Assessing Fit in Ordinal Factor Analysis Models: SRMR vs. RMSEA

Dexin Shi

University of South Carolina

Alberto Maydeu-Olivares

University of South Carolina

University of Barcelona

Yves Rosseel

Ghent University

Structural Equation Modeling (2019)

Author Note

Correspondence concerning this article should be addressed to Dexin Shi, Dept. of Psychology, University of South Carolina. Barnwell College. 1512 Pendleton St. Columbia, SC 29208. E-mail: shid@mailbox.sc.edu. The authors acknowledge the Research Computing Center at the University of South Carolina and the High-Performance Computing Center at Weifang University for providing the computing resources used to obtain the results reported in this paper. This research was supported by the National Science Foundation under Grant No. SES-1659936.

Abstract

This study introduces the statistical theory of using the Standardized Root Mean Squared Error (SRMR) to test close fit in ordinal factor analysis. We also compare the accuracy of confidence intervals (CIs) and tests of close fit based on the Standardized Root Mean Squared Error (SRMR) with those obtained based on the Root Mean Squared Error of Approximation (RMSEA). We use Unweighted Least Squares (ULS) estimation with a mean and variance corrected test statistic. The current (biased) implementation for the RMSEA never rejects that a model fits closely when data are binary and almost invariably rejects the model in large samples if data consist of five categories. The unbiased RMSEA produces better rejection rates, but it is only accurate enough when the number of variables is small (e.g., $p = 10$) and the degree of misfit is small. In contrast, across all simulated conditions, the tests of close fit based on the SRMR yield acceptable type I error rates. SRMR tests of close fit are also more powerful than those using the unbiased RMSEA.

Keywords: Ordinal Factor analysis; SRMR; RMSEA; Close fit

Assessing Close Fit in Ordinal Factor Analysis Models: SRMR vs. RMSEA

When fitting a structural equations model (SEM), most often the model is to some degree incorrect or misspecified (Box, 1979; MacCallum, 2003). Effect sizes of model misfit are population parameters that quantify the extent of a model misfit (Maydeu-Olivares, 2017). When a model fits poorly, incorrect substantive inferences are likely to be drawn (Maydeu-Olivares, 2017; Saris, Satorra, & van der Veld, 2009). Consequently, either using practical experience (Browne & Cudeck, 1993) or carefully designed population studies (Dexin Shi, Maydeu-Olivares, & DiStefano, 2018), reference (aka cutoff) values for effect sizes of model misfit have been offered to distinguish between misspecified models that are actionable (aka close fitting models) and models that should be discarded. As for any parameter, it is important to convey the accuracy with which the size of model misfit is estimated. This is accomplished by providing its confidence interval. Also, when using the reference (aka cutoff) population values of close fit that have been provided, researchers may be interested in obtaining a *p*-value that the estimated effect size of model misfit does not exceed the recommended cutoff.

The Root Mean Squared Error of Approximation (RMSEA: Steiger & Lind, 1980) is currently the most widely used effect size[1] of model misfit (Kenny, Kaniskan & McCoach, 2015; Savalei, 2012), and it is routinely reported in papers that apply SEM models (Hancock, Stapleton, & Mueller, 2018; McDonald & Ho, 2002). Asymptotic statistical theory is available for constructing confidence intervals for the RMSEA (Browne & Cudeck, 1993; Steiger, 1989, 1990), as well as tests of close fit of the type $H_0 : RMSEA \le c$, where c is a cutoff population value. In practice, the most commonly used cut-off values for the RMSEA is based on the recommendations put forth by Browne and Cudeck (1993, p. 144), who stated that “practical experience had made us feel that a value of the RMSEA of about 0.05 or less would indicate a close fit of the model in relation to the degrees of freedom”. However, the RMSEA is an unstandardized effect size of model misfit (Maydeu-Olivares, 2017). As a result, any population value (say 0.05) has a different meaning depending on the structure and size of the model (Chen, Curran, Bollen, Kirby, & Paxton, 2008; Savalei, 2012; Shi, Lee, & Maydeu-Olivares, 2018). For instance, suppose the population model has two factors with an interfactor correlation of 0.95, most researchers would consider a one-factor model to be a close-fitting model. However, because the population RMSEA is an unstandardized parameter, it may exceed the recommended reference value of .05 or not, depending on the number of indicators per factor, holding any other determinant constant (Maydeu-Olivares, 2017).

