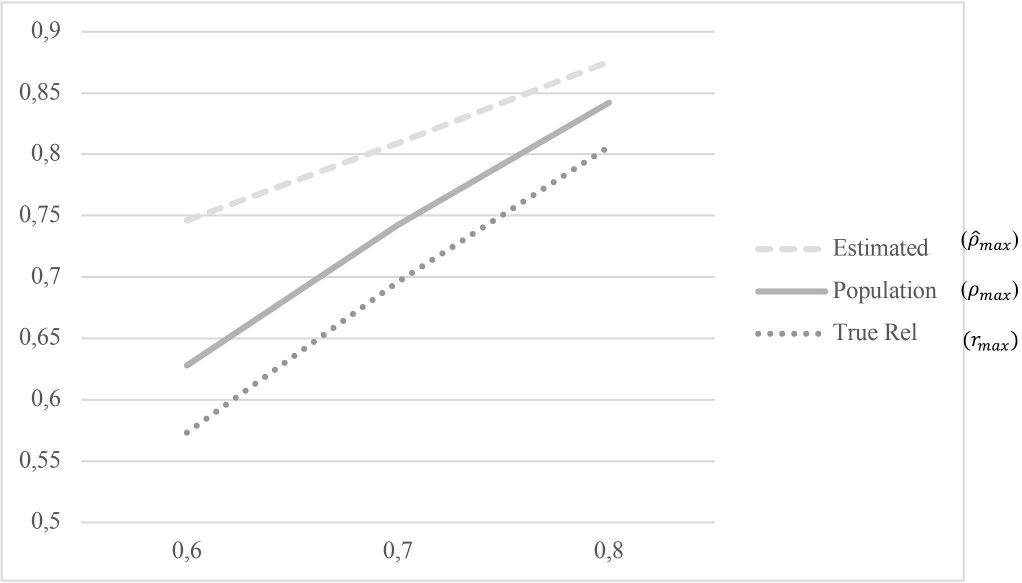


Panel (a). Equal Loading Condition

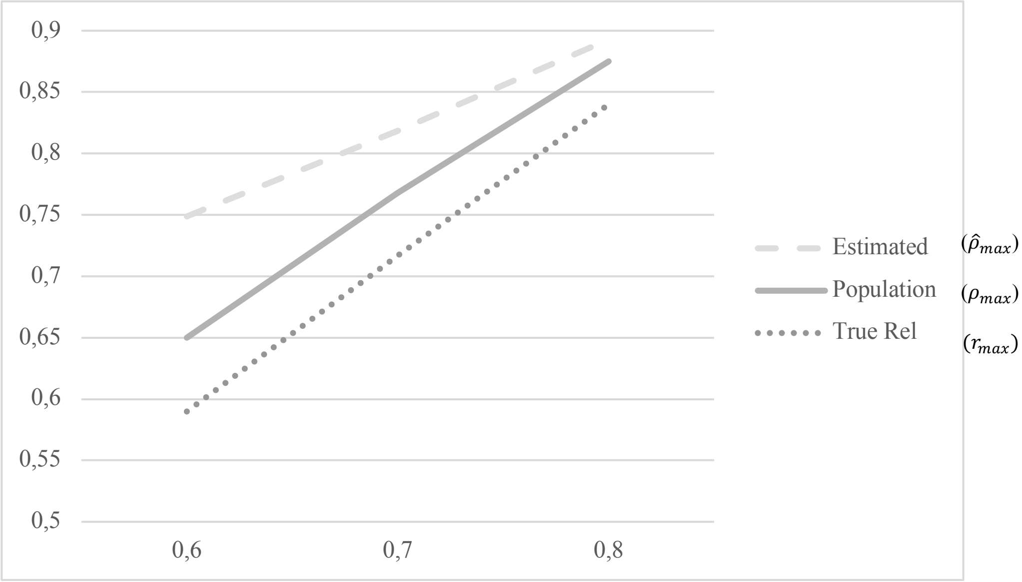


Panel (b). Unequal Loading Condition

Figure 1. Maximal Reliability Estimates over Sample Size

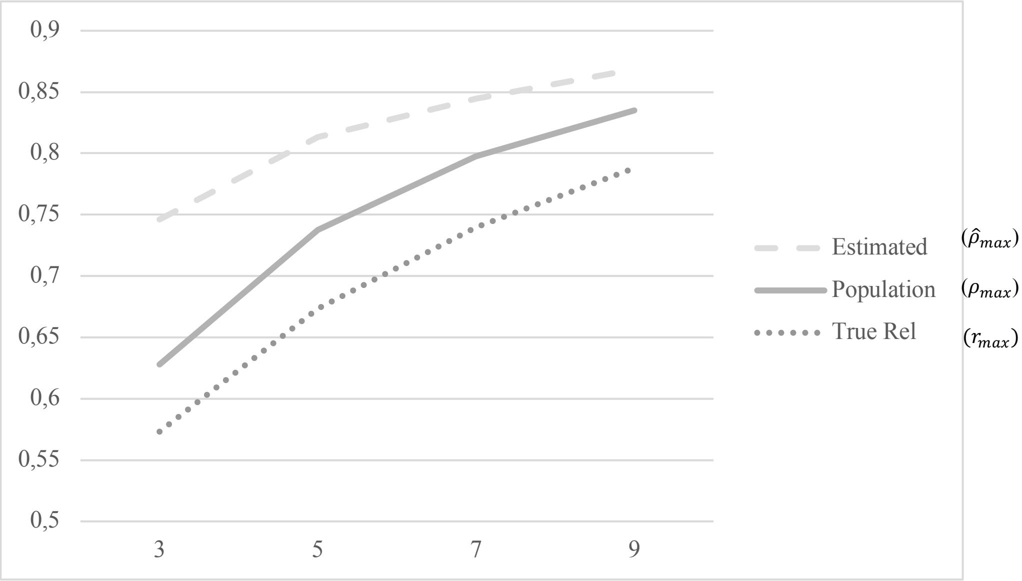


Panel (a). Equal Loading Condition

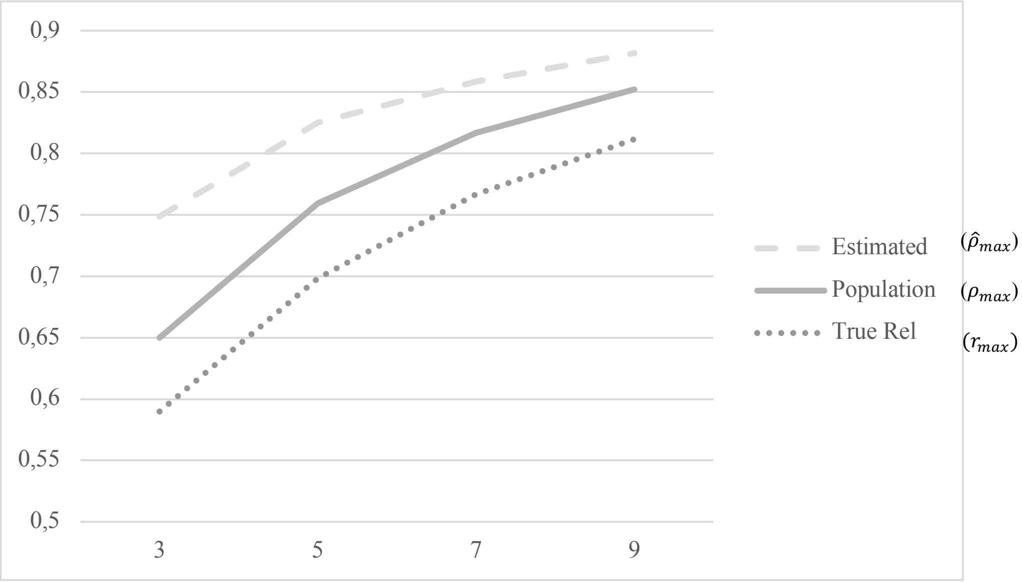


Panel (b). Unequal Loading Condition

Figure 2. Maximal Reliability Estimates over Loading Strength (N = 25, 3 indicators)

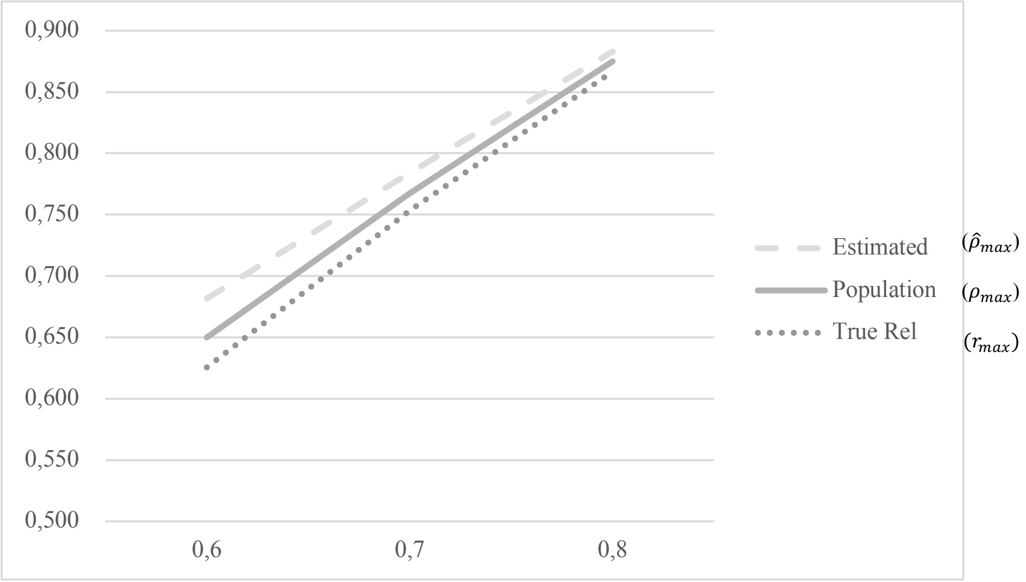


Panel (a). Equal Loading Condition

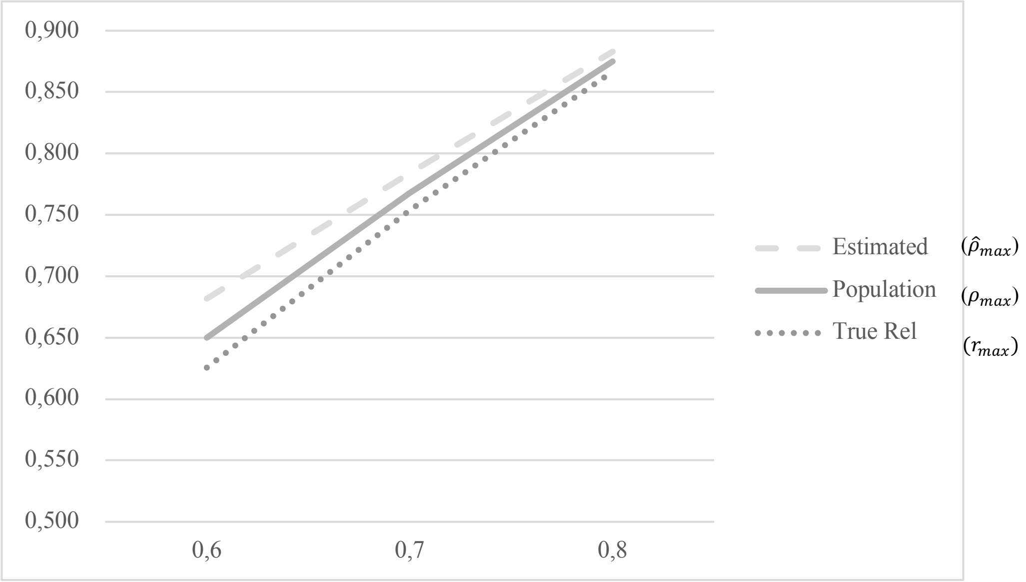


Panel (b). Unequal Loading Condition

Figure 3. Maximal Reliability Estimates over Number of Indicators (N = 25, 0.6 loadings)

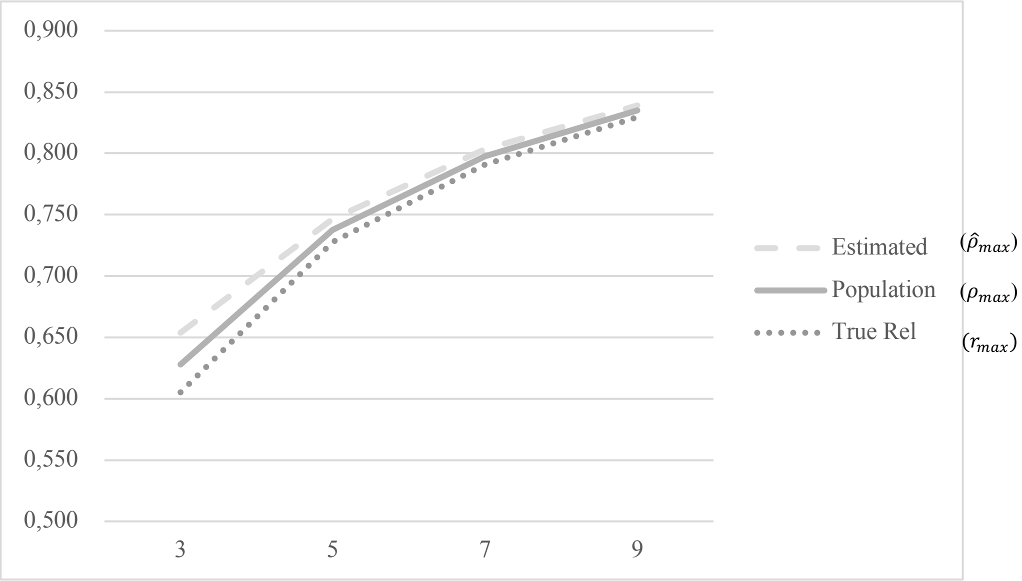


Panel (a). Equal Loading Condition

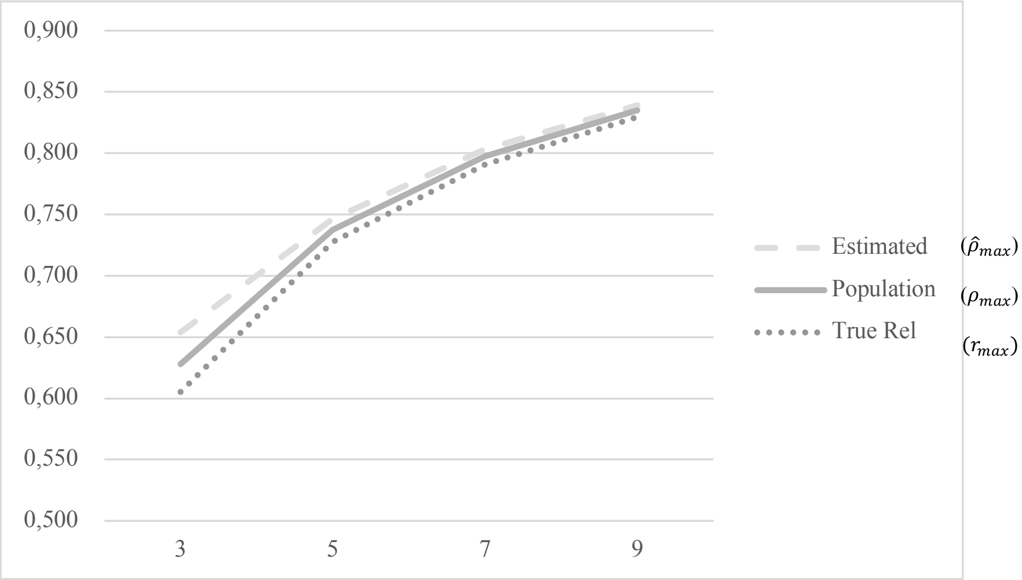


Panel (b). Unequal Loading Condition

Figure 4. Maximal Reliability Estimates over Loading Strength (N = 150, 3 indicators)