To overcome problems associated with interpreting the magnitude of the unstandardized effect sizes of misfit, Maydeu-Olivares (2017) advocated the use of standardized effect sizes of model misfit such as the Standardized Root Mean Squared Error (SRMR). Unlike the RMSEA, the population SRMR is easy to interpret: it can be crudely interpreted as the average standardized residual covariance. Maydeu-Olivares (2017) also derived the unbiased estimate of the population SRMR and provided the asymptotic statistical theory to obtain confidence intervals as well as to perform close fit tests. In this respect, the SRMR is comparable to the RMSEA.

Previous studies have shed some light in understanding the performance of the RMSEA and SRMR in assessing model misfit. A number of simulation studies has examined the performance of point estimates, confidence intervals and tests of close fit using the RMSEA

(Curran, Bollen, Chen, Paxton, & Kirby, 2003; Fan, Thompson, & Wang, 1999; Hu & Bentler, 1998; Kenny, Kaniskan, & McCoach, 2015; Kenny & McCoach, 2003; Nasser & Wisenbaker, 2003; Nevitt & Hancock, 2000). In general, under the conditions studied, the RMSEA yields relatively accurate point estimates and CIs when the sample size is large enough (e.g., $N \geq 200$). Maydeu-Olivares (2017) investigated the accuracy of point estimates, CIs and *p*-values for tests of close fit for the SRMR in small models with normal data, and showed that the performance of SRMR was adequate, except for conditions where the level of model misspecification was small (population SRMR = .025) and the sample size was small ($N \leq 100$). Recently, using a more comprehensive simulation design with both normal and non-normal data, Maydeu-Olivares, Shi and Rosseel (2018) compared the performance of RMSEA and SRMR in assessing close fit in factor analysis models. Results showed that compared to the RMSEA, the SRMR generally yields more accurate empirical rejection rates for tests of close fit and better coverage for its population value, especially when the number of observed variables (p) is large. CIs and tests of close fit based on the RMSEA were only accurate in the smallest models considered (i.e., p = 10).

Although suggestive, the findings and recommendations from previous studies are limited. This is because virtually all previous studies focused on factor analysis models for continuous outcome variables using maximum likelihood estimation. However, it would be fair to say that most data in the social sciences are ordinal (e.g., responses to survey items, educational tests, psychological questionnaires) and therefore, in principle, it should be treated as discrete. Strictly speaking, fitting a common factor analysis model to ordinal data introduces model misspecification, as the relationship between a discrete outcome and a continuous factor cannot be linear (Maydeu-Olivares, 2005; Maydeu-Olivares, Cai, & Hernández, 2011). To take into account the ordinal nature of the data, under the SEM framework, a better approach is to fit an ordinal factor analysis model, especially if the data have fewer than five ordered categories (DiStefano & Morgan, 2014; Muthén & Kaplan, 1992; Rhemtulla, Brosseau-Liard, & Savalei, 2012).

When fitting ordinal factor analysis models, sequential estimators involving tetrachoric or polychoric correlations are generally recommended. More specifically, assuming that the observed ordinal data is the result of categorizing underlying normal variates, polychoric correlations[2] can be estimated to measure the association between the underlying normal variates before categorization (see Olsson, 1979). Let $\mathbf{P}$ denotes the unknown population polychoric correlation matrix, and let $\mathbf{P}_0 = \mathbf{P}(\boldsymbol{\theta})$ be correlation structure implied by the fitted model, which depends on a vector of q parameters $\boldsymbol{\theta}$ to be estimated from the data. When an ordinal factor analysis model does not involve constraints on the threshold structure[3], it can be estimated by minimizing

$$F(\mathbf{R},\mathbf{P}_0) = \left(\mathbf{r}-\boldsymbol{\rho}_0\right)' \hat{\mathbf{W}} \left(\mathbf{r}-\boldsymbol{\rho}_0\right) \tag{1}$$

where $\mathbf{r}$ and $\boldsymbol{\rho}_0$ are the vectors of the sample and model-implied polychoric correlation matrices (i.e., elements below the diagonal of the matrix $\mathbf{R}$ and $\mathbf{P}_0$), and $\hat{\mathbf{W}} \xrightarrow{p} \mathbf{W}$, a fixed weight matrix. Different choices of the weight matrix $\hat{\mathbf{W}}$ lead to different estimators. In Unweighted Least Squares (ULS: Muthén, 1993), the weight matrix is an identity matrix. In Diagonally Weighted Least Squares (DWLS: Jöreskog & Sörbom, 1988; Muthén, du Toit, & Spisic, 1997), the weight matrix is $\hat{\mathbf{W}} = (diag(\hat{\boldsymbol{\Gamma}}))^{-1}$, where $\boldsymbol{\Gamma}$ is the asymptotic covariance matrix of the estimated polychoric correlations. Finally, in Weighted Least Squares (WLS: Muthén, 1978, 1984), the weight matrix is $\hat{\mathbf{W}} = \hat{\boldsymbol{\Gamma}}^{-1}$. WLS is seldom used as it requires very large samples (Muthén &

Kaplan, 1992). Consequently, ULS and DWLS have emerged in the literature as recommended options (Bandalos, 2014; Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009; Savalei & Rhemtulla, 2013). For WLS, the estimated minimum of the fit function (see Eq. 1) multiplied by sample size follows a chi-square distribution in large samples. This is the so-called chi-square test of model fit. However, when ULS or DWLS estimation is employed, the test of exact fit of the model is obtained by applying a mean or a mean and variance correction to the test statistic (Muthén, 1993). ULS and DWLS estimation with mean and variance corrected test statistics are denoted as ULSMV and WLSMV respectively in the popular SEM software program Mplus (Muthén & Muthén, 2017); DWLS estimation with a mean corrected statistic is denoted as WLSM, and parameter estimates and standard errors for WLSM and WLSMV are identical. Several studies have shown that ULSMV generally provides more accurate parameter estimates, standard errors, and chi-square test statistics than WLSMV (Forero et al., 2009; Savalei & Rhemtulla, 2013; Shi, DiStefano, McDaniel, & Jiang, 2018). Therefore, in this study, we focus on the ULS estimator with a mean and variance corrected chi-square statistic.

Although ordinal factor models are widely applied across various disciplines in the behavioral and social sciences, the accuracy with which close fit assessment can be performed in these models has not been systematically examined. When mean and variance-adjusted chi-square test statistics are used (as in ULS or DWLS estimation), current implementations of the RMSEA in most SEM software programs (e.g., Mplus) provides an inconsistent estimate of the parameter (see Brosseau-Liard & Savalei, 2014; Brosseau-Liard, Savalei, & Li, 2012). We will return to this point later in the paper.

The correct procedure to estimate the population RMSEA and obtain its CI with mean and variance- adjusted test statistics has been provided by Savalei (2018) in the context of models for continuous outcomes. Xia (2016) applied Savalei's formula in ordinal factor analysis models and showed that the sample estimates of RMSEA under ULSMV were generally consistent, especially when the sample size was large ($N \geq 500$). However, no results were provided in terms of accuracy of confidence intervals, nor tests of close fit. Additionally, statistical theory for obtaining confidence intervals and tests of close fit of ordinal factor analysis models using the SRMR is needed.

In the current study, we expand Maydeu-Olivares (2017)'s work and introduce statistical theory that enables using the SRMR to test close fit in ordinal factor analysis models. In addition, we aim to investigate and compare the accuracy of statistical procedures to obtain confidence intervals and tests of close fit between the SRMR and the RMSEA. To this end, we also discuss how to compute point estimates and confidence intervals (CIs) for the RMSEA in these models. Next, we report a simulation study using ULS estimation. We close our presentation by offering a numeric example and providing suggestions to applied researchers.

Statistical Theory for the SRMR

Let p denote the number of observed variables. For covariance structure models, the population Standardized Root Mean Squared Residual (SRMR) is defined as (Maydeu-Olivares, 2017)

$$SRMR = \sqrt{\frac{1}{t}\sum_{i \leq j} \frac{\left(\sigma_{ij} - \sigma_{ij}^{0}\right)^{2}}{\sqrt{\sigma_{ii}\sigma_{jj}}}}\,, \qquad (2)$$

where σ_{ij} denotes the population covariance between variables i and j (or variance if $i = j$) and σ_{ij}^{0} denotes the population covariance (or variance) under the fitted model, and $t = p(p+1)/2$. In ordinal factor analysis a correlation structure is fitted. As a result, in these models $\sigma_{ij} = \rho_{ij}$, the population polychoric correlation between variables *i* and j , $\sigma_{ij} = 1$ if *i* = *j*, and $\sigma_{ij}^0 = \rho_{ij}^0$ denote the polychoric correlation under the fitted ordinal factor analysis model. As a result, in correlation structure models such as those fitted in ordinal factor analysis, the population SRMR simplifies to

$$SRMR_p = \sqrt{\frac{\boldsymbol{\varepsilon}_r'\boldsymbol{\varepsilon}_r}{t}} = \sqrt{\frac{1}{t}\sum_{i\le j}\left(\rho_{ij} - \rho_{ij}^0\right)^2} \tag{3}$$

Here, $\boldsymbol{\varepsilon}_r$ is the $(t-p)$ dimensional vector of population residual correlations and the population SRMR (2) can be crudely interpreted as the average population residual correlation. Notice the division by *t* in Equation (3) for consistency with the definition of Equation (2).