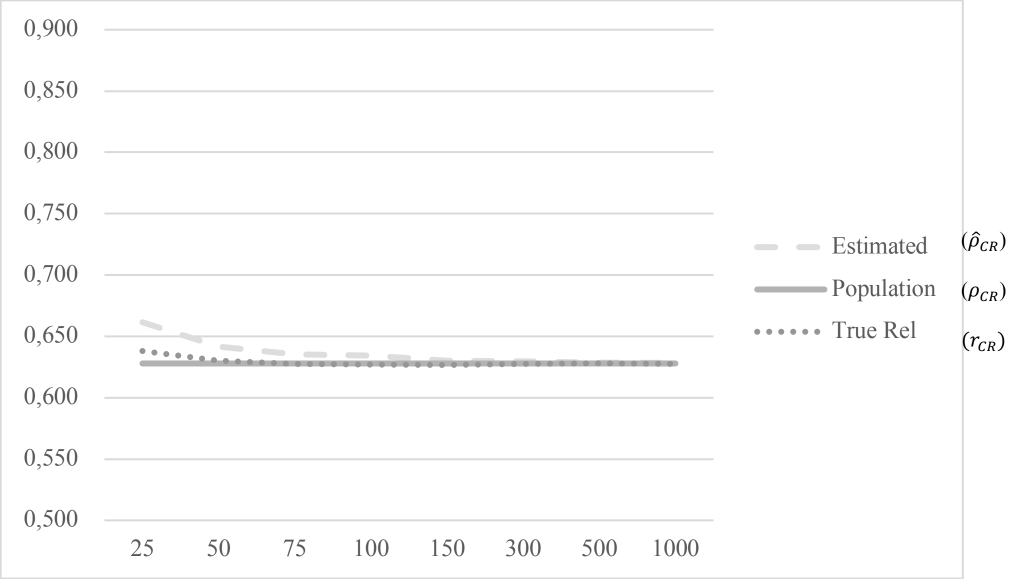


Panel (a). Equal Loading Condition

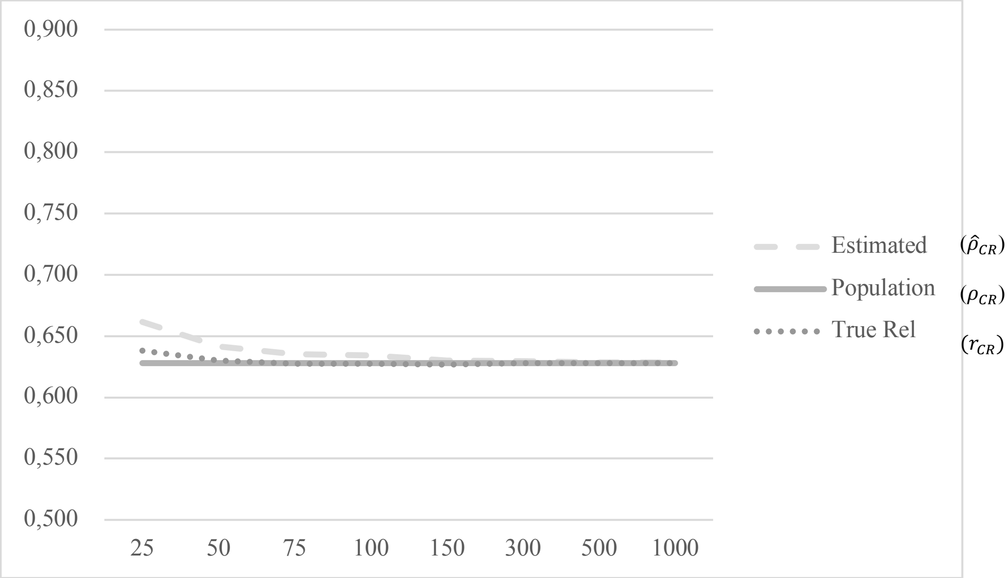


Panel (b). Unequal Loading Condition

Figure 5. Maximal Reliability Estimates over Number of Indicators (N = 150, 0.6 loadings)

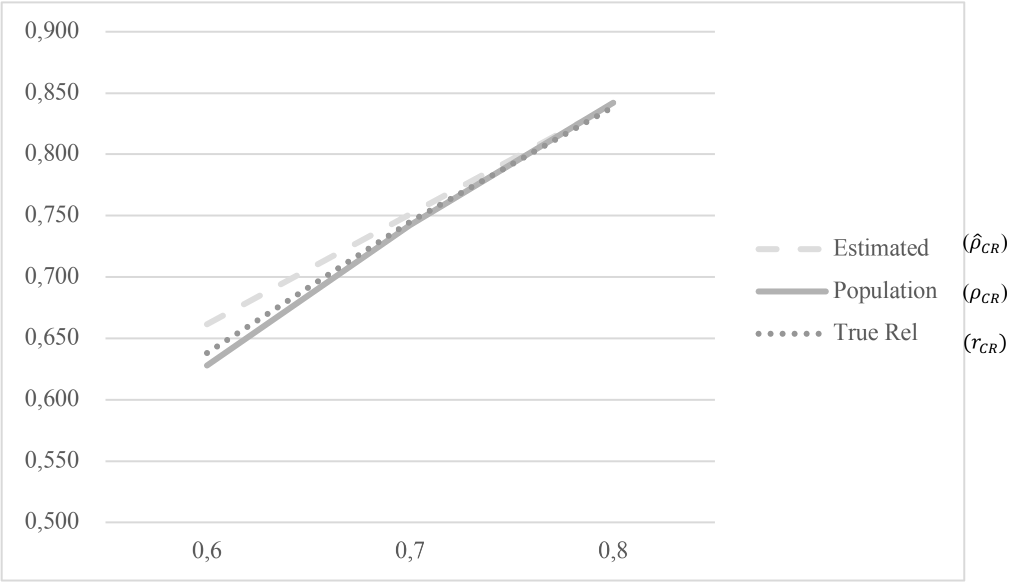


Panel (a). Equal Loading Condition

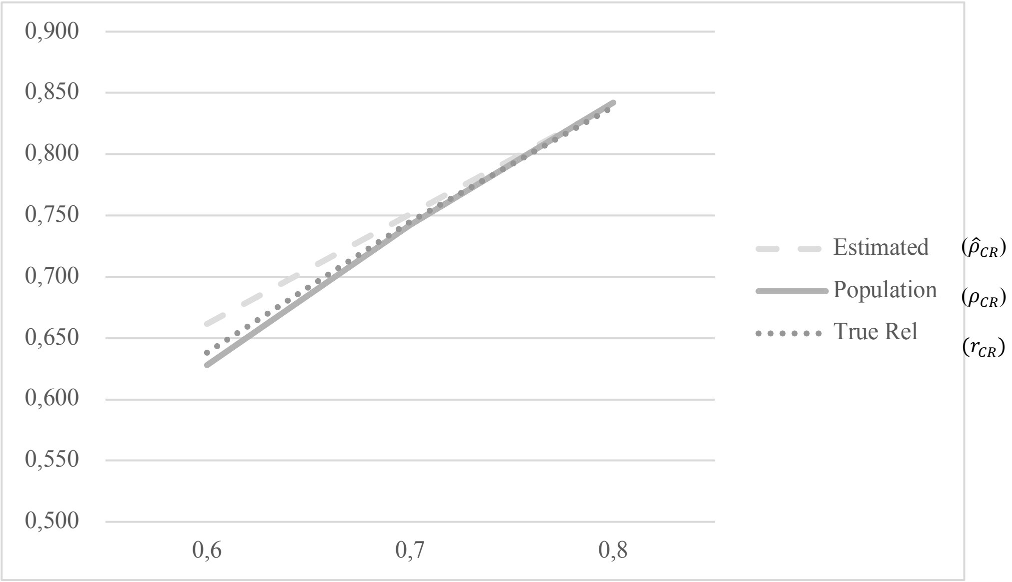


Panel (b). Unequal Loading Condition

Figure 6. Composite Reliability Estimates over Sample Size

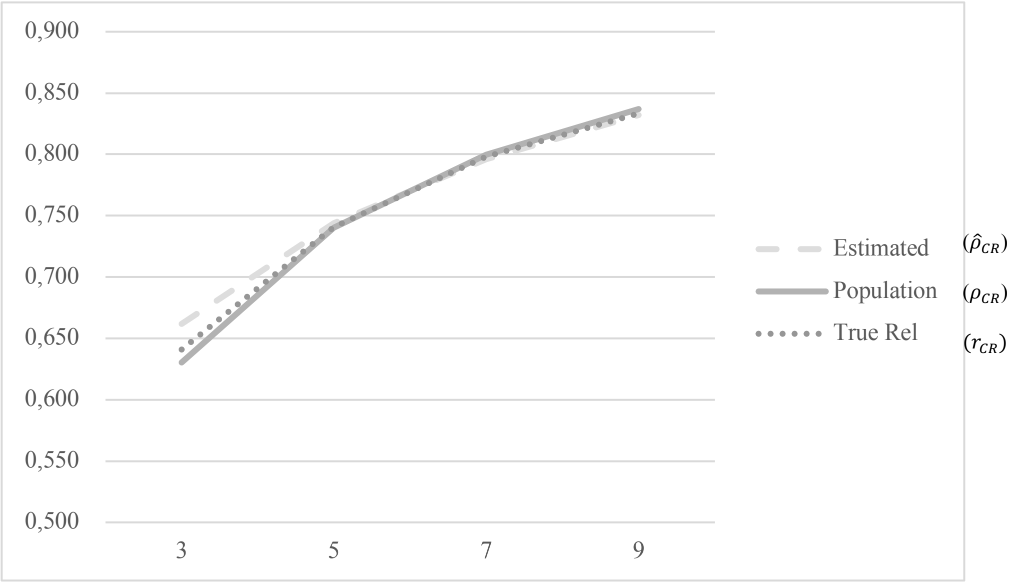


Panel (a). Equal Loading Condition

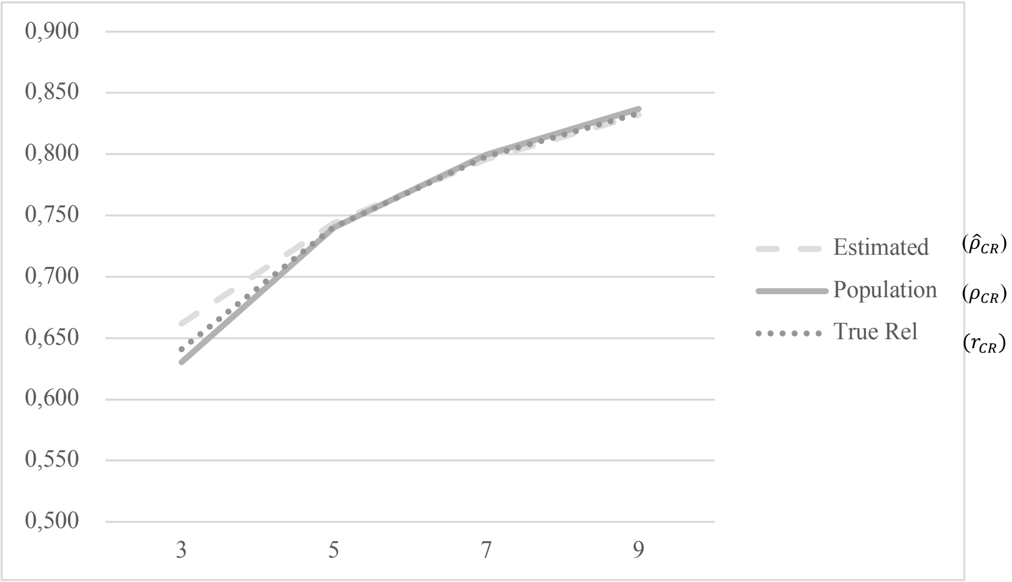


Panel (b). Unequal Loading Condition

Figure 7. Composite Reliability Estimates over Loading Strength (N = 25, 3 indicators)



Panel (a). Equal Loading Condition



Panel (b). Unequal Loading Condition

Figure 8. Composite Reliability Estimates over Number of Indicators (N = 25, 0.6 loadings)

Supplemental Material, "A Note on the Finite Sample Behavior of Maximal Reliability"

Appendix A – Equivalence of Maximal Reliability Weights and Regression Weights

In the first example employed by Raykov et al. (2015), the vectors of loadings, error variances and weights were as follows:

$$\boldsymbol{\lambda} = [1.00, 1.25, 1.50, 1.75, 2.00]$$

$$\boldsymbol{\theta} = [2.00, 2.50, 3.00, 3.50, 4.00]$$

$$\mathbf{w}_{max} = [0.50, 0.50, 0.50, 0.50, 0.50]$$

Using the formula for maximal reliability presented in the main body of this research (Eq. 2), the resulting value of this index is 0.7894737 (≅ 0.789 in the original). Using these same vectors, we can calculate the value of the maximal reliability index using the regression formulation as follows:

$$\boldsymbol{\Sigma}_{yy} = \boldsymbol{\lambda}\boldsymbol{\lambda}^{T} + \boldsymbol{\psi}$$

$$= \begin{bmatrix} 3.0000 & 1.2500 & 1.5000 & 1.7500 & 2.0000 \\ 1.2500 & 4.0625 & 1.8750 & 2.1875 & 2.5000 \\ 1.5000 & 1.8750 & 5.2500 & 2.6250 & 3.0000 \\ 1.7500 & 2.1875 & 2.6250 & 6.5625 & 3.5000 \\ 2.0000 & 2.5000 & 3.0000 & 3.5000 & 8.0000 \end{bmatrix}$$

$$\mathbf{w}_{max} = \boldsymbol{\Sigma}_{\boldsymbol{yy}}^{\mathbf{-1}}\boldsymbol{\lambda}$$

$$= [0.1052632, 0.1052632, 0.1052632, 0.1052632, 0.1052632]$$

$$\rho_{max} = \mathbf{w}_{\boldsymbol{max}}{}^{\boldsymbol{t}}\boldsymbol{\lambda}$$

$$= 0.7894737$$

Thus under both the traditional maximal reliability computations and the regression formulation presented here, exactly the same value for the maximal reliability index is obtained.

It is also easy to see that both sets of weights are the same (in this first example, equal weights are deemed optimal) up to a scaling constant.

Consider now the second example, where:

$$\boldsymbol{\lambda} = [1.00, 1.25, 1.50, 1.75, 2.00]$$

$$\boldsymbol{\theta} = [1.00, 2.00, 3.00, 4.00, 5.00]$$

$$\mathbf{w}_{max} = [1.000, 0.6250, 0.5000, 0.4375, 0.4000]$$

The population value of the maximal reliability index in this case would be 0.803818 (≅ 0.804 in the original). The vector of weights that can be obtained following the regression form is:

$$\begin{aligned}\mathbf{w}_{max} &= \boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-\mathbf{1}}\boldsymbol{\lambda} \\ &= [0.19619865, 0.12262416, 0.09809933, 0.08583691, 0.07847946]\end{aligned}$$

The proportional relationship among the different weights is the same in both cases (e.g., the ratio of the first weight to the second is 1.6 in both cases, and so on), but the approaches differ in how the composite is scaled. In fact, the second vector of weights can be obtained by multiplying the first by 0.1961987. Using the second set of weights to calculate the maximal reliability index using the regression formulas leads to the same value as originally provided in the example.

Supplemental Material, "A Cautionary Note on the Finite Sample Behavior of Maximal Reliability"

Appendix B – Full Simulation Results for Unequal Loadings

Table B1

Full Simulation Results for Unequal Loadings (Bias)

Sample Size	Indicators	Loadings	Diifference between estimate, population value, and sample statistic					
			Maximal reliability			Composite reliability		
			Estimate vs. pop. ($\hat{\rho}_{max} - \rho_{max}$)	Sample reliability vs. pop ($r_{max} - \rho_{max}$)	Est. vs. sample reliability ($\hat{\rho}_{max} - r_{max}$)	Estimate vs. pop. ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample reliability vs. pop ($r_{CR} - \rho_{CR}$)	Est. vs. sample reliability ($\hat{\rho}_{CR} - r_{CR}$)
25	3	0.6	15.2 %	-9.3 %	24.5 %	4.6 %	1.3 %	3.3 %
50	3	0.6	10.7 %	-7.6 %	18.3 %	2.5 %	0.2 %	2.2 %
75	3	0.6	8.1 %	-6.3 %	14.4 %	1.3 %	-0.2 %	1.5 %
100	3	0.6	6.8 %	-5.2 %	12.0 %	1.1 %	-0.1 %	1.2 %
150	3	0.6	4.8 %	-3.7 %	8.5 %	0.6 %	-0.1 %	0.7 %
300	3	0.6	2.4 %	-1.9 %	4.3 %	0.3 %	0.0 %	0.3 %
500	3	0.6	1.4 %	-1.2 %	2.6 %	0.1 %	0.0 %	0.2 %
1000	3	0.6	0.7 %	-0.6 %	1.3 %	0.1 %	0.0 %	0.1 %
25	5	0.6	8.4 %	-8.4 %	16.9 %	0.3 %	-0.2 %	0.5 %
50	5	0.6	4.2 %	-4.6 %	8.8 %	-0.4 %	-0.4 %	0.0 %
75	5	0.6	2.6 %	-2.9 %	5.5 %	-0.2 %	-0.2 %	0.0 %
100	5	0.6	1.8 %	-2.1 %	3.9 %	-0.2 %	-0.2 %	0.0 %
150	5	0.6	1.1 %	-1.4 %	2.5 %	-0.1 %	-0.2 %	0.1 %
300	5	0.6	0.6 %	-0.6 %	1.2 %	0.0 %	0.0 %	0.0 %
500	5	0.6	0.3 %	-0.4 %	0.7 %	0.0 %	-0.1 %	0.0 %
1000	5	0.6	0.2 %	-0.2 %	0.3 %	0.0 %	0.0 %	0.0 %
25	7	0.6	5.1 %	-6.5 %	11.6 %	-0.6 %	-0.4 %	-0.2 %
50	7	0.6	2.1 %	-2.8 %	4.9 %	-0.4 %	-0.3 %	-0.1 %
75	7	0.6	1.3 %	-1.8 %	3.1 %	-0.3 %	-0.3 %	0.0 %
100	7	0.6	0.9 %	-1.3 %	2.2 %	-0.2 %	-0.2 %	0.0 %
150	7	0.6	0.6 %	-0.8 %	1.4 %	-0.1 %	-0.1 %	0.0 %
300	7	0.6	0.3 %	-0.4 %	0.7 %	-0.1 %	-0.1 %	0.0 %
500	7	0.6	0.2 %	-0.3 %	0.4 %	-0.1 %	0.0 %	0.0 %
1000	7	0.6	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
25	9	0.6	3.4 %	-5.0 %	8.3 %	-0.8 %	-0.5 %	-0.2 %
50	9	0.6	1.3 %	-2.1 %	3.4 %	-0.4 %	-0.3 %	-0.1 %
75	9	0.6	0.9 %	-1.3 %	2.2 %	-0.2 %	-0.2 %	0.0 %
100	9	0.6	0.6 %	-1.0 %	1.6 %	-0.2 %	-0.1 %	0.0 %
150	9	0.6	0.4 %	-0.6 %	1.0 %	-0.1 %	-0.1 %	0.0 %
300	9	0.6	0.2 %	-0.3 %	0.5 %	-0.1 %	-0.1 %	0.0 %
500	9	0.6	0.1 %	-0.2 %	0.3 %	0.0 %	0.0 %	0.0 %
1000	9	0.6	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
25	3	0.7	6.8 %	-6.8 %	13.6 %	1.0 %	-0.1 %	1.1 %

50	3	0.7	4.8 %	-4.7 %	9.4 %	0.5 %	-0.2 %	0.7 %
75	3	0.7	3.7 %	-3.6 %	7.3 %	0.2 %	-0.3 %	0.5 %
100	3	0.7	3.0 %	-2.8 %	5.7 %	0.2 %	-0.2 %	0.4 %
150	3	0.7	2.1 %	-1.9 %	4.0 %	0.1 %	-0.1 %	0.2 %
300	3	0.7	1.0 %	-0.9 %	2.0 %	0.0 %	0.0 %	0.1 %
500	3	0.7	0.6 %	-0.6 %	1.2 %	0.0 %	0.0 %	0.0 %
1000	3	0.7	0.3 %	-0.3 %	0.6 %	0.0 %	0.0 %	0.0 %
25	5	0.7	3.4 %	-4.7 %	8.1 %	-0.5 %	-0.6 %	0.1 %
50	5	0.7	1.7 %	-2.2 %	3.9 %	-0.4 %	-0.3 %	0.0 %
75	5	0.7	1.1 %	-1.4 %	2.4 %	-0.2 %	-0.2 %	0.0 %
100	5	0.7	0.7 %	-1.0 %	1.7 %	-0.2 %	-0.2 %	0.0 %
150	5	0.7	0.5 %	-0.7 %	1.1 %	-0.1 %	-0.1 %	0.0 %
300	5	0.7	0.2 %	-0.3 %	0.5 %	0.0 %	0.0 %	0.0 %
500	5	0.7	0.1 %	-0.2 %	0.3 %	0.0 %	0.0 %	0.0 %
1000	5	0.7	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
25	7	0.7	2.0 %	-3.2 %	5.2 %	-0.6 %	-0.5 %	-0.1 %
50	7	0.7	0.8 %	-1.3 %	2.1 %	-0.3 %	-0.2 %	-0.1 %
75	7	0.7	0.5 %	-0.9 %	1.4 %	-0.2 %	-0.2 %	0.0 %
100	7	0.7	0.4 %	-0.6 %	1.0 %	-0.1 %	-0.1 %	0.0 %
150	7	0.7	0.3 %	-0.4 %	0.6 %	-0.1 %	-0.1 %	0.0 %
300	7	0.7	0.1 %	-0.2 %	0.3 %	0.0 %	-0.1 %	0.0 %
500	7	0.7	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
1000	7	0.7	0.0 %	-0.1 %	0.1 %	0.0 %	0.0 %	0.0 %
25	9	0.7	1.3 %	-2.3 %	3.6 %	-0.5 %	-0.4 %	-0.1 %
50	9	0.7	0.5 %	-1.0 %	1.5 %	-0.2 %	-0.2 %	0.0 %
75	9	0.7	0.4 %	-0.6 %	1.0 %	-0.1 %	-0.1 %	0.0 %
100	9	0.7	0.3 %	-0.5 %	0.7 %	-0.1 %	-0.1 %	0.0 %
150	9	0.7	0.2 %	-0.3 %	0.5 %	-0.1 %	-0.1 %	0.0 %
300	9	0.7	0.1 %	-0.2 %	0.2 %	0.0 %	0.0 %	0.0 %
500	9	0.7	0.1 %	-0.1 %	0.1 %	0.0 %	0.0 %	0.0 %
1000	9	0.7	0.0 %	0.0 %	0.1 %	0.0 %	0.0 %	0.0 %
25	3	0.8	2.0 %	-3.9 %	5.9 %	-0.1 %	-0.6 %	0.5 %
50	3	0.8	1.6 %	-2.4 %	4.0 %	0.0 %	-0.3 %	0.3 %
75	3	0.8	1.4 %	-1.8 %	3.1 %	-0.1 %	-0.2 %	0.2 %
100	3	0.8	1.2 %	-1.4 %	2.6 %	0.0 %	-0.1 %	0.1 %
150	3	0.8	0.9 %	-0.9 %	1.8 %	0.0 %	-0.1 %	0.0 %
300	3	0.8	0.5 %	-0.5 %	0.9 %	0.0 %	0.0 %	0.0 %
500	3	0.8	0.3 %	-0.3 %	0.6 %	0.0 %	0.0 %	0.0 %
1000	3	0.8	0.1 %	-0.1 %	0.3 %	0.0 %	0.0 %	0.0 %
25	5	0.8	1.1 %	-2.0 %	3.1 %	-0.4 %	-0.5 %	0.1 %
50	5	0.8	0.7 %	-1.0 %	1.6 %	-0.2 %	-0.2 %	0.0 %
75	5	0.8	0.4 %	-0.6 %	1.1 %	-0.1 %	-0.1 %	0.0 %
100	5	0.8	0.3 %	-0.4 %	0.7 %	-0.1 %	-0.1 %	0.0 %
150	5	0.8	0.2 %	-0.3 %	0.5 %	-0.1 %	-0.1 %	0.0 %
300	5	0.8	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
500	5	0.8	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
1000	5	0.8	0.0 %	0.0 %	0.1 %	0.0 %	0.0 %	0.0 %
25	7	0.8	0.7 %	-1.2 %	1.9 %	-0.3 %	-0.3 %	0.0 %
50	7	0.8	0.3 %	-0.5 %	0.9 %	-0.2 %	-0.1 %	0.0 %
75	7	0.8	0.2 %	-0.4 %	0.6 %	-0.1 %	-0.1 %	0.0 %
100	7	0.8	0.1 %	-0.3 %	0.4 %	-0.1 %	-0.1 %	0.0 %
150	7	0.8	0.1 %	-0.2 %	0.3 %	-0.1 %	0.0 %	0.0 %
300	7	0.8	0.0 %	-0.1 %	0.1 %	0.0 %	0.0 %	0.0 %
500	7	0.8	0.0 %	-0.1 %	0.1 %	0.0 %	0.0 %	0.0 %
1000	7	0.8	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
25	9	0.8	0.4 %	-0.9 %	1.3 %	-0.3 %	-0.3 %	0.0 %

50	9	0.8	0.2 %	-0.4 %	0.6 %	-0.1 %	-0.1 %	0.0 %
75	9	0.8	0.1 %	-0.2 %	0.4 %	-0.1 %	-0.1 %	0.0 %
100	9	0.8	0.1 %	-0.2 %	0.3 %	-0.1 %	-0.1 %	0.0 %
150	9	0.8	0.1 %	-0.1 %	0.2 %	0.0 %	0.0 %	0.0 %
300	9	0.8	0.0 %	-0.1 %	0.1 %	0.0 %	0.0 %	0.0 %
500	9	0.8	0.0 %	0.0 %	0.1 %	0.0 %	0.0 %	0.0 %
1000	9	0.8	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %

Note: The values shown are calculated as (average estimate for a condition – population value) / population value. Est. = Estimate, Pop. = Population

Table B2

Full Simulation Results for Unequal Loadings (RMSE)

Sample Size	Indicators	Loadings	Difference between population and sample statistics			
			Maximal reliability		Composite reliability	
			Population RMSE ($\hat{\rho}_{max} - \rho_{max}$)	Sample RMSE ($r_{max} - \rho_{max}$)	Population RMSE ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample RMSE ($r_{CR} - \rho_{CR}$)
25	3	0.6	24.80%	23.80%	19.00%	18.60%
50	3	0.6	19.70%	17.70%	13.80%	13.20%
75	3	0.6	16.90%	14.30%	11.40%	10.80%
100	3	0.6	14.80%	12.30%	9.80%	9.50%
150	3	0.6	12.30%	9.50%	8.20%	7.60%
300	3	0.6	8.00%	6.20%	5.70%	5.40%
500	3	0.6	5.70%	4.50%	4.40%	4.30%
1000	3	0.6	3.80%	2.90%	3.20%	3.00%
25	5	0.6	12.60%	17.90%	10.90%	12.40%
50	5	0.6	8.70%	11.20%	8.10%	8.80%
75	5	0.6	6.70%	8.20%	6.60%	7.20%
100	5	0.6	5.70%	6.60%	5.70%	6.20%
150	5	0.6	4.50%	5.00%	4.60%	5.00%
300	5	0.6	3.20%	3.40%	3.20%	3.50%
500	5	0.6	2.40%	2.50%	2.40%	2.70%
1000	5	0.6	1.70%	1.70%	1.80%	1.90%
25	7	0.6	8.00%	14.20%	8.50%	9.60%
50	7	0.6	5.10%	7.70%	5.80%	6.60%
75	7	0.6	4.20%	5.60%	4.60%	5.30%
100	7	0.6	3.50%	4.70%	3.90%	4.60%
150	7	0.6	2.90%	3.50%	3.30%	3.60%
300	7	0.6	2.10%	2.40%	2.30%	2.60%
500	7	0.6	1.60%	1.80%	1.80%	2.00%
1000	7	0.6	1.10%	1.30%	1.30%	1.40%
25	9	0.6	5.60%	11.10%	6.70%	7.90%
50	9	0.6	3.90%	5.90%	4.40%	5.30%
75	9	0.6	3.20%	4.30%	3.60%	4.20%
100	9	0.6	2.70%	3.60%	3.00%	3.70%
150	9	0.6	2.20%	2.80%	2.40%	3.00%
300	9	0.6	1.50%	2.00%	1.70%	2.00%
500	9	0.6	1.20%	1.50%	1.30%	1.60%
1000	9	0.6	0.80%	1.10%	1.00%	1.10%
25	3	0.7	13.90%	16.30%	11.80%	12.40%
50	3	0.7	10.90%	11.20%	8.50%	8.60%
75	3	0.7	9.30%	8.70%	6.80%	7.00%
100	3	0.7	7.90%	7.20%	5.80%	6.00%

150	3	0.7	6.60%	5.50%	4.80%	4.80%
300	3	0.7	4.30%	3.50%	3.40%	3.50%
500	3	0.7	3.30%	2.60%	2.70%	2.70%
1000	3	0.7	2.20%	1.80%	1.90%	1.90%
25	5	0.7	6.80%	10.80%	7.10%	8.10%
50	5	0.7	4.70%	6.20%	4.90%	5.50%
75	5	0.7	3.60%	4.60%	4.00%	4.50%
100	5	0.7	3.20%	3.80%	3.40%	3.90%
150	5	0.7	2.50%	2.90%	2.70%	3.00%
300	5	0.7	1.80%	2.00%	1.90%	2.20%
500	5	0.7	1.30%	1.50%	1.40%	1.70%
1000	5	0.7	0.90%	1.10%	1.10%	1.20%
25	7	0.7	4.40%	7.80%	5.30%	6.20%
50	7	0.7	2.90%	4.20%	3.30%	4.00%
75	7	0.7	2.40%	3.10%	2.80%	3.20%
100	7	0.7	2.00%	2.60%	2.30%	2.80%
150	7	0.7	1.70%	2.00%	1.80%	2.30%
300	7	0.7	1.10%	1.50%	1.30%	1.60%
500	7	0.7	0.90%	1.10%	1.00%	1.30%
1000	7	0.7	0.70%	0.80%	0.70%	0.90%
25	9	0.7	3.10%	5.80%	4.00%	4.90%
50	9	0.7	2.20%	3.20%	2.60%	3.20%
75	9	0.7	1.80%	2.40%	2.10%	2.60%
100	9	0.7	1.50%	2.10%	1.80%	2.20%
150	9	0.7	1.20%	1.50%	1.40%	1.80%
300	9	0.7	0.90%	1.10%	1.00%	1.20%
500	9	0.7	0.70%	0.90%	0.80%	1.00%
1000	9	0.7	0.40%	0.50%	0.60%	0.70%
25	3	0.8	6.90%	9.10%	6.70%	7.30%
50	3	0.8	5.40%	5.70%	4.60%	5.10%
75	3	0.8	4.80%	4.30%	3.80%	4.00%
100	3	0.8	4.20%	3.70%	3.20%	3.60%
150	3	0.8	3.70%	2.70%	2.60%	2.70%
300	3	0.8	2.50%	1.80%	1.80%	2.00%
500	3	0.8	1.90%	1.40%	1.40%	1.50%
1000	3	0.8	1.30%	0.90%	1.10%	1.10%
25	5	0.8	3.20%	5.00%	4.00%	4.80%
50	5	0.8	2.40%	2.80%	2.70%	3.10%
75	5	0.8	1.80%	2.20%	2.10%	2.60%
100	5	0.8	1.60%	1.70%	1.80%	2.10%
150	5	0.8	1.30%	1.40%	1.40%	1.80%
300	5	0.8	0.90%	0.90%	1.00%	1.20%
500	5	0.8	0.60%	0.60%	0.80%	0.90%
1000	5	0.8	0.40%	0.40%	0.60%	0.70%
25	7	0.8	2.10%	3.30%	2.90%	3.50%
50	7	0.8	1.40%	1.80%	1.80%	2.30%

75	7	0.8	1.20%	1.40%	1.50%	1.80%
100	7	0.8	0.90%	1.20%	1.20%	1.50%
150	7	0.8	0.70%	0.90%	1.00%	1.30%
300	7	0.8	0.50%	0.60%	0.80%	0.90%
500	7	0.8	0.40%	0.50%	0.50%	0.60%
1000	7	0.8	0.30%	0.30%	0.40%	0.40%
25	9	0.8	1.50%	2.40%	2.10%	2.80%
50	9	0.8	0.90%	1.40%	1.40%	1.80%
75	9	0.8	0.80%	1.00%	1.20%	1.40%
100	9	0.8	0.60%	0.80%	1.00%	1.30%
150	9	0.8	0.50%	0.70%	0.70%	1.00%
300	9	0.8	0.40%	0.50%	0.50%	0.70%
500	9	0.8	0.30%	0.30%	0.40%	0.50%
1000	9	0.8	0.20%	0.20%	0.30%	0.40%

Note: Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Supplemental Material, “A Note on the Finite Sample Behavior of Maximal Reliability”

Appendix C – Full Simulation Results for Equal Loadings

Table C1

Full Simulation Results for Equal Loadings (Bias)