In finite samples, let r_{ij} be the sample estimated polychoric correlation, $\hat{\rho}_{ij}$ denote the model implied polychoric correlation, $\hat{\boldsymbol{\varepsilon}}_r$ be the $(t-p)$ vector of the residual correlations with elements $r_{ij} - \hat{\rho}_{ij}$, and $\boldsymbol{\Xi}_r$ represent the asymptotic covariance matrix of $\hat{\boldsymbol{\varepsilon}}_r$. When fitting ordinal factor analysis model under ULS estimation, with $\mathbf{H} = \boldsymbol{\Delta}(\boldsymbol{\Delta}'\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}'$, $\boldsymbol{\Xi}_r$ is computed as

$$\boldsymbol{\Xi}_r = N^{-1}(\mathbf{I} - \mathbf{H})\boldsymbol{\Gamma}(\mathbf{I} - \mathbf{H})' \tag{4}$$

Here, $\boldsymbol{\Gamma}$ denotes the asymptotic covariance matrix of $\sqrt{N}\mathbf{r}$, and $\boldsymbol{\Delta}$ is the Jacobian matrix $\boldsymbol{\Delta} = \left[\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\theta}'}\right]_{\boldsymbol{\theta}_0}$.

Maydeu-Olivares (2017) showed that regardless of the discrepancy function used, an asymptotically unbiased estimate of the population SRMR of Equation 2 is

$$SRMR_u = \hat{k}^{-1}\sqrt{\frac{\max\left(\hat{\boldsymbol{\varepsilon}}_r{}'\hat{\boldsymbol{\varepsilon}}_r - \mathrm{tr}(\hat{\boldsymbol{\Xi}}_r),0\right)}{t}}, \quad \hat{k} = 1 - \frac{\mathrm{tr}(\hat{\boldsymbol{\Xi}}_r^2) + 2\hat{\boldsymbol{\varepsilon}}_r\hat{\boldsymbol{\Xi}}_r\hat{\boldsymbol{\varepsilon}}_r}{4\left(\hat{\boldsymbol{\varepsilon}}_r{}'\hat{\boldsymbol{\varepsilon}}_r\right)^2}. \tag{5}$$

Currently, for ordinal factor analysis models, most SEM software programs compute a sample counterpart of the population SRMR in (1)

$$SRMR_b = \sqrt{\frac{\hat{\boldsymbol{\varepsilon}}_r{}'\hat{\boldsymbol{\varepsilon}}_r}{t}} = \sqrt{\frac{1}{t}\sum_{i \le j}\left(r_{ij} - \hat{\rho}_{ij}\right)^2}\,. \tag{6}$$

However, in finite samples, the sample SRMR of Equation 5 is a biased estimate of the population SRMR. In large samples, its expected value under parameter drift assumptions is (Shi, Maydeu-Olivares, and DiStefano, 2018)

$$E[SRMR_b] = \sqrt{\frac{\operatorname{tr}(\Xi_r) + \boldsymbol{\varepsilon}_r'\boldsymbol{\varepsilon}_r}{t}\,\frac{8[\operatorname{tr}(\Xi_s) + \boldsymbol{\varepsilon}_r'\boldsymbol{\varepsilon}_r]^2 - 2\operatorname{tr}(\Xi_r) - 4\boldsymbol{\varepsilon}_r'\Xi_r\boldsymbol{\varepsilon}_r}{8[\operatorname{tr}(\Xi_r) + \boldsymbol{\varepsilon}_r'\boldsymbol{\varepsilon}_r]^2}}\,, \tag{7}$$

and therefore the bias of the sample SRMR*b* is $SRMR_p - E[SRMR_b]$.