Sample Size	Indicators	Loadings	Relative difference					
			Maximal reliability			Composite reliability		
			Estimate vs. pop. ($\hat{\rho}_{max} - \rho_{max}$)	Sample reliability vs. pop ($r_{max} - \rho_{max}$)	Est. vs. sample reliability ($\hat{\rho}_{max} - r_{max}$)	Estimate vs. pop. ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample reliability vs. pop ($r_{CR} - \rho_{CR}$)	Est. vs. sample reliability ($\hat{\rho}_{CR} - r_{CR}$)
25	3	0.6	18.2%	-8.9%	27.1%	4.9%	1.5%	3.4%
50	3	0.6	12.3%	-7.8%	20.1%	2.4%	0.3%	2.1%
75	3	0.6	8.6%	-6.5%	15.1%	1.1%	-0.2%	1.3%
100	3	0.6	6.6%	-5.1%	11.7%	0.9%	-0.2%	1.0%
150	3	0.6	4.2%	-3.4%	7.6%	0.4%	-0.1%	0.5%
300	3	0.6	1.8%	-1.6%	3.4%	0.2%	0.0%	0.2%
500	3	0.6	1.0%	-1.0%	2.0%	0.1%	0.0%	0.1%
1000	3	0.6	0.5%	-0.5%	1.0%	0.1%	0.0%	0.1%
25	5	0.6	10.2%	-9.4%	19.5%	0.3%	-0.1%	0.5%
50	5	0.6	4.4%	-5.0%	9.4%	-0.4%	-0.4%	0.0%
75	5	0.6	2.6%	-3.1%	5.7%	-0.2%	-0.2%	0.0%
100	5	0.6	1.8%	-2.2%	4.1%	-0.2%	-0.2%	0.0%
150	5	0.6	1.2%	-1.4%	2.6%	-0.1%	-0.2%	0.1%
300	5	0.6	0.6%	-0.7%	1.3%	0.0%	0.0%	0.0%
500	5	0.6	0.3%	-0.4%	0.8%	0.0%	-0.1%	0.0%
1000	5	0.6	0.2%	-0.2%	0.4%	0.0%	0.0%	0.0%
25	7	0.6	6.1%	-7.4%	13.4%	-0.7%	-0.4%	-0.2%
50	7	0.6	2.4%	-3.2%	5.5%	-0.4%	-0.3%	-0.1%
75	7	0.6	1.5%	-2.1%	3.5%	-0.3%	-0.3%	0.0%
100	7	0.6	1.1%	-1.5%	2.5%	-0.2%	-0.2%	0.0%
150	7	0.6	0.7%	-0.9%	1.6%	-0.1%	-0.1%	0.0%
300	7	0.6	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%
500	7	0.6	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
1000	7	0.6	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	9	0.6	4.0%	-5.7%	9.7%	-0.8%	-0.5%	-0.3%
50	9	0.6	1.6%	-2.4%	4.0%	-0.4%	-0.3%	-0.1%
75	9	0.6	1.0%	-1.5%	2.5%	-0.2%	-0.2%	0.0%
100	9	0.6	0.7%	-1.1%	1.8%	-0.2%	-0.2%	0.0%
150	9	0.6	0.5%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
300	9	0.6	0.2%	-0.4%	0.6%	-0.1%	-0.1%	0.0%

500	9	0.6	0.2%	-0.2%	0.3%	0.0%	0.0%	0.0%
1000	9	0.6	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	3	0.7	8.9%	-6.9%	15.8%	0.9%	0.0%	0.9%
50	3	0.7	5.6%	-4.8%	10.4%	0.4%	-0.2%	0.6%
75	3	0.7	3.5%	-3.5%	7.0%	0.0%	-0.3%	0.3%
100	3	0.7	2.7%	-2.6%	5.2%	0.1%	-0.2%	0.3%
150	3	0.7	1.6%	-1.6%	3.3%	0.0%	-0.1%	0.1%
300	3	0.7	0.7%	-0.8%	1.5%	0.0%	0.0%	0.0%
500	3	0.7	0.4%	-0.5%	0.9%	0.0%	0.0%	0.0%
1000	3	0.7	0.2%	-0.2%	0.5%	0.0%	0.0%	0.0%
25	5	0.7	4.2%	-5.5%	9.8%	-0.6%	-0.6%	0.0%
50	5	0.7	1.7%	-2.5%	4.2%	-0.4%	-0.3%	0.0%
75	5	0.7	1.1%	-1.5%	2.6%	-0.2%	-0.2%	0.0%
100	5	0.7	0.8%	-1.1%	1.9%	-0.2%	-0.2%	0.0%
150	5	0.7	0.5%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
300	5	0.7	0.3%	-0.3%	0.6%	0.0%	0.0%	0.0%
500	5	0.7	0.1%	-0.2%	0.4%	0.0%	0.0%	0.0%
1000	5	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	7	0.7	2.5%	-3.8%	6.3%	-0.6%	-0.5%	-0.1%
50	7	0.7	1.0%	-1.5%	2.5%	-0.3%	-0.2%	-0.1%
75	7	0.7	0.6%	-1.0%	1.6%	-0.2%	-0.2%	0.0%
100	7	0.7	0.4%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
150	7	0.7	0.3%	-0.5%	0.7%	-0.1%	-0.1%	0.0%
300	7	0.7	0.1%	-0.2%	0.4%	0.0%	-0.1%	0.0%
500	7	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	7	0.7	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
25	9	0.7	1.6%	-2.8%	4.4%	-0.5%	-0.4%	-0.1%
50	9	0.7	0.7%	-1.2%	1.8%	-0.2%	-0.2%	0.0%
75	9	0.7	0.4%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
100	9	0.7	0.3%	-0.5%	0.9%	-0.1%	-0.1%	0.0%
150	9	0.7	0.2%	-0.4%	0.6%	-0.1%	-0.1%	0.0%
300	9	0.7	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
500	9	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	9	0.7	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
25	3	0.8	4.0%	-4.2%	8.2%	-0.2%	-0.4%	0.2%
50	3	0.8	2.3%	-2.5%	4.7%	-0.1%	-0.2%	0.1%
75	3	0.8	1.4%	-1.7%	3.0%	-0.1%	-0.2%	0.1%
100	3	0.8	1.0%	-1.2%	2.2%	-0.1%	-0.1%	0.1%
150	3	0.8	0.6%	-0.8%	1.4%	-0.1%	-0.1%	0.0%
300	3	0.8	0.3%	-0.4%	0.7%	0.0%	0.0%	0.0%
500	3	0.8	0.2%	-0.2%	0.4%	0.0%	0.0%	0.0%
1000	3	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	5	0.8	1.6%	-2.7%	4.3%	-0.5%	-0.5%	0.0%
50	5	0.8	0.7%	-1.1%	1.8%	-0.2%	-0.2%	0.0%
75	5	0.8	0.4%	-0.7%	1.1%	-0.1%	-0.1%	0.0%
100	5	0.8	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%

150	5	0.8	0.2%	-0.4%	0.6%	-0.1%	-0.1%	0.0%
300	5	0.8	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
500	5	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	5	0.8	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%
25	7	0.8	1.0%	-1.8%	2.7%	-0.4%	-0.3%	0.0%
50	7	0.8	0.4%	-0.7%	1.1%	-0.2%	-0.1%	0.0%
75	7	0.8	0.3%	-0.5%	0.7%	-0.1%	-0.1%	0.0%
100	7	0.8	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
150	7	0.8	0.1%	-0.2%	0.3%	-0.1%	0.0%	0.0%
300	7	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
500	7	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
1000	7	0.8	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%
25	9	0.8	0.6%	-1.3%	1.9%	-0.3%	-0.3%	0.0%
50	9	0.8	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%
75	9	0.8	0.2%	-0.4%	0.5%	-0.1%	-0.1%	0.0%
100	9	0.8	0.1%	-0.3%	0.4%	-0.1%	-0.1%	0.0%
150	9	0.8	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
300	9	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
500	9	0.8	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%
1000	9	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Note: The values shown are calculated as (average estimate for a condition – population value) / population value. Est. = Estimate, Pop. = Population

Table C2

Full Simulation Results for Equal Loadings (RMSE)

Sample Size	Indicators	Loadings	Difference between population and sample statistics			
			Maximal reliability		Composite reliability	
			Population RMSE ($\hat{\rho}_{max} - \rho_{max}$)	Sample RMSE ($r_{max} - \rho_{max}$)	Population RMSE ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample RMSE ($r_{CR} - \rho_{CR}$)
25	3	0.6	27.20%	24.40%	19.10%	18.60%
50	3	0.6	20.70%	18.50%	14.00%	13.20%
75	3	0.6	16.60%	15.30%	11.80%	10.80%
100	3	0.6	13.70%	12.90%	10.00%	9.60%
150	3	0.6	10.40%	9.70%	8.40%	7.60%
300	3	0.6	6.40%	6.20%	5.90%	5.40%
500	3	0.6	4.80%	4.60%	4.60%	4.30%
1000	3	0.6	3.30%	3.00%	3.20%	3.00%
25	5	0.6	14.00%	19.50%	11.00%	12.50%
50	5	0.6	8.80%	12.20%	8.30%	8.90%
75	5	0.6	6.90%	8.90%	6.80%	7.20%
100	5	0.6	5.80%	7.20%	5.70%	6.20%
150	5	0.6	4.60%	5.60%	4.60%	5.00%
300	5	0.6	3.30%	3.70%	3.30%	3.50%
500	5	0.6	2.40%	2.80%	2.40%	2.70%
1000	5	0.6	1.80%	1.90%	1.80%	1.90%
25	7	0.6	8.90%	15.90%	8.70%	9.80%
50	7	0.6	5.50%	8.70%	5.80%	6.60%
75	7	0.6	4.50%	6.30%	4.60%	5.40%
100	7	0.6	3.90%	5.30%	4.00%	4.60%
150	7	0.6	3.10%	4.00%	3.30%	3.80%
300	7	0.6	2.30%	2.80%	2.30%	2.60%
500	7	0.6	1.80%	2.10%	1.80%	2.00%
1000	7	0.6	1.30%	1.50%	1.30%	1.40%
25	9	0.6	6.20%	12.60%	6.80%	7.90%
50	9	0.6	4.20%	6.70%	4.60%	5.40%
75	9	0.6	3.50%	4.90%	3.60%	4.20%
100	9	0.6	3.00%	4.20%	3.00%	3.70%
150	9	0.6	2.40%	3.20%	2.50%	3.00%
300	9	0.6	1.70%	2.20%	1.80%	2.20%
500	9	0.6	1.30%	1.70%	1.30%	1.70%
1000	9	0.6	1.00%	1.20%	1.00%	1.20%
25	3	0.7	15.50%	17.10%	12.10%	12.40%
50	3	0.7	11.20%	12.00%	8.60%	8.80%
75	3	0.7	8.50%	9.30%	7.10%	7.00%
100	3	0.7	7.00%	7.50%	6.10%	6.20%

150	3	0.7	5.40%	5.80%	5.00%	4.80%
300	3	0.7	3.60%	3.80%	3.50%	3.50%
500	3	0.7	2.70%	2.80%	2.70%	2.70%
1000	3	0.7	1.90%	1.90%	1.90%	1.90%
25	5	0.7	7.60%	12.40%	7.20%	8.20%
50	5	0.7	5.00%	7.10%	5.00%	5.70%
75	5	0.7	4.00%	5.20%	4.00%	4.50%
100	5	0.7	3.40%	4.20%	3.40%	3.90%
150	5	0.7	2.80%	3.40%	2.80%	3.10%
300	5	0.7	1.90%	2.30%	1.90%	2.20%
500	5	0.7	1.40%	1.70%	1.40%	1.70%
1000	5	0.7	1.10%	1.20%	1.10%	1.20%
25	7	0.7	4.80%	9.30%	5.40%	6.30%
50	7	0.7	3.20%	4.90%	3.40%	4.10%
75	7	0.7	2.60%	3.70%	2.80%	3.30%
100	7	0.7	2.30%	3.10%	2.30%	2.90%
150	7	0.7	1.80%	2.40%	2.00%	2.30%
300	7	0.7	1.30%	1.70%	1.40%	1.60%
500	7	0.7	1.00%	1.30%	1.00%	1.30%
1000	7	0.7	0.70%	0.90%	0.70%	0.90%
25	9	0.7	3.60%	7.00%	4.10%	5.00%
50	9	0.7	2.50%	3.80%	2.70%	3.30%
75	9	0.7	2.00%	2.90%	2.10%	2.60%
100	9	0.7	1.80%	2.50%	1.80%	2.20%
150	9	0.7	1.50%	1.90%	1.50%	1.80%
300	9	0.7	1.00%	1.30%	1.00%	1.30%
500	9	0.7	0.80%	1.00%	0.80%	1.00%
1000	9	0.7	0.60%	0.70%	0.60%	0.70%
25	3	0.8	8.20%	10.30%	7.00%	7.50%
50	3	0.8	5.60%	6.80%	4.90%	5.10%
75	3	0.8	4.30%	5.00%	3.90%	4.00%
100	3	0.8	3.40%	4.20%	3.30%	3.60%
150	3	0.8	2.70%	3.10%	2.70%	2.90%
300	3	0.8	1.90%	2.10%	1.90%	2.00%
500	3	0.8	1.40%	1.70%	1.40%	1.50%
1000	3	0.8	1.10%	1.10%	1.10%	1.10%
25	5	0.8	4.00%	6.80%	4.10%	4.90%
50	5	0.8	2.70%	3.80%	2.70%	3.20%
75	5	0.8	2.10%	2.90%	2.10%	2.60%
100	5	0.8	1.80%	2.30%	1.90%	2.20%
150	5	0.8	1.40%	1.90%	1.40%	1.80%
300	5	0.8	1.00%	1.20%	1.00%	1.20%
500	5	0.8	0.80%	1.00%	0.80%	1.00%
1000	5	0.8	0.60%	0.70%	0.60%	0.70%
25	7	0.8	2.60%	4.80%	2.90%	3.60%
50	7	0.8	1.70%	2.60%	1.80%	2.30%

75	7	0.8	1.40%	2.10%	1.50%	1.80%
100	7	0.8	1.20%	1.70%	1.30%	1.60%
150	7	0.8	1.00%	1.30%	1.10%	1.30%
300	7	0.8	0.80%	1.00%	0.80%	0.90%
500	7	0.8	0.50%	0.80%	0.50%	0.60%
1000	7	0.8	0.40%	0.50%	0.40%	0.50%
25	9	0.8	1.90%	3.60%	2.20%	2.90%
50	9	0.8	1.40%	2.00%	1.40%	1.80%
75	9	0.8	1.10%	1.60%	1.20%	1.50%
100	9	0.8	1.00%	1.40%	1.00%	1.30%
150	9	0.8	0.70%	1.10%	0.70%	1.00%
300	9	0.8	0.50%	0.70%	0.50%	0.70%
500	9	0.8	0.40%	0.50%	0.40%	0.50%
1000	9	0.8	0.30%	0.40%	0.30%	0.40%

Note: Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Supplemental Material, “A Note on the Finite Sample Behavior of Maximal Reliability”

Appendix D – Annotated Simulation and Analysis Code

```
# Loads the lavaan (for data analysis) and MASS (for data generation) libraries

library(lavaan)
library(MASS)

# Defines the function that calculates the reliability of a composite based on loadings (l),
# errors (r), and weights (w)

cr.implied <- function(l, r, w){

  numer <- sum(w * l) ^ 2
  denom <- numer + sum((w ^ 2) * r)
  cr <- numer / denom
  return(cr)

}

# These are the different simulation conditions and values of the levels within each

SSIZE <- c(25, 50, 75, 100, 150, 300, 500, 1000)
INDICATORS <- c(3, 5, 7, 9)
LOADINGS <- c(.6, .7, .8)
EQFACTOR <- c(0, 1)
LOAD.SD <- .1
REPLICATIONS <- 10000

design.matrix <- expand.grid(ssize = SSIZE,
                             indic = INDICATORS,
                             loads = LOADINGS,
                             eqfactor = EQFACTOR)

# The following block of code is repeated for each different condition in the simulation

for(i in 1:nrow(design.matrix)){

  design <- design.matrix[i,]

  # This defines the single-factor analysis model

  lav.model <- paste(paste('T =~ ', paste('X', 1:design$indic, sep = '', collapse = ' + '), sep = ''),
                     'T ~~ 1*T', sep = '\n')

  # This defines the single-factor analysis model with constrained error variances

  lav.model.negvar <- paste(lav.model,
                            paste('X', 1:design$indic, ' ~~ ', 'varx', 1:design$indic,'*X',
                                  1:design$indic, sep = '', collapse = '\n'),
                            paste('varx', 1:design$indic, ' > 0', sep = '', collapse = '\n'),
                            sep = '\n')

  # Defines the loadings, whether equal or unequal

  if(design$eqfactor == 1){

    loadings.vector <- rep(design$loads, design$indic)

  } else {

    loadings.vector <- as.numeric(scale(1:design$indic) * LOAD.SD + design$loads)
```

```
  }

  # Defines the vector of errors that gives the indicators unit variance

  errors.vector <- (1 - loadings.vector**2)

  # The following block of code returns a list of latent variable case values and indicator data

  set.seed(1)

  sim.data <- lapply(1:REPLICATIONS, function(x){

    lv <- rnorm(design$ssize, 0, sqrt(1))

    eresid <- mvrnorm(n = design$ssize,
                      mu = rep(0, design$indic),
                      Sigma = diag(errors.vector))

    dat <- lv %*% t(loadings.vector) + eresid

    colnames(dat) <- c(paste('X', 1:design$indic, sep = ''))

    list(lv = lv, data = dat)

  })

  # The following block of code performs the main analyses

  results <- lapply(sim.data, function(x){

    neg.var <- FALSE

    # Executes a confirmatory factor analysis with the unconstrained model and factor variance equal to 1

    lav.cfa <- cfa(lav.model, data = as.data.frame(x$data), std.lv=TRUE)

    # Checks for convergence; the rest of the code is not executed if the analysis did not converge

    if(inspect(lav.cfa, 'converged')){

      # Extracts the indicator errors from the results of the confirmatory factor analysis

      errors <- coef(lav.cfa)[c(paste('X', 1:design$indic, '~~X', 1:design$indic, sep = ''))]

      # Checks if any errors are negative; the code inside this block only executes if one or more errors
are negative
          # The neg.var flag is set to identify the replication in later analysis
          # The confirmatory factor analysis is run again but using the constrained model previously
defined
          # The errors from the new analysis (now constrained to be positive) are extracted

      if(any(errors < 0)){
        neg.var <- TRUE
        lav.cfa <- cfa(lav.model.negvar, data = as.data.frame(x$data), std.lv=TRUE)
        errors <- coef(lav.cfa)[c(paste('varx', 1:design$indic, sep = ''))]
      }

      # Extracts the loading estimates from the results of the confirmatory factor analysis

      loadings <- coef(lav.cfa)[c(paste('T=~X', 1:design$indic, sep = ''))]

      # Calculates the maximal reliability weights as the ratio of loadings to errors

      weights <- loadings / errors

      # Calculates the estimated maximal reliability based on loading and error estimates, and weights

      mr.estim <- cr.implied(l = loadings,
```

```
                             r = errors,
                             w = weights)

      # Calculates the estimated composite reliability based on loading and error estimates, and unit
weights

      cr.estim <- cr.implied(l = loadings,
                             r = errors,
                             w = 1)

      # Creates a composite of the indicators using the maximal reliability weights

      mr.comp <- x$data %*% weights

      # Creates a composite of the indicators using unit/equal weights

      eq.comp <- rowSums(x$data)

      # Calculates the true reliability of the maximal reliability composite as the squared correlation
      # of the composite and case values of the latent variable (from the data generation process)

      mr.true <- cor(mr.comp, x$lv)**2

      # Calculates the true reliability of the equal weights composite as the squared correlation
      # of the composite and case values of the latent variable (from the data generation process)

      cr.true <- cor(eq.comp, x$lv)**2

      # Returns a list of all results

      list(loadings = loadings,
           errors = errors,
           cr.estim = cr.estim,
           mr.estim = mr.estim,
           cr.true = cr.true,
           mr.true = mr.true,
           neg.var = neg.var)

    } else {

      # If the confirmatory factor analysis did not converge, no results are returned

      NULL

    }

  })

  # The results object is saved for later processing

  save(results, file = paste('Results_', i, '.Rdata', sep = ''))
  print(paste('Completed: ', i, ' of ', nrow(design.matrix) , sep = ''))

}
```

Supplemental Material, “A Note on the Finite Sample Behavior of Maximal Reliability”

Appendix E – Complete Results from the Simulation Including Replications that Were Rerun with Constrained Estimation

Table E1

Results for Equal Loadings (Selected Conditions)

Sample Size	Indicators	Loadings	Difference between population and sample statistics							
			Maximal reliability				Composite reliability			
			Population $(\hat{\rho}_{max} - \rho_{max})$		Sample $(r_{max} - \rho_{max})$		Population $(\hat{\rho}_{CR} - \rho_{CR})$		Sample $(r_{CR} - \rho_{CR})$	
			Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
25	3	0.6	28.4%	37.4%	-14.1%	28.7%	5.8%	18.8%	0.6%	19.0%
50	3	0.6	16.0%	26.3%	-10.0%	21.0%	2.4%	13.9%	-0.1%	13.4%
75	3	0.6	10.3%	19.4%	-7.3%	16.6%	1.3%	11.6%	-0.2%	10.8%
25	5	0.6	12.0%	16.5%	-10.7%	22.2%	-0.1%	11.7%	-0.3%	12.9%
50	5	0.6	4.5%	8.9%	-5.0%	12.5%	-0.4%	8.4%	-0.3%	8.9%
75	5	0.6	2.6%	6.9%	-3.0%	8.9%	-0.2%	6.8%	-0.2%	7.2%
25	3	0.7	12.8%	19.7%	-9.0%	19.9%	1.2%	12.0%	-0.2%	12.7%
50	3	0.7	6.3%	12.5%	-5.4%	12.9%	0.4%	8.6%	-0.3%	8.8%
75	3	0.7	3.7%	8.9%	-3.5%	9.6%	0.1%	7.1%	-0.2%	7.0%
25	5	0.7	4.7%	8.0%	-5.6%	13.2%	-0.5%	7.5%	-0.5%	8.3%
50	5	0.7	1.7%	5.0%	-2.3%	7.1%	-0.3%	5.0%	-0.3%	5.7%
75	5	0.7	1.1%	4.0%	-1.5%	5.2%	-0.2%	4.0%	-0.2%	4.5%
25	3	0.8	5.2%	9.4%	-5.3%	11.8%	-0.2%	7.0%	-0.5%	7.6%
50	3	0.8	2.4%	5.7%	-2.5%	6.9%	-0.1%	4.9%	-0.2%	5.1%
75	3	0.8	1.5%	4.3%	-1.5%	5.1%	0.0%	3.9%	0.0%	4.0%
25	5	0.8	1.7%	4.0%	-2.6%	6.8%	-0.4%	4.2%	-0.3%	4.9%
50	5	0.8	0.7%	2.7%	-1.1%	3.8%	-0.2%	2.7%	-0.2%	3.2%
75	5	0.8	0.4%	2.1%	-0.7%	2.9%	-0.1%	2.1%	-0.1%	2.6%

Note: The Mean values shown are calculated as (average estimate for a condition – population value) / population value. Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table E2

Results for Unequal Loadings (Selected Conditions)

Sample Size	Indicators	Loadings	Difference between population and sample statistics							
			Maximal reliability				Composite reliability			
			Population $(\hat{\rho}_{max} - \rho_{max})$		Sample $(r_{max} - \rho_{max})$		Population $(\hat{\rho}_{CR} - \rho_{CR})$		Sample $(r_{CR} - \rho_{CR})$	
			Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
25	3	0.6	25.7%	35.4%	-13.9%	27.8%	5.5%	18.7%	0.5%	18.7%
50	3	0.6	16.2%	27.1%	-10.0%	20.2%	2.7%	13.6%	-0.1%	13.3%
75	3	0.6	11.7%	22.0%	-7.6%	16.0%	1.7%	11.4%	0.0%	10.8%
25	5	0.6	10.8%	15.4%	-9.9%	20.4%	-0.3%	11.6%	-0.4%	12.7%
50	5	0.6	4.2%	9.1%	-4.6%	11.6%	-0.3%	8.2%	-0.3%	8.9%
75	5	0.6	2.5%	6.8%	-2.8%	8.3%	-0.2%	6.6%	-0.2%	7.2%
25	3	0.7	11.7%	18.8%	-9.1%	18.5%	1.1%	11.7%	-0.5%	12.5%
50	3	0.7	6.7%	13.8%	-5.6%	12.1%	0.4%	8.5%	-0.4%	8.6%
75	3	0.7	4.5%	10.9%	-3.9%	9.3%	0.2%	6.8%	-0.3%	7.0%
25	5	0.7	4.2%	7.6%	-5.0%	11.6%	-0.6%	7.3%	-0.6%	8.2%
50	5	0.7	1.6%	4.7%	-2.2%	6.2%	-0.4%	4.9%	-0.3%	5.5%
75	5	0.7	1.0%	3.6%	-1.3%	4.6%	-0.2%	4.0%	-0.2%	4.5%
25	3	0.8	4.4%	8.9%	-4.8%	9.9%	-0.1%	6.7%	-0.6%	7.5%
50	3	0.8	2.7%	6.7%	-2.7%	6.1%	-0.1%	4.7%	-0.2%	5.0%
75	3	0.8	1.9%	5.6%	-1.9%	4.6%	-0.1%	3.8%	-0.1%	4.0%
25	5	0.8	1.5%	3.7%	-2.2%	5.2%	-0.4%	4.1%	-0.4%	4.8%
50	5	0.8	0.7%	2.4%	-0.9%	2.9%	-0.2%	2.7%	-0.1%	3.1%
75	5	0.8	0.4%	1.8%	-0.6%	2.2%	-0.2%	2.1%	-0.1%	2.6%

Note: The Mean values shown are calculated as (average estimate for a condition – population value) / population value. Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table E3

Full Simulation Results for Unequal Loadings

Sample Size	Indicators	Loadings	Relative difference					
			Maximal reliability			Composite reliability		
			Estimate vs. pop. ($\hat{\rho}_{max} - \rho_{max}$)	Sample reliability vs. pop ($r_{max} - \rho_{max}$)	Est. vs. sample reliability ($\hat{\rho}_{max} - r_{max}$)	Estimate vs. pop. ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample reliability vs. pop ($r_{CR} - \rho_{CR}$)	Est. vs. sample reliability ($\hat{\rho}_{CR} - r_{CR}$)
25	3	0.6	25.7%	-13.9%	39.6%	5.5%	0.5%	5.0%
50	3	0.6	16.2%	-10.0%	26.2%	2.7%	-0.1%	2.8%
75	3	0.6	11.7%	-7.6%	19.3%	1.7%	0.0%	1.7%
100	3	0.6	8.5%	-5.8%	14.3%	1.2%	-0.2%	1.4%
150	3	0.6	5.6%	-4.0%	9.6%	0.7%	-0.1%	0.8%
300	3	0.6	2.6%	-2.0%	4.6%	0.3%	0.0%	0.3%
500	3	0.6	1.4%	-1.2%	2.6%	0.1%	-0.1%	0.2%
1000	3	0.6	0.7%	-0.6%	1.3%	0.1%	0.0%	0.1%
25	5	0.6	10.8%	-9.9%	20.7%	-0.3%	-0.4%	0.1%
50	5	0.6	4.2%	-4.6%	8.8%	-0.3%	-0.3%	0.0%
75	5	0.6	2.5%	-2.8%	5.3%	-0.2%	-0.2%	0.0%
100	5	0.6	1.9%	-2.0%	3.9%	-0.1%	-0.1%	0.0%
150	5	0.6	1.1%	-1.3%	2.4%	-0.1%	0.0%	-0.1%
300	5	0.6	0.5%	-0.6%	1.1%	-0.1%	0.0%	-0.1%
500	5	0.6	0.3%	-0.4%	0.7%	-0.1%	-0.1%	0.0%
1000	5	0.6	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
25	7	0.6	5.6%	-6.9%	12.5%	-0.9%	-0.4%	-0.5%
50	7	0.6	2.1%	-2.9%	5.0%	-0.4%	-0.4%	0.0%
75	7	0.6	1.3%	-1.7%	3.0%	-0.2%	-0.2%	0.0%
100	7	0.6	1.0%	-1.3%	2.3%	-0.2%	-0.2%	0.0%
150	7	0.6	0.6%	-0.8%	1.4%	-0.1%	-0.1%	0.0%
300	7	0.6	0.3%	-0.4%	0.7%	0.0%	0.0%	0.0%
500	7	0.6	0.2%	-0.2%	0.4%	0.0%	0.0%	0.0%
1000	7	0.6	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	9	0.6	3.6%	-5.0%	8.6%	-0.7%	-0.5%	-0.2%
50	9	0.6	1.4%	-2.0%	3.4%	-0.4%	-0.2%	-0.2%
75	9	0.6	0.9%	-1.3%	2.2%	-0.3%	-0.2%	-0.1%
100	9	0.6	0.6%	-1.0%	1.6%	-0.2%	-0.2%	0.0%
150	9	0.6	0.4%	-0.6%	1.0%	-0.1%	0.0%	-0.1%
300	9	0.6	0.2%	-0.3%	0.5%	-0.1%	0.0%	-0.1%
500	9	0.6	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
1000	9	0.6	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	3	0.7	11.7%	-9.1%	20.8%	1.1%	-0.5%	1.6%
50	3	0.7	6.7%	-5.6%	12.3%	0.4%	-0.4%	0.8%
75	3	0.7	4.5%	-3.9%	8.4%	0.2%	-0.3%	0.5%
100	3	0.7	3.3%	-3.0%	6.3%	0.1%	-0.2%	0.3%

150	3	0.7	2.3%	-1.9%	4.2%	0.2%	0.0%	0.2%
300	3	0.7	1.0%	-1.0%	2.0%	0.0%	-0.1%	0.1%
500	3	0.7	0.6%	-0.5%	1.1%	0.0%	0.0%	0.0%
1000	3	0.7	0.3%	-0.3%	0.6%	0.0%	0.0%	0.0%
25	5	0.7	4.2%	-5.0%	9.2%	-0.6%	-0.6%	0.0%
50	5	0.7	1.6%	-2.2%	3.8%	-0.4%	-0.3%	-0.1%
75	5	0.7	1.0%	-1.3%	2.3%	-0.2%	-0.2%	0.0%
100	5	0.7	0.7%	-1.0%	1.7%	-0.2%	-0.2%	0.0%
150	5	0.7	0.5%	-0.6%	1.1%	-0.1%	-0.1%	0.0%
300	5	0.7	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
500	5	0.7	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
1000	5	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	7	0.7	2.1%	-3.1%	5.2%	-0.6%	-0.4%	-0.2%
50	7	0.7	0.9%	-1.3%	2.2%	-0.3%	-0.2%	-0.1%
75	7	0.7	0.5%	-0.8%	1.3%	-0.2%	-0.2%	0.0%
100	7	0.7	0.4%	-0.6%	1.0%	-0.1%	-0.1%	0.0%
150	7	0.7	0.2%	-0.4%	0.6%	-0.1%	-0.1%	0.0%
300	7	0.7	0.1%	-0.2%	0.3%	-0.1%	-0.1%	0.0%
500	7	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	7	0.7	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
25	9	0.7	1.3%	-2.2%	3.5%	-0.5%	-0.4%	-0.1%
50	9	0.7	0.5%	-1.0%	1.5%	-0.3%	-0.2%	-0.1%
75	9	0.7	0.3%	-0.6%	0.9%	-0.2%	-0.2%	0.0%
100	9	0.7	0.2%	-0.5%	0.7%	-0.1%	-0.1%	0.0%
150	9	0.7	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
300	9	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
500	9	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	9	0.7	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25	3	0.8	4.4%	-4.8%	9.2%	-0.1%	-0.6%	0.5%
50	3	0.8	2.7%	-2.7%	5.4%	-0.1%	-0.2%	0.1%
75	3	0.8	1.9%	-1.9%	3.8%	-0.1%	-0.1%	0.0%
100	3	0.8	1.5%	-1.5%	3.0%	0.0%	-0.1%	0.1%
150	3	0.8	1.0%	-1.0%	2.0%	0.0%	-0.1%	0.1%
300	3	0.8	0.5%	-0.5%	1.0%	0.0%	-0.1%	0.1%
500	3	0.8	0.3%	-0.3%	0.6%	0.0%	0.0%	0.0%
1000	3	0.8	0.2%	-0.1%	0.3%	0.0%	0.0%	0.0%
25	5	0.8	1.5%	-2.2%	3.7%	-0.4%	-0.4%	0.0%
50	5	0.8	0.7%	-0.9%	1.6%	-0.2%	-0.1%	-0.1%
75	5	0.8	0.4%	-0.6%	1.0%	-0.2%	-0.1%	-0.1%
100	5	0.8	0.3%	-0.4%	0.7%	-0.1%	-0.1%	0.0%
150	5	0.8	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
300	5	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
500	5	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
1000	5	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25	7	0.8	0.8%	-1.3%	2.1%	-0.4%	-0.3%	-0.1%
50	7	0.8	0.3%	-0.6%	0.9%	-0.2%	-0.2%	0.0%

75	7	0.8	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
100	7	0.8	0.2%	-0.2%	0.4%	-0.1%	0.0%	-0.1%
150	7	0.8	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
300	7	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
500	7	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
1000	7	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25	9	0.8	0.5%	-0.9%	1.4%	-0.3%	-0.2%	-0.1%
50	9	0.8	0.2%	-0.4%	0.6%	-0.1%	-0.1%	0.0%
75	9	0.8	0.1%	-0.2%	0.3%	-0.1%	-0.1%	0.0%
100	9	0.8	0.1%	-0.2%	0.3%	0.0%	-0.1%	0.1%
150	9	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
300	9	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
500	9	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
1000	9	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Note: The values shown are calculated as (average estimate for a condition – population value) / population value.