Confidence intervals for the SRMR and tests of close fit can be obtained using the unbiased estimate and a reference normal distribution. Specifically, in large samples, a (100 – α)% confidence interval for the SRMR can be obtained using

$$\Pr\left(SRMR_u - z_{\alpha/2}SE(SRMR_u) \le SRMR \le SRMR_u + z_{\alpha/2}SE(SRMR_u)\right) = 1 - \alpha\,, \tag{8}$$

where SE () denotes asymptotic standard error, which is given as (Maydeu-Olivares, 2017):

$$\mathrm{SE}(SRMR_u) = \sqrt{k^{-2}\,\frac{\operatorname{tr}(\Xi_r^2) + 2\hat{\boldsymbol{\varepsilon}}_r'\Xi_r\hat{\boldsymbol{\varepsilon}}_r}{2t\,\hat{\boldsymbol{\varepsilon}}_r{}'\hat{\boldsymbol{\varepsilon}}_r}}\,. \tag{9}$$

In addition, a statistical test of the close fit of a model, of the type $H_0: SRMR \le c$ vs. $H_1: SRMR > c$, where $c > 0$ is a reference cut-off value for close fit can be conducted. *P*-values are obtained using $p = 1 - \Phi(z)$, where $\Phi()$ denotes a standard normal distribution function and $z = \dfrac{SRMR_u - c}{SE(SRMR_u)}$.

Statistical Theory for the RMSEA

The population Root Mean Squared Error of approximation (RMSEA: Browne & Cudeck, 1993; Steiger, 1989, 1990) is defined as

$$RMSEA = \sqrt{\frac{F_0}{df}} \tag{10}$$

where F_0 denotes the discrepancy between the population and model-implied covariance (or correlation) matrices. Thus, the RMSEA adjusts the discrepancy by the degrees of freedom of the model. It is noted that the definition of the RMSEA depends on the discrepancy used to estimate the model. There is not a unique population RMSEA. Rather, there are as many population RMSEA parameters as discrepancies can be used to estimate the model. When fitting an ordinal factor analysis model using WLS the population RMSEA is $RMSEA_{WLS} = \sqrt{\frac{\boldsymbol{\varepsilon}_r' \boldsymbol{\Gamma}^{-1} \boldsymbol{\varepsilon}_r}{df}}$; when using DWLS the population RMSEA is $RMSEA_{DWLS} = \sqrt{\frac{\boldsymbol{\varepsilon}_r' (diag(\boldsymbol{\Gamma}))^{-1} \boldsymbol{\varepsilon}_r}{df}}$; whereas for the ULS fitting function the population RMSEA is

$$RMSEA_{ULS} = \sqrt{\frac{\boldsymbol{\varepsilon}_r' \boldsymbol{\varepsilon}_r}{df}} = \sqrt{\frac{1}{df} \sum_{i \le j} \left(\rho_{ij} - \rho_{ij}^0 \right)^2} . \tag{11}$$

The sample RMSEA computed in many SEM programs (e.g., Mplus) is

$$RMSEA_b = \sqrt{\max\left(\frac{X^2 - df}{N \times df}, 0 \right)} . \tag{12}$$

Where X^2 denotes the chi-square test statistic of model fit. Confidence intervals are then computed using

$$\left(\sqrt{\max\left(\frac{\hat{L}}{N \times df}, 0 \right)}; \sqrt{\max\left(\frac{\hat{U}}{N \times df} \right), 0} \right) \tag{13}$$

where $\hat{L}$ and $\hat{U}$ are the solution to

$$F_{\chi^2}\left(X^2;df,\hat{L}\right)=1-\alpha\text{, and } F_{\chi^2}\left(X^2;df,\hat{U}\right)=\alpha \tag{14}$$

respectively, and $F_{\chi^2}\left(\cdot;df,\lambda\right)$ is the non-central chi-square distribution function with df degrees of freedom and non-centrality parameter λ. Finally, *p*-values for the test of close fit are computed using

$$p=1-F_{\chi^2}\left(X^2;df,N\times df\times c^2\right). \tag{15}$$

These formulae can be used to estimate the population RMSEA$_{\text{WLS}}$, but should not be used to estimate RMSEA$_{\text{DWLS}}$ or RMSEA$_{\text{ULS}}$, as it does not consistently estimate these population RMSEA parameters, they estimate different parameters (Brosseau-Liard & Savalei, 2014; Brosseau-Liard, Savalei, & Li, 2012). Since in this paper we focus on the ULS estimator with a mean and variance corrected chi-square statistic, and therefore on the estimation of the RMSEA$_{\text{ULS}}$, we refer to the formula (12) to (15) as “biased” formula (for this RMSEA).