Table E4

Full Simulation Results for Equal Loadings

Sample Size	Indicators	Loadings	Relative difference					
			Maximal reliability			Composite reliability		
			Estimate vs. pop. ($\hat{\rho}_{max} - \rho_{max}$)	Sample reliability vs. pop ($r_{max} - \rho_{max}$)	Est. vs. sample reliability ($\hat{\rho}_{max} - r_{max}$)	Estimate vs. pop. ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample reliability vs. pop ($r_{CR} - \rho_{CR}$)	Est. vs. sample reliability ($\hat{\rho}_{CR} - r_{CR}$)
25	3	0.6	28.4%	-14.1%	42.5%	5.8%	0.6%	5.2%
50	3	0.6	16.0%	-10.0%	26.0%	2.4%	-0.1%	2.5%
75	3	0.6	10.3%	-7.3%	17.6%	1.3%	-0.2%	1.5%
100	3	0.6	7.5%	-5.5%	13.0%	1.0%	-0.2%	1.2%
150	3	0.6	4.2%	-3.7%	7.9%	0.4%	-0.2%	0.6%
300	3	0.6	1.9%	-1.6%	3.5%	0.2%	0.0%	0.2%
500	3	0.6	1.1%	-0.9%	2.0%	0.1%	0.0%	0.1%
1000	3	0.6	0.5%	-0.5%	1.0%	0.1%	0.0%	0.1%
25	5	0.6	12.0%	-10.7%	22.7%	-0.1%	-0.3%	0.2%
50	5	0.6	4.5%	-5.0%	9.5%	-0.4%	-0.3%	-0.1%
75	5	0.6	2.6%	-3.0%	5.6%	-0.2%	-0.2%	0.0%
100	5	0.6	1.8%	-2.2%	4.0%	-0.2%	-0.2%	0.0%
150	5	0.6	1.2%	-1.4%	2.6%	-0.1%	-0.1%	0.0%
300	5	0.6	0.6%	-0.7%	1.3%	0.0%	0.0%	0.0%
500	5	0.6	0.3%	-0.4%	0.7%	0.0%	-0.1%	0.1%
1000	5	0.6	0.2%	-0.2%	0.4%	0.0%	0.0%	0.0%
25	7	0.6	6.3%	-8.0%	14.3%	-1.1%	-0.7%	-0.4%
50	7	0.6	2.4%	-3.2%	5.6%	-0.4%	-0.3%	-0.1%
75	7	0.6	1.4%	-2.0%	3.4%	-0.3%	-0.3%	0.0%
100	7	0.6	1.0%	-1.5%	2.5%	-0.2%	-0.2%	0.0%
150	7	0.6	0.7%	-0.9%	1.6%	-0.1%	0.0%	-0.1%
300	7	0.6	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%
500	7	0.6	0.2%	-0.3%	0.5%	0.0%	0.0%	0.0%
1000	7	0.6	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	9	0.6	4.1%	-6.0%	10.1%	-0.9%	-0.5%	-0.4%
50	9	0.6	1.6%	-2.4%	4.0%	-0.4%	-0.3%	-0.1%
75	9	0.6	1.0%	-1.5%	2.5%	-0.2%	-0.2%	0.0%
100	9	0.6	0.7%	-1.1%	1.8%	-0.2%	-0.2%	0.0%
150	9	0.6	0.5%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
300	9	0.6	0.3%	-0.3%	0.6%	0.0%	0.0%	0.0%
500	9	0.6	0.2%	-0.2%	0.4%	0.0%	0.0%	0.0%
1000	9	0.6	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	3	0.7	12.8%	-9.0%	21.8%	1.2%	-0.2%	1.4%
50	3	0.7	6.3%	-5.4%	11.7%	0.4%	-0.3%	0.7%
75	3	0.7	3.7%	-3.5%	7.2%	0.1%	-0.2%	0.3%
100	3	0.7	2.6%	-2.6%	5.2%	0.0%	-0.2%	0.2%

150	3	0.7	1.6%	-1.7%	3.3%	0.0%	-0.1%	0.1%
300	3	0.7	0.8%	-0.7%	1.5%	0.0%	0.0%	0.0%
500	3	0.7	0.4%	-0.5%	0.9%	0.0%	-0.1%	0.1%
1000	3	0.7	0.2%	-0.3%	0.5%	0.0%	0.0%	0.0%
25	5	0.7	4.7%	-5.6%	10.3%	-0.5%	-0.5%	0.0%
50	5	0.7	1.7%	-2.3%	4.0%	-0.3%	-0.3%	0.0%
75	5	0.7	1.1%	-1.5%	2.6%	-0.2%	-0.2%	0.0%
100	5	0.7	0.8%	-1.1%	1.9%	-0.1%	-0.1%	0.0%
150	5	0.7	0.5%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
300	5	0.7	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
500	5	0.7	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
1000	5	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	7	0.7	2.5%	-3.8%	6.3%	-0.6%	-0.5%	-0.1%
50	7	0.7	1.0%	-1.5%	2.5%	-0.2%	-0.2%	0.0%
75	7	0.7	0.6%	-1.0%	1.6%	-0.2%	-0.2%	0.0%
100	7	0.7	0.5%	-0.7%	1.2%	-0.1%	-0.1%	0.0%
150	7	0.7	0.3%	-0.4%	0.7%	0.0%	0.0%	0.0%
300	7	0.7	0.1%	-0.2%	0.3%	-0.1%	-0.1%	0.0%
500	7	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	7	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	9	0.7	1.6%	-2.7%	4.3%	-0.5%	-0.4%	-0.1%
50	9	0.7	0.7%	-1.1%	1.8%	-0.2%	-0.1%	-0.1%
75	9	0.7	0.4%	-0.8%	1.2%	-0.2%	-0.2%	0.0%
100	9	0.7	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%
150	9	0.7	0.2%	-0.4%	0.6%	-0.1%	-0.1%	0.0%
300	9	0.7	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
500	9	0.7	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	9	0.7	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25	3	0.8	5.2%	-5.3%	10.5%	-0.2%	-0.5%	0.3%
50	3	0.8	2.4%	-2.5%	4.9%	-0.1%	-0.2%	0.1%
75	3	0.8	1.5%	-1.5%	3.0%	0.0%	0.0%	0.0%
100	3	0.8	1.0%	-1.2%	2.2%	-0.1%	-0.1%	0.0%
150	3	0.8	0.6%	-0.8%	1.4%	-0.1%	-0.1%	0.0%
300	3	0.8	0.3%	-0.4%	0.7%	0.0%	0.0%	0.0%
500	3	0.8	0.2%	-0.2%	0.4%	0.0%	-0.1%	0.1%
1000	3	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
25	5	0.8	1.7%	-2.6%	4.3%	-0.4%	-0.3%	-0.1%
50	5	0.8	0.7%	-1.1%	1.8%	-0.2%	-0.2%	0.0%
75	5	0.8	0.4%	-0.7%	1.1%	-0.1%	-0.1%	0.0%
100	5	0.8	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%
150	5	0.8	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
300	5	0.8	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
500	5	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
1000	5	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
25	7	0.8	1.0%	-1.7%	2.7%	-0.3%	-0.3%	0.0%
50	7	0.8	0.4%	-0.7%	1.1%	-0.1%	-0.1%	0.0%

75	7	0.8	0.3%	-0.4%	0.7%	-0.1%	-0.1%	0.0%
100	7	0.8	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
150	7	0.8	0.1%	-0.2%	0.3%	-0.1%	-0.1%	0.0%
300	7	0.8	0.1%	-0.1%	0.2%	0.0%	0.0%	0.0%
500	7	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
1000	7	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25	9	0.8	0.7%	-1.3%	2.0%	-0.3%	-0.2%	-0.1%
50	9	0.8	0.3%	-0.5%	0.8%	-0.1%	-0.1%	0.0%
75	9	0.8	0.2%	-0.3%	0.5%	-0.1%	-0.1%	0.0%
100	9	0.8	0.1%	-0.2%	0.3%	0.0%	-0.1%	0.1%
150	9	0.8	0.1%	-0.2%	0.3%	0.0%	0.0%	0.0%
300	9	0.8	0.0%	-0.1%	0.1%	0.0%	0.0%	0.0%
500	9	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
1000	9	0.8	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Note: The values shown are calculated as (average estimate for a condition – population value) / population value.

Table E5

Full Simulation Results for Unequal Loadings (RMSE)

Sample Size	Indicators	Loadings	Difference between population and sample statistics			
			Maximal reliability		Composite reliability	
			Population RMSE $(\hat{\rho}_{max} - \rho_{max})$	Sample RMSE $(r_{max} - \rho_{max})$	Population RMSE $(\hat{\rho}_{CR} - \rho_{CR})$	Sample RMSE $(r_{CR} - \rho_{CR})$
25	3	0.6	35.4%	27.8%	18.7%	18.7%
50	3	0.6	27.1%	20.2%	13.6%	13.3%
75	3	0.6	22.0%	16.0%	11.4%	10.8%
100	3	0.6	18.0%	13.2%	9.8%	9.5%
150	3	0.6	13.8%	9.8%	8.2%	7.6%
300	3	0.6	8.2%	6.2%	5.7%	5.4%
500	3	0.6	5.7%	4.5%	4.4%	4.3%
1000	3	0.6	3.8%	2.9%	3.2%	3.0%
25	5	0.6	15.4%	20.4%	11.6%	12.7%
50	5	0.6	9.1%	11.6%	8.2%	8.9%
75	5	0.6	6.8%	8.3%	6.6%	7.2%
100	5	0.6	5.7%	6.6%	5.7%	6.2%
150	5	0.6	4.5%	5.0%	4.6%	5.0%
300	5	0.6	3.2%	3.4%	3.2%	3.5%
500	5	0.6	2.4%	2.5%	2.4%	2.7%
1000	5	0.6	1.7%	1.7%	1.8%	1.9%
25	7	0.6	8.8%	15.4%	9.0%	9.9%
50	7	0.6	5.3%	7.7%	5.8%	6.6%
75	7	0.6	4.2%	5.6%	4.6%	5.3%
100	7	0.6	3.5%	4.7%	3.9%	4.6%
150	7	0.6	2.9%	3.5%	3.3%	3.6%
300	7	0.6	2.1%	2.4%	2.3%	2.6%
500	7	0.6	1.6%	1.8%	1.8%	2.0%
1000	7	0.6	1.1%	1.3%	1.3%	1.4%
25	9	0.6	5.9%	11.7%	6.9%	7.9%
50	9	0.6	3.9%	5.9%	4.4%	5.3%
75	9	0.6	3.2%	4.3%	3.6%	4.2%
100	9	0.6	2.7%	3.6%	3.0%	3.7%
150	9	0.6	2.2%	2.8%	2.4%	3.0%
300	9	0.6	1.5%	2.0%	1.7%	2.0%
500	9	0.6	1.2%	1.5%	1.3%	1.6%
1000	9	0.6	0.8%	1.1%	1.0%	1.1%
25	3	0.7	18.8%	18.5%	11.7%	12.5%
50	3	0.7	13.8%	12.1%	8.5%	8.6%
75	3	0.7	10.9%	9.3%	6.8%	7.0%
100	3	0.7	9.0%	7.4%	5.8%	6.0%

150	3	0.7	6.9%	5.6%	4.8%	4.8%
300	3	0.7	4.3%	3.5%	3.4%	3.5%
500	3	0.7	3.3%	2.6%	2.7%	2.7%
1000	3	0.7	2.2%	1.8%	1.9%	1.9%
25	5	0.7	7.6%	11.6%	7.3%	8.2%
50	5	0.7	4.7%	6.2%	4.9%	5.5%
75	5	0.7	3.6%	4.6%	4.0%	4.5%
100	5	0.7	3.2%	3.8%	3.4%	3.9%
150	5	0.7	2.5%	2.9%	2.7%	3.0%
300	5	0.7	1.8%	2.0%	1.9%	2.2%
500	5	0.7	1.3%	1.5%	1.4%	1.7%
1000	5	0.7	0.9%	1.1%	1.1%	1.2%
25	7	0.7	4.5%	8.0%	5.3%	6.2%
50	7	0.7	2.9%	4.2%	3.3%	4.0%
75	7	0.7	2.4%	3.1%	2.8%	3.2%
100	7	0.7	2.0%	2.6%	2.3%	2.8%
150	7	0.7	1.7%	2.0%	1.8%	2.3%
300	7	0.7	1.1%	1.5%	1.3%	1.6%
500	7	0.7	0.9%	1.1%	1.0%	1.3%
1000	7	0.7	0.7%	0.8%	0.7%	0.9%
25	9	0.7	3.2%	5.8%	4.0%	4.9%
50	9	0.7	2.2%	3.2%	2.6%	3.2%
75	9	0.7	1.8%	2.4%	2.1%	2.6%
100	9	0.7	1.5%	2.1%	1.8%	2.2%
150	9	0.7	1.2%	1.5%	1.4%	1.8%
300	9	0.7	0.9%	1.1%	1.0%	1.2%
500	9	0.7	0.7%	0.9%	0.8%	1.0%
1000	9	0.7	0.4%	0.5%	0.6%	0.7%
25	3	0.8	8.9%	9.9%	6.7%	7.5%
50	3	0.8	6.7%	6.1%	4.7%	5.0%
75	3	0.8	5.6%	4.6%	3.8%	4.0%
100	3	0.8	4.8%	3.8%	3.2%	3.6%
150	3	0.8	3.8%	2.9%	2.6%	2.7%
300	3	0.8	2.5%	1.8%	1.8%	2.0%
500	3	0.8	1.9%	1.4%	1.4%	1.5%
1000	3	0.8	1.3%	0.9%	1.1%	1.1%
25	5	0.8	3.7%	5.2%	4.1%	4.8%
50	5	0.8	2.4%	2.9%	2.7%	3.1%
75	5	0.8	1.8%	2.2%	2.1%	2.6%
100	5	0.8	1.6%	1.7%	1.8%	2.1%
150	5	0.8	1.3%	1.4%	1.4%	1.8%
300	5	0.8	0.9%	0.9%	1.0%	1.2%
500	5	0.8	0.6%	0.6%	0.8%	0.9%
1000	5	0.8	0.4%	0.4%	0.6%	0.7%
25	7	0.8	2.1%	3.4%	2.9%	3.5%
50	7	0.8	1.4%	1.8%	1.8%	2.3%

75	7	0.8	1.2%	1.4%	1.5%	1.8%
100	7	0.8	0.9%	1.2%	1.2%	1.5%
150	7	0.8	0.7%	0.9%	1.0%	1.3%
300	7	0.8	0.5%	0.6%	0.8%	0.9%
500	7	0.8	0.4%	0.5%	0.5%	0.6%
1000	7	0.8	0.3%	0.3%	0.4%	0.4%
25	9	0.8	1.5%	2.4%	2.1%	2.8%
50	9	0.8	0.9%	1.4%	1.4%	1.8%
75	9	0.8	0.8%	1.0%	1.2%	1.4%
100	9	0.8	0.6%	0.8%	1.0%	1.3%
150	9	0.8	0.5%	0.7%	0.7%	1.0%
300	9	0.8	0.4%	0.5%	0.5%	0.7%
500	9	0.8	0.3%	0.3%	0.4%	0.5%
1000	9	0.8	0.2%	0.2%	0.3%	0.4%