More specifically, under ULS estimation, $N\hat{F}_{ULS}$ does not follow a chi-square distribution. To assess the exact fit of the model, $N\hat{F}_{ULS}$ is adjusted by its asymptotic mean or by its mean and variance (Muthén, 1993; Satorra & Bentler, 1994). The mean and variance adjusted chi-square (X^2_{ULSMV}) used in this paper is (Asparouhov & Muthén, 2010)

$$X^2_{ULSMV}=\frac{1}{\hat{a}_{ULS}}N\hat{F}_{ULS}+\hat{b}_{ULS} \tag{16}$$

where $\hat{a}_{ULS}$ and $\hat{b}_{ULS}$ are the scaling correction factor and the shift parameter. Recently, Savalei (2018) has provided the correct formula to estimate the population RMSEA and obtained its confidence interval when a mean and variance adjusted test statistic is used. The ULS population RMSEA (11) should be estimated using

$$RMSEA_u = \sqrt{\max\left(\frac{\hat{a}(X^2_{ULSMV} - df)}{N \times df}, 0\right)}. \tag{17}$$

Provided that X^2_{ULSMV} can be well approximated in large samples using a non-central chi-square distribution under parameter drift assumptions, a (100 – α)% confidence interval (CI) for the RMSEA is given by

$$\left(\sqrt{\max\left(\frac{\hat{a} \times \hat{L}}{N \times df}, 0\right)}; \sqrt{\max\left(\frac{\hat{a} \times \hat{U}}{N \times df}\right), 0}\right) \tag{18}$$

where $\hat{L}$ and $\hat{U}$ are the solution to

$$F_{\chi^2}\left(X^2_{ULSMV}; df, \hat{L}\right) = 1 - \alpha \text{ , and } F_{\chi^2}\left(X^2_{ULSMV}; df, \hat{U}\right) = \alpha \tag{19}$$

Respectively. Alternatively, researchers may be interested in testing the close fit of an ordinal factor analysis model using a hypothesis of the type: $H_0 : RMSEA \le c$ vs. $H_1 : RMSEA > c$, where *c* is an arbitrary value of the population RMSEA. In this case, the asymptotic *p*-value for the test of close fit is

$$p = 1 - F_{\chi^2}\left(X^2_{ULSMV}; df, N \times df \times c^2 / \hat{a}\right) \tag{20}$$

Simulation Study

We conducted a simulation study to evaluate the accuracy of the point estimates, confidence intervals (CIs), and tests of close fit using the SRMR and the RMSEA under ULS estimation of ordinal factor analysis models. We first generated multivariate normal data with mean zero and a correlation structure conforming to an independent cluster model with two correlated factors. Each factor was specified to have the same number of indicators. To create ordinal outcomes, we then discretized the continuous data into categories using thresholds. We set the population values for all factor loadings to 0.70 and all error variances to

$1-0.70^2=0.51$. For both factors, the factor variances were set one. The model misspecification was introduced by ignoring the multidimensional structure (i.e., fitting a one-factor model to two-factor data). The levels of model misspecification were manipulated by altering the population values of the inter-factor correlations. Given that the estimated model was unidimensional, a smaller inter-factor correlation indicated a greater level of misspecification. The detailed simulation conditions are described below.

Simulation Conditions

The simulation conditions were obtained by manipulating four variables:

Sample size. Sample sizes included 100, 200, 500, and 1,000.

Model size. The model size was indicated by the total number of observed variables (p; Shi, Lee, & Terry, 2015; 2018), including small ($p = 10$), medium ($p = 20$), and large ($p = 30$) models.

Number of categories. The observed data were categorized to be either binary ($k = 2$) or polytomous ($k = 5$) by altering threshold values. Specifically, when the data were binary, the thresholds used led to expected probabilities of 60% and 40%. For data with five categories, thresholds used led to expected probabilities of 10%, 20%, 40%, 20%, and 10% for ordered categories 0 through 4, respectively. The threshold values used were based on previous simulation studies in ordinal factor analysis (Forero et al., 2009).

Levels of Model Misspecification. Five levels of model misspecification were obtained by manipulating the inter-factor correlation in the population model. The population inter-factor correlations ranged from 0.50 to 0.90, in increments of .10. As a representation of the different situations of model misfit that might occur in real data analysis, we considered various levels of model misspecification. For example, when the population model is a two factor model with an interfactor correlation of 0.90 and a one-factor model is fitted to the data, most researchers would consider the model misfit to be substantively ignorable. On the other hand, when the true model is a two dimensional model with an interfactor correlation of 0.50, a unidimensional structure would fit the data poorly and thereby should be rejected.

For every possible condition, the population values of SRMR and RMSEA were computed using the known population polychoric correlation matrices. As shown in Table 1, across the conditions considered in the current study, the population SRMR values ranged from .022 to .118; whereas the population RMSEA ranged from .028 to .138. It is noted that the population RMSEA tended to be slightly larger than the population SRMR obtained from the same simulation condition. In addition, for the same level of model misspecification, as the number of observed variable (*p*) increased, the population RMSEA decreased; however, the population SRMR slightly increased as more observed variables (*p*) were added to the model.

In total, the simulation study consisted of a crossed design with 120 conditions (4 sample sizes × 3 model sizes × 2 number of categories × 5 levels of model misspecification). For each simulation condition, 1,000 replications were generated using *simsem* package in R (Pornprasertmanit, Miller, & Schoemann, 2013; R Development Core Team, 2015).

Data Analysis

A one-factor model was fitted to each simulated dataset by unweighted least squares (ULS) estimation using the *lavaan* package in R (R Development Core Team, 2015; Rosseel, 2012). Across replications, we computed the sample estimates and obtained the 90% CIs of both SRMR and RMSEA. In addition, we computed the *p* values of tests of close fit for the SRMR and RMSEA. First, we tested the hypothesis that the SRMR and RMSEA were less than or equal to its population value to examine empirical Type I errors of the tests of close fit. Second, to examine the empirical power of the tests of close fit, we tested that the SRMR and RMSEA were less than or equal to fixed cutoff values (c), where the cut-offs (c) were defined as the corresponding population SRMR or RMSEA when fitting a one-factor model to two-factor data with inter-factor correlations $\rho = .90$. That is, for conditions where $\rho < .90$, the empirical rejection rate (i.e. $p < .05$) reveals the power of the test of close fit (to reject a model that does not fit closely). For the RMSEA, the sample estimates, along with the CIs and p values for the test of close fit were computed based on the mean- and variance adjusted chi-square test statistics (i.e., ULSMV) as described in equation (15) –as provided by current software, as well as using (20), the correct formula.

Results

Table 1 summarizes the population values and the average sample estimates (based on both the biased and the unbiased formula) for the SRMR and RMSEA across all simulated conditions. The sample estimates that yielded a relative bias greater than 10% are highlighted in bold. In general, both the RMSEA and SRMR yield average estimates close to their population values, as long as the correct (unbiased) formula were used, Equation (5) for the SRMR and Equation (17) for the RMSEA, even when the sample size was small (e.g., N =100). In addition, in small samples (N =100), the population SRMR and RMSEA was more accurately estimated with five categories than with two. For example, when $p = 10$, $\rho = .80$, and $N = 100$, the average sample RMSEA was .047 for binary data and .050 for polytomous data, whereas the population value of RMSEA was .055.

On the other hand, when the biased formula was applied, Equation (6) for the SRMR and Equation (12) for the RMSEA, the average sample RMSEA and SRMR could be substantially biased from their population counterparts. To be more specific, using the biased (incorrect) formula, the population RMSEA tended to be underestimated for binary data and overestimated for polytomous data. For the SRMR, the population values were consistently overestimated, suggesting that the model fits poorer than what it does. For the same condition as the previous example ($p = 10$, $\rho = .80$, and $N = 100$), the average (biased) sample RMSEA was .037 and .066 for binary data and polytomous data, respectively. However, the population RMSEA was .055. The average (biased) sample SRMR was .096 (binary data) and .068 (polytomous data), whereas the population value for SRMR was .044. It is interesting to note that for the SRMR, the estimates based on the biased formula converged to its population value as N increased, as Equation (6) still estimates the correct population parameter asymptotically. However, this did not hold for the RMSEA. The biased formula for the RMSEA estimates a different parameter than Equation (11).

Besides the accuracy of the point estimates, we also evaluated the performance of estimating the 90% CIs around the population SRMR and RMSEA. The coverage rates for the 90% CIs are provided in Table 2[4]. Coverage rates between 85% and 95% were considered acceptable. Cases with unacceptable coverage rates were highlighted in bold in Table 2. We see in this table that the CIs for the unbiased RMSEA could perform poorly, especially when fitting large models (e.g., $p \geq 20$), or/and models with a high level of misspecification (e.g., $\rho = .70$). Under the conditions mentioned above, it was surprising to observe that the coverage rates for the RMSEA did not improve, and could even become worse as sample size (N) increased. For example, when $p = 10$ and $\rho = .60$, for five response category data, the coverage rate for the 90% CI of RMSEA dropped from .70 ($N = 100$) to .61 ($N = 1{,}000$). CIs using the biased formula performed much worse, especially when the sample was large (e.g., $N \geq 500$). For the same condition, as N increased from 100 to 1,000, the 90% CI using the biased formula dropped from

66% to 0%. Since the biased formula actually estimates a different parameter than the ULS based population RMSEA (defined in Eq. 11), it is not surprising to observe that the accuracy of the CIs of the biased RMSEA become worse as sample size increases.