Note: Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Table E6

Full Simulation Results for Equal Loadings (RMSE)

Sample Size	Indicators	Loadings	Difference between population and sample statistics			
			Maximal reliability		Composite reliability	
			Population RMSE ($\hat{\rho}_{max} - \rho_{max}$)	Sample RMSE ($r_{max} - \rho_{max}$)	Population RMSE ($\hat{\rho}_{CR} - \rho_{CR}$)	Sample RMSE ($r_{CR} - \rho_{CR}$)
25	3	0.6	37.4%	28.7%	18.8%	19.0%
50	3	0.6	26.3%	21.0%	13.9%	13.4%
75	3	0.6	19.4%	16.6%	11.6%	10.8%
100	3	0.6	15.0%	13.4%	10.0%	9.6%
150	3	0.6	10.5%	9.9%	8.4%	7.6%
300	3	0.6	6.4%	6.2%	5.9%	5.4%
500	3	0.6	4.8%	4.6%	4.6%	4.3%
1000	3	0.6	3.3%	3.0%	3.2%	3.0%
25	5	0.6	16.5%	22.2%	11.7%	12.9%
50	5	0.6	8.9%	12.5%	8.4%	8.9%
75	5	0.6	6.9%	8.9%	6.8%	7.2%
100	5	0.6	5.8%	7.2%	5.7%	6.2%
150	5	0.6	4.6%	5.6%	4.6%	5.0%
300	5	0.6	3.3%	3.7%	3.3%	3.5%
500	5	0.6	2.4%	2.8%	2.4%	2.7%
1000	5	0.6	1.8%	1.9%	1.8%	1.9%
25	7	0.6	9.5%	17.2%	9.2%	10.0%
50	7	0.6	5.6%	8.7%	5.9%	6.6%
75	7	0.6	4.5%	6.3%	4.6%	5.4%
100	7	0.6	3.9%	5.3%	4.0%	4.6%
150	7	0.6	3.1%	4.0%	3.3%	3.8%
300	7	0.6	2.3%	2.8%	2.3%	2.6%
500	7	0.6	1.8%	2.1%	1.8%	2.0%
1000	7	0.6	1.3%	1.5%	1.3%	1.4%
25	9	0.6	6.5%	13.3%	7.1%	8.0%
50	9	0.6	4.2%	6.7%	4.6%	5.4%
75	9	0.6	3.5%	4.9%	3.6%	4.2%
100	9	0.6	3.0%	4.2%	3.0%	3.7%
150	9	0.6	2.4%	3.2%	2.5%	3.0%
300	9	0.6	1.7%	2.2%	1.8%	2.2%
500	9	0.6	1.3%	1.7%	1.3%	1.7%
1000	9	0.6	1.0%	1.2%	1.0%	1.2%
25	3	0.7	19.7%	19.9%	12.0%	12.7%
50	3	0.7	12.5%	12.9%	8.6%	8.8%
75	3	0.7	8.9%	9.6%	7.1%	7.0%
100	3	0.7	7.1%	7.7%	6.1%	6.2%

150	3	0.7	5.4%	5.8%	5.0%	4.8%
300	3	0.7	3.6%	3.8%	3.5%	3.5%
500	3	0.7	2.7%	2.8%	2.7%	2.7%
1000	3	0.7	1.9%	1.9%	1.9%	1.9%
25	5	0.7	8.0%	13.2%	7.5%	8.3%
50	5	0.7	5.0%	7.1%	5.0%	5.7%
75	5	0.7	4.0%	5.2%	4.0%	4.5%
100	5	0.7	3.4%	4.2%	3.4%	3.9%
150	5	0.7	2.8%	3.4%	2.8%	3.1%
300	5	0.7	1.9%	2.3%	1.9%	2.2%
500	5	0.7	1.4%	1.7%	1.4%	1.7%
1000	5	0.7	1.1%	1.2%	1.1%	1.2%
25	7	0.7	4.9%	9.4%	5.4%	6.3%
50	7	0.7	3.2%	4.9%	3.4%	4.1%
75	7	0.7	2.6%	3.7%	2.8%	3.3%
100	7	0.7	2.3%	3.1%	2.3%	2.9%
150	7	0.7	1.8%	2.4%	2.0%	2.3%
300	7	0.7	1.3%	1.7%	1.4%	1.6%
500	7	0.7	1.0%	1.3%	1.0%	1.3%
1000	7	0.7	0.7%	0.9%	0.7%	0.9%
25	9	0.7	3.6%	7.1%	4.1%	5.0%
50	9	0.7	2.5%	3.8%	2.7%	3.3%
75	9	0.7	2.0%	2.9%	2.1%	2.6%
100	9	0.7	1.8%	2.5%	1.8%	2.2%
150	9	0.7	1.5%	1.9%	1.5%	1.8%
300	9	0.7	1.0%	1.3%	1.0%	1.3%
500	9	0.7	0.8%	1.0%	0.8%	1.0%
1000	9	0.7	0.6%	0.7%	0.6%	0.7%
25	3	0.8	9.4%	11.8%	7.0%	7.6%
50	3	0.8	5.7%	6.9%	4.9%	5.1%
75	3	0.8	4.3%	5.1%	3.9%	4.0%
100	3	0.8	3.4%	4.2%	3.3%	3.6%
150	3	0.8	2.7%	3.1%	2.7%	2.9%
300	3	0.8	1.9%	2.1%	1.9%	2.0%
500	3	0.8	1.4%	1.7%	1.4%	1.5%
1000	3	0.8	1.1%	1.1%	1.1%	1.1%
25	5	0.8	4.0%	6.8%	4.2%	4.9%
50	5	0.8	2.7%	3.8%	2.7%	3.2%
75	5	0.8	2.1%	2.9%	2.1%	2.6%
100	5	0.8	1.8%	2.3%	1.9%	2.2%
150	5	0.8	1.4%	1.9%	1.4%	1.8%
300	5	0.8	1.0%	1.2%	1.0%	1.2%
500	5	0.8	0.8%	1.0%	0.8%	1.0%
1000	5	0.8	0.6%	0.7%	0.6%	0.7%
25	7	0.8	2.6%	4.8%	2.9%	3.6%
50	7	0.8	1.7%	2.6%	1.8%	2.3%

75	7	0.8	1.4%	2.1%	1.5%	1.8%
100	7	0.8	1.2%	1.7%	1.3%	1.6%
150	7	0.8	1.0%	1.3%	1.1%	1.3%
300	7	0.8	0.8%	1.0%	0.8%	0.9%
500	7	0.8	0.5%	0.8%	0.5%	0.6%
1000	7	0.8	0.4%	0.5%	0.4%	0.5%
25	9	0.8	1.9%	3.6%	2.2%	2.9%
50	9	0.8	1.4%	2.0%	1.4%	1.8%
75	9	0.8	1.1%	1.6%	1.2%	1.5%
100	9	0.8	1.0%	1.4%	1.0%	1.3%
150	9	0.8	0.7%	1.1%	0.7%	1.0%
300	9	0.8	0.5%	0.7%	0.5%	0.7%
500	9	0.8	0.4%	0.5%	0.4%	0.5%
1000	9	0.8	0.3%	0.4%	0.3%	0.4%

Note: Root mean square error (RMSE) values shown are calculated as square root of average squared difference between estimate and population value / population value.

Supplemental Material, “A Note on the Finite Sample Behavior of Maximal Reliability”

Appendix F – List of Reviewed Articles

Authors	Publication	Year	Sample Size
Neff, Kristin D.; Whittaker, T.; Karl, Anke	Journal of Personality Assessment	Forth.	215
Derby, Dustin C.; Weinert, Daniel J.	Organizational Justice	2016	500
Ruzgar, Nursel Selver; Kocak, Akin; Ruzgar, Bahadtin	WSEAS Transactions on Business and Economics	2015	720
Reid, M. Jamila; Webster-Stratton, Carolyn; Baydar, Nazli	Journal of Clinical Child and Adolescent Psychology	2004	882
Vautier, Stephane	Journal of Personality Assessment	2004	888
Vautier, Stephane; Callahan, Stacey; Moncany, Delphine; Sztulman, Henri	Structural Equation Modeling	2004	1017
Buehner, Markus; Krumm, Stefan; Pick, Marion	Intelligence	2005	124
Yang, Sung-Un; Grunig, James E.	Journal of Communication Management	2005	317
Brunner, Martin; SÜβ, Heinz-Martin	Educational and Psychological Measurement	2005	1233
Buehner, Markus; Krumm, Stefan; Ziegler, Matthias; Pluecken, Tonja	Journal of Individual Differences	2006	121
Bühner, Markus; König, Cornelius J.; Pick, Marion; Krumm, Stefan	Human Performance	2006	135
Hofferth, Sandra L.	Sociological Methodology	2006	605
Jensen, G. L.; Vellas, Bruno; Garry, Phillip; Charney, Pam; others	The Journal of Nutrition, Health & Aging	2006	1324
Miller, Carla K.; Gutschall, Melissa Davis; Lawrence, Frank	Public Health Nutrition	2007	108
Loyens, Sofie MM; Rikers, Remy MJP; Schmidt, Henk G.	Studies in Higher Education	2007	180
Loyens, Sofie MM; Rikers, Remy MJP; Schmidt, Henk G.	European Journal of Psychology of Education	2007	209
Yang, Sung-Un	Journal of Public Relations Research	2007	300
Lew, Magdeleine DN; Schmidt, Henk G.	Proceedings Ascilite Singapore	2007	327
Gignac, Gilles E.; Palmer, Benjamin R.; Stough, Con	Journal of Personality Assessment	2007	363
Gignac, Gilles E.; Bates, Timothy C.; Jang, Kerry L.	Personality and Individual Differences	2007	539
Whiteside-Mansell, Leanne; Ayoub, Catherine; McKelvey, Lorraine;	Parenting: Science and Practice	2007	1122

Faldowski, Richard A.; Hart, Andrea; Shears, Jeffery			
Loyens, Sofie MM; Rikers, Remy MJP; Schmidt, Henk G.	Instructional Science	2008	98
Miller, Carla K.; Gutschall, Melissa	Health Education & Behavior	2008	103
Yun, Seong-Hun	Journal of Public Relations Research	2008	113
Shears, Jeffrey K.; Whiteside-Mansell, Leanne; McKelvey, Lorraine; Selig, James	Social Work Research	2008	315
Ziegler, Matthias; Buehner, Markus	Educational and Psychological Measurement	2008	341
Myers, Nicholas D.; Feltz, Deborah L.; Chase, Melissa A.; Reckase, Mark D.; Hancock, Gregory R.	Educational and Psychological Measurement	2008	799
Derby, Dustin C.; Smith, Thomas J.	Measurement and Evaluation in Counseling and Development	2008	1680
Owen, Steven V.; Toepperwein, Mary Anne; Marshall, Carolyn E.; Lichtenstein, Michael J.; Blalock, Cheryl L.; Liu, Yan; Pruski, Linda A.; Grimes, Kandi	Science Education	2008	1754
Thompson, Amanda L.; Mendez, Michelle A.; Borja, Judith B.; Adair, Linda S.; Zimmer, Catherine R.; Bentley, Margaret E.	Appetite	2009	150
Levy, Susan S.; Readdy, R. Tucker	Measurement in Physical Education and Exercise Science	2009	151
Rotgans, Jerome; Schmidt, Henk	Educational Studies	2009	155
Mitchell, Mary M.; Knowlton, Amy	AIDS Patient Care and STDs	2009	207
Loyens, Sofie MM; Rikers, Remy MJP; Schmidt, Henk G.	British Journal of Educational Psychology	2009	212
Silvia, Paul J.; Martin, Christopher; Nusbaum, Emily C.	Thinking Skills and Creativity	2009	226
Hofer, Adriana Rossiter; Knemeyer, A. Michael; Dresner, Martin E.	Journal of Business Logistics	2009	265
Rossiter Hofer, Adriana; Knemeyer, A. Michael	The International Journal of Logistics Management	2009	265
Cadiz, David; Sawyer, John E.; Griffith, Terri L.	Educational and Psychological Measurement	2009	583
Chen, Lung Hung; Wu, Chia-Huei; Kee, Ying Hwa; Lin, Meng-Shyan; Shui, Shang-Hsueh	Contemporary Educational Psychology	2009	691
Schaefer, Victoria A.; Meece, Judith L.	National Research Centre on Rural Education Support, University of North Carolina, Chapel Hill. Paper presented in the Annual Meeting of	2009	2095

	the American Educational Research Association		
Greene, Jeffrey A.; Brown, Scott C.	The International Journal of Aging and Human Development	2009	2715
Falbo, Toni; Kim, Sunghun; Chen, Kuan-yi	Developmental psychology	2009	3968
Silvia, Paul J.; Sanders, Camilla E.	Learning and Individual Differences	2010	129
Péloquin, Katherine; Lafontaine, Marie-France	Journal of Personality Assessment	2010	133
Schönberger, Michael; Ponsford, Jennie	Psychiatry Research	2010	140
Beaujean, A. Alexander; Firmin, Michael W.; Michonski, Jared D.; Berry, Theodore; Johnson, Courtney	Assessment	2010	142
Greene, Jeffrey Alan; Costa, Lara-Jeane; Robertson, Jane; Pan, Yi; Deekens, Victor M.	Computers & Education	2010	170
Phillips, Ann G.; Silvia, Paul J.	Personality and Individual Differences	2010	245
Cates, Dewaynna A.; Mathis, Christopher J.; Randle, Natasha W.	Journal of Managerial Issues	2010	268
Aragon, Stephen J.; McGuinn, Laura; Bavin, Stefoni A.; Gesell, Sabina B.	Journal for Healthcare Quality	2010	311
Weerts, David J.; Cabrera, Alberto F.; Sanford, Thomas	Research in Higher Education	2010	541
Moreno, Juan A.; González-Cutre, David; Sicilia, Álvaro; Spray, Christopher M.	Psychology of Sport and Exercise	2010	727
Greene, Jeffrey Alan; Torney-Purta, Judith; Azevedo, Roger	Journal of Educational Psychology	2010	740
Rotgans, Jerome I.; Schmidt, Henk G.	The Asia-Pacific Education Researcher	2010	1166
Martin, Frank H.	Rehabilitation Counseling Bulletin	2010	1200
Rotgans, Jerome I.; Schmidt, Henk G.	Learning and Instruction	2011	69
Molla, Alemayehu; Cooper, Vanessa; Pittayachawan, Siddhi	Communications of the AIS	2011	143
Jackson, Dennis L.; Singleton-Jackson, Jill A.; Frey, Marc P.	American International Journal of Contemporary Research	2011	159
Baker, Susan D.; Mathis, Christopher J.; Stites-Doe, Susan	Journal of Managerial Issues	2011	200
Rotgans, Jerome I.; Schmidt, Henk G.	Advances in Health Sciences Education	2011	208
Slater, Stanley F.; Olson, Eric M.; Finnegan, Carol	Marketing letters	2011	217
Kamp, Rachelle JA; Dolmans, Diana HJM; Van Berkel, Henk JM;	Medical Teacher	2011	223