On the other hand, CIs coverage rates for the SRMR were closer to the acceptable level across most of the simulation conditions. The performance of SRMR improved as the sample size increased and when the level of model misspecification increased. In addition, the SRMR CIs generally produced better coverage rates for data with five categories than those obtained from binary data. Note that the only situations where unbiased CIs of RMSEA seemed to perform better than of SRMR involved fitting models with very small levels of model misspecification (i.e., $\rho = .90$), especially in samples of $N \leq 500$.

Table 3 depicts the 5% empirical rejection rates (type I error rates) of tests that RMSEA and SRMR are less than or equal to their population values. As reported above, the population values ranged from .028 to .138 for RMSEA and from .022 to .118 for SRMR. At the significance level of 5%, we considered empirical rejection rates between .01 and .10 acceptable. Cases with unacceptable empirical rejection rates were highlighted in bold in the table. It is noted that for the unbiased RMSEA, the empirical rejection rates were reasonably accurate only when the model size was small (e.g., $p = 10$) and the level of model misspecification was small (i.e., $\rho = .90$). For larger models with large specification error, the empirical rejection rates for RMSEA were larger than the nominal level (i.e., inflated type I error rates), even in large samples. For example, when $p = 30$, $\rho = .70$, and $N = 1{,}000$, the empirical rejection rates at $\alpha = 5\%$ were 25% ($c = 2$) and 29% ($c = 5$). For the biased RMSEA, the null hypothesis is almost never rejected (i.e., rejection rates approach 0%) with binary data. However, when the number of response alternatives was five, severely inflated type I error rates were observed. For example, when $\rho \leq$

.80, and *N* = 1,000, the null hypothesis is almost surely rejected (rejection rates approach 100%). Finally, for the SRMR, acceptable type I error rates were observed across all simulated conditions: empirical rejection rates at α = 5% ranged from 4% to 9 %.

Table 4 summarizes the power of tests of close fit. More specifically, at a 5% significance level, this table depicts the empirical rejection rates of tests that the RMSEA and SRMR are less than or equal to fixed cutoffs (*c*). In this study, the cut-offs (*c*) were defined as the corresponding population SRMR or RMSEA when fitting a one-factor model to two-factor data with inter-factor correlations ρ = .90. When the level of model misspecification is severe (i.e., ρ = .60 or ρ = .50), the power rates approached 100% across most of the simulated conditions. Therefore, only results for ρ = .80 or ρ = .70 are shown in this table. For both the SRMR and RMSEA, power rates increase as the level of misspecification increased (i.e., smaller ρ), the sample size increased, and the number of response categories increased (from 2 to 5); the latter is consistent with recent research (Maydeu-Olivares, Fairchild, & Hall, 2017). In addition, compared to the unbiased RMSEA, the SRMR generally yielded higher rejection rates in small samples (i.e., *N* < 500). For example, when fitting a one-factor model to two-factor binary data with an inter-factor correlation of 0.8, *p* = 10 and *N* = 100, the power rate for the unbiased RMSEA was .12, whereas the power rate for SRMR was .25. It was also noted that when *N* reached 500, the difference between the power rates of RMSEA and SRMR became negligible. However, empirical rejection rates across different test statistics can only be compared if they have similar empirical Type I errors, which is not the case for all conditions examined. In particular, empirical rejection rates for the biased RMSEA cannot be compared to the other two statistics as the biased RMSEA has very different Type I errors.

A Numerical Example: Fitting Ordinal Factor Analysis Models to the Life Orientation Test

We offer a numeric example to demonstrate the process of obtaining the unbiased sample estimates and CIs for the SRMR and RMSEA. The R code and data used in this example are provided as supplementary materials to this article. In this example, we use data gathered from 389 respondents by Chang, D'Zurilla, and Maydeu-Olivares (1994). The scale of interest is the Life Orientation Test (LOT: Scheier & Carver, 1985), which was designed to measure generalized outcome expectancies. Responses are on a five-point Likert-type scale ranging from zero (strongly disagree) to four (strongly agree). The scale consists of eight items, four of the items are positively worded, while the remaining four items are negatively worded.

The LOT was designed to measure a unidimensional construct. Therefore, we first fitted a one-factor ordinal factor analysis model using ULSMV estimation to the data. We obtained an unbiased RMSEA of .138 with a 90% CI (.126, .151). Based on the simulation results, the confidence interval for RMSEA may not be trustworthy. It was also noted