Schmidt, Henk G.			
Silvia, Paul J.	Thinking Skills and Creativity	2011	226
Ni, Lan; Wang, Qi	Journal of Public Relations Research	2011	246
Putwain, Dave; Symes, Wendy	Learning and Individual Differences	2011	273
Coote, Leonard V.	Journal of Business Research	2011	275
Danay, Erik; Ziegler, Matthias	Journal of Research in Personality	2011	406
Rotgans, Jerome I.; Schmidt, Henk G.	Teaching and Teacher Education	2011	498
Tan, Ai-Girl; Li, Joe; Rotgans, Jerome	The Open Education Journal	2011	545
Shen, Hongmei	Journalism & Mass Communication Quarterly	2011	583
Myers, Nicholas D.; Beauchamp, Mark R.; Chase, Melissa A.	Journal of Sports Sciences	2011	748
Myers, Nicholas; Feltz, Deborah; Chase, Melissa	Research Quarterly for Exercise and Sport	2011	799
Brown, Allison R.; Finney, Sara J.	International Journal of Testing	2011	896
Gurt, Jochen; Schwennen, Christian; Elke, Gabriele	Work & Stress	2011	1027
Winterstein, Beate P.; Silvia, Paul J.; Kwapil, Thomas R.; Kaufman, James C.; Reiter-Palmon, Roni; Wigert, Benjamin	Personality and Individual Differences	2011	1144
Silvia, Paul J.; Kaufman, James C.; Reiter-Palmon, Roni; Wigert, Benjamin	Personality and Individual Differences	2011	1304
Sockalingam, Nachamma; Rotgans, Jerome I.; Schmidt, Henk G.	Advances in Health Sciences Education	2011	5949
Shen, Hongmei; Kim, Jeong-Nam	Public Relations Journal	2012	120
Silvia, Paul J.; Beaty, Roger E.	Intelligence	2012	132
Beaty, Roger E.; Silvia, Paul J.	Psychology of Aesthetics, Creativity, and the Arts	2012	133
Karantzas, Gery C.; Davison, Tanya E.; McCabe, Marita P.; Mellor, David; Beaton, Paul	Journal of Affective Disorders	2012	149
Mitchell, Mary M.; Knowlton, Amy	AIDS and Behavior	2012	215
Mu, Guanglun Michael	AERA Annual Meeting	2012	230
Beierlein, Constanze; Davidov, Eldad; Schmidt, Peter	Survey Research Methods	2012	325
Wang, Qi; Fink, Edward L.; Cai, Deborah A.	Human Communication Research	2012	352
Jiang, Hua	Public Relations Review	2012	396
Appaneal, Renee N.	Research Quarterly for Exercise and Sport	2012	433
Jackson, Ben; Myers, Nicholas D.; Taylor, Ian M.; Beauchamp, Mark R.	Journal of Sport and Exercise Psychology	2012	516
Sockalingam, Nachamma; Rotgans,	International Journal of Teaching and	2012	517

Jerome; Schmidt, Henk	Learning in Higher Education		
Trépanier, Sarah-Geneviève; Fernet, Claude; Austin, Stéphanie	Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement	2012	568
Shen, Hongmei; Kim, Jeong-Nam	Journal of Public Relations Research	2012	608
Gross, Georgina M.; Silvia, Paul J.; Barrantes-Vidal, Neus; Kwapil, Thomas R.	Schizophrenia Research	2012	780
Fernet, Claude; Guay, Frédéric; Senécal, Caroline; Austin, Stéphanie	Teaching and Teacher Education	2012	806
Zubrick, S. R.; Lawrence, D.; Mitrou, F.; Christensen, D.; Taylor, C. L.	Psychological Medicine	2012	1064
Rotgans, Jerome I.; Schmidt, Henk G.	International Journal of Teaching and Learning in Higher Education	2012	1166
Silvia, Paul J.; Wigert, Benjamin; Reiter-Palmon, Roni; Kaufman, James C.	Psychology of Aesthetics, Creativity, and the Arts	2012	1304
Fall, Anna-Mária; Roberts, Greg	Journal of Adolescence	2012	14781
Cahill, Helen; Coffey, Julia; Lester, Leanne; Midford, Richard; Ramsden, Robyn; Venning, Lynne	Health Education Journal	2013	75
Silvia, Paul J.; Beaty, Roger E.; Nusbaum, Emily C.	Intelligence	2013	131
Luciano, Juan V.; Aguado, Jaume; Serrano-Blanco, Antoni; Calandre, Elena P.; Rodriguez-Lopez, Carmen M.	Arthritis Care & Research	2013	179
Wang, Jennifer M.; Rubin, Kenneth H.; Laursen, Brett; Booth-LaForce, Cathryn; Rose-Krasnor, Linda	Journal of Clinical Child & Adolescent Psychology	2013	204
Schleimer, Stephanie C.; Pedersen, Torben	Journal of Management Studies	2013	213
Austin, Stéphanie; Guay, Frédéric; Senécal, Caroline; Fernet, Claude; Nouwen, Arie	Journal of Psychosomatic Research	2013	237
Jung, Jae Yup	Research in Higher Education	2013	349
Ziegler, Matthias; Kemper, Christoph; Rammstedt, Beatrice	Journal of Individual Differences	2013	527
Gralewski, Jacek; Karwowski, Maciej	The Journal of Creative Behavior	2013	589
Arias, José Luis; Alonso, José Ignacio; Yuste, Juan Luis	Universitas Psychologica	2013	667
Hopkins, Katrina D.; Taylor, Catherine L.; Zubrick, Stephen R.	American Journal of Orthopsychiatry	2013	1073
Trépanier, Sarah-Geneviève; Fernet, Claude; Austin, Stéphanie	Work & Stress	2013	1179
Willoughby, Michael T.; Pek, Jolynn;	Psychological Assessment	2013	1292

Blair, Clancy B.			
Dunn, Karee E.; Airola, Denise T.; Lo, Wen-Juo; Garrison, Mickey	Contemporary Educational Psychology	2013	1728
Yarnell, Lisa M.; Sargeant, Marsha N.; Prescott, Carol A.; Tilley, Jacqueline L.; Farver, Jo Ann M.; Mednick, Sarnoff A.; Venables, Peter H.; Raine, Adrian; Luczak, Susan E.	Assessment	2013	1795
Mitchell, Mary M.; Bradshaw, Catherine P.	Journal of School Psychology	2013	1902
Shi, Qi; Steen, Sam; Weiss, Brandi A.	The Family Journal	2013	2719
Allen, Joanne; Inder, Kerry J.; Lewin, Terry J.; Attia, John R.; Kelly, Brian J.	Health and Quality of Life Outcomes	2013	2740
Rotgans, Jerome I.; Schmidt, Henk G.	Learning and Instruction	2014	32
Baggetta, Peter; Alexander, Patricia A.	Sport, Exercise, and Performance Psychology	2014	58
Wang, Qi; Ni, Lan; De la Flor, Maria	Journal of Public Relations Research	2014	90
Silvia, Paul J.; Beaty, Roger E.; Nusbaum, Emily C.; Eddington, Kari M.; Kwapil, Thomas R.	Biological Psychology	2014	111
Beaty, Roger E.; Silvia, Paul J.; Nusbaum, Emily C.; Jauk, Emanuel; Benedek, Mathias	Memory & Cognition	2014	147
Fernet, Claude; Lavigne, Geneviève L.; Vallerand, Robert J.; Austin, Stéphanie	Work & Stress	2014	175
Dinsmore, Daniel L.; Baggetta, Peter; Doyle, Stephanie; Loughlin, Sandra M.	The Journal of Experimental Education	2014	178
Yi-Cheon Yim, Mark; L. Sauer, Paul; Williams, Jerome; Lee, Se-Jin; Macrury, Iain	International Marketing Review	2014	179
Pappu, Ravi; Cornwell, T. Bettina	Journal of the Academy of Marketing Science	2014	195
Wang, Jennifer M.	Child Psychiatry & Human Development	2014	200
Dumas, Denis; Dunbar, Kevin N.	Thinking Skills and Creativity	2014	201
Gunnell, Katie E.; Crocker, Peter RE; Mack, Diane E.; Wilson, Philip M.; Zumbo, Bruno D.	Psychology of Sport and Exercise	2014	203
Schleimer, Stephanie C.; Coote, Leonard V.; Riege, Andreas	Journal of Business Research	2014	213
Sánchez, R. Arteaga; Cortijo, Virginia; Javed, Uzma	Computers & Education	2014	214
Jacobs, Kate E.; Roodenburg, John	Intelligence	2014	222

Mu, Guanglun Michael	Journal of Multilingual and Multicultural Development	2014	230
Mu, Guanglun Michael	Language and Education	2014	230
Mu, Guanglun Michael	The Australian Educational Researcher	2014	230
Perera, Harsha N.; McIlveen, Peter	Journal of Vocational Behavior	2014	236
Rossiter Hofer, Adriana; J. Smith, Ronn; R. Murphy, Paul	The International Journal of Logistics Management	2014	281
Chen, Lung Hung; Chang, Yen-Ping	Current Psychology	2014	293
Ferdous, Ahmed Shahriar; Polonsky, Michael	Journal of Strategic Marketing	2014	295
De Laet, Steven; Colpin, Hilde; Goossens, Luc; Van Leeuwen, Karla; Verschueren, Karine	Journal of Psychoeducational Assessment	2014	297
Capezio, Alessandra; Cui, Lin; Hu, Helen Wei; Shields, John	Asia Pacific Journal of Management	2014	300
Sayegh, Philip; Knight, Bob G.	The Gerontologist	2014	311
Sanders, Matthew R.; Morawska, Alina; Haslam, Divna M.; Filus, Ania; Fletcher, Renee	Child Psychiatry & Human Development	2014	347
Lee, Christine S.; Huggins, Anne Corinne; Therriault, David J.	Psychology of Aesthetics, Creativity, and the Arts	2014	413
Muhamad, Haslina; Roodenburg, John; Moore, Dennis	Association for Transpersonal Psychology	2014	437
McNally, Brenton; Bradley, Graham L.	Accident Analysis & Prevention	2014	586
Hodis, Flaviu A.	Journal of Psychoeducational Assessment	2014	605
Litmanen, Topi; Loyens, Sofie MM; Sjöblom, Kirsi; Lonka, Kirsti; others	Creative Education	2014	610
Trépanier, Sarah-Geneviève; Fernet, Claude; Austin, Stéphanie; Forest, Jacques; Vallerand, Robert J.	Motivation and Emotion	2014	745
Fan, Di; Cui, Lin; Zhang, Mike Mingqiong; Zhu, Cherrie Jiuhua; Härtel, Charmine EJ; Nyland, Chris	The International Journal of Human Resource Management	2014	1488
Préville, Michel; Lamoureux-Lamarche, Catherine; Vasiliadis, Helen-Maria; Grenier, Sébastien; Potvin, Olivier; Quesnel, Louise; Gontijo-Guerra, Samantha; Mechakra-Tahiri, Samia Djemaa; Berbiche, Djamal	The Canadian Journal of Psychiatry	2014	1765
Préville, Michel; Mechakra-Tahiri, Samia Djemaa; Vasiliadis, Helen-Maria; Mathieu, Véronique; Quesnel,	The Canadian Journal of Psychiatry	2014	1765

Louise; Gontijo-Guerra, Samantha; Lamoureux-Lamarche, Catherine; Berbiche, Djamal			
Quesnel, Louise	Canadian Journal of Psychiatry	2014	1765
Gnambs, Timo; Hanfstingl, Barbara	European Journal of Psychological Assessment	2014	2138
Rotgans, Jerome I.; Schmidt, Henk G.	Higher Education Studies	2014	4068
Derby, Dustin C.; Smith, Thomas J.	Measurement and Evaluation in Counseling and Development	2014	10904
Rotgans, Jerome I.	Health Professions Education	2015	65
Chng, Esther; Yew, Elaine HJ; Schmidt, Henk G.	Advances in Health Sciences Education	2015	77
Mejia, Anilena; Filus, Ania; Calam, Rachel; Morawska, Alina; Sanders, Matthew R.	Child Psychiatry & Human Development	2015	174
Mack, Julian; Herrberg, Marlene; Hetzel, Andreas; Wallesch, Claus Werner; Bengel, Jürgen; Schulz, Moritz; Rohde, Nathalie; Schönberger, Michael	Neuropsychological Rehabilitation	2015	196
Can, Gürhan	Psychological Reports	2015	237
Jankowska, Dorota M.; Karwowski, Maciej	Frontiers in Psychology	2015	261
Ni, Lan; Wang, Qi; De la Flor, Maria	Journal of Communication Management	2015	268
Hofer, Adriana Rossiter	Transportation Journal	2015	281
Haslam, Divna; Filus, Ania; Morawska, Alina; Sanders, Matthew R.; Fletcher, Renee	Child Psychiatry & Human Development	2015	305
Vecchione, Michele; Döring, Anna K.; Alessandri, Guido; Marsicano, Gilda; Bardi, Anat	Social Development	2015	310
Silva, Desiree; Houghton, Stephen; Hagemann, Erika; Jacoby, Peter; Jongeling, Brad; Bower, Carol	Community Mental Health Journal	2015	358
Prilleltensky, Isaac; Dietz, Samantha; Prilleltensky, Ora; Myers, Nicholas D.; Rubenstein, Carolyn L.; Jin, Ying; McMahon, Adam	Journal of Community Psychology	2015	426
Hasenfratz, Liat; Benish-Weisman, Maya; Steinberg, Tami; Knafo-Noam, Ariel	Development and Psychopathology	2015	431
Trépanier, Sarah-Geneviève; Fernet, Claude; Austin, Stéphanie	Journal of Occupational Health Psychology	2015	508
Trépanier, Sarah-Geneviève; Forest, Jacques; Fernet, Claude; Austin,	Work & Stress	2015	669

Stéphanie			
Trépanier, Sarah-Geneviève; Fernet, Claude; Austin, Stéphanie; Ménard, Julie	Burnout Research	2015	1159
Daniel, Ella; Dys, Sebastian P.; Buchmann, Marlis; Malti, Tina	Social Development	2015	1258
Sperry, Sarah H.; Walsh, Molly A.; Kwapil, Thomas R.	Personality and Individual Differences	2015	2713
Salari, Raziye; Filus, Ania	Prevention Science	2016	122
Gralewski, Jacek; Karwowski, Maciej	The Journal of Creative Behavior	2016	131
Dinsmore, Daniel L.; Alexander, Patricia A.	The Journal of Experimental Education	2016	151
Smith, Kelly A.; Barstead, Matthew G.; Rubin, Kenneth H.	Journal of Youth and Adolescence	2016	181
Lampard, Amy M.; Nishi, Akihiro; Baskin, Monica L.; Carson, Tiffany L.; Davison, Kirsten K.	Behavioral Medicine	2016	195
Neff, Kristin D.	Mindfulness	2016	215
Sharma, Umesh; Jacobs, Dt Kate	Teaching and Teacher Education	2016	245
Dittman, Cassandra K.; Burke, Kylie; Filus, Ania; Haslam, Divna; Ralph, Alan	Journal of Adolescence	2016	279
Lundgren, Berndt Allan; Lundgren, Berndt Allan; Wallentin, Fan Yang; Wallentin, Fan Yang	Journal of European Real Estate Research	2016	282
Draaijer, Silvester; Schoonenboom, Judith; Beishuizen, Jos; Schuwirth, Lambert	Thinking Skills and Creativity	2016	470
Zhang, Qinghua; Huang, Feifei; Liu, Zaoling; Zhang, Na; Mahapatra, Tanmay; Tang, Weiming; Lei, Yang; Dai, Yali; Tang, Songyuan; Zhang, Jingping	PloS One	2016	550
Way, Kirsten A.; Jimmieson, Nerina L.; Bordia, Prashant	International Journal of Conflict Management	2016	562
Shen, Hongmei	Journalism & Mass Communication Quarterly	2016	583
Fernet, Claude; Trépanier, Sarah-Geneviève; Austin, Stéphanie; Levesque-Côté, Julie	Teaching and Teacher Education	2016	586
Gnambs, Timo; Hanfstingl, Barbara	Educational Psychology	2016	600
Guo, Mingchun; Morawska, Alina; Filus, Ania	Assessment	2016	650
Gunnell, Katie E.; Bélanger, Mathieu; Brunet, Jennifer	Health Psychology	2016	842
Dwivedi, Abhishek; Merrilees, Bill	Australasian Marketing Journal	2016	966

Benedek, Mathias; Nordtvedt, Nora; Jauk, Emanuel; Koschmieder, Corinna; Pretsch, Jürgen; Krammer, Georg; Neubauer, Aljoscha C.	Thinking Skills and Creativity	2016	1118
Willoughby, Michael T.; Blair, Clancy B.	Psychological Assessment	2016	1292
Jimmieson, Nerina L.; Tucker, Michelle K.; White, Katherine M.; Liao, Jenny; Campbell, Megan; Brain, David; Page, Katie; Barnett, Adrian G.; Graves, Nicholas	Safety Science	2016	2335
Askell-Williams, Helen; Skrzypiec, Grace; Jin, Yan; Owens, Larry; Zhao, Xueqin; Du, Wenping; Cao, Fei; Xing, Lihong	International Journal of Emotional Education	2016	2756
Roos, J. Micah	Social Currents	2016	4861
Prati, Gabriele; Zani, Bruna; Pietrantoni, Luca; Scudiero, Diego; Perone, Patrizia; Cosmaro, Lella; Cerioli, Alessandra; Oldrini, Massimo	Quality & Quantity	2016	6675
Datu, Jesus Alfonso D.; King, Ronnel B.; Valdez, Jana Patricia M.	Journal of School Psychology	2017	400
Hodis, Flaviu A.; Hattie, John AC; Hodis, Georgeta M.	Personality and Individual Differences	2017	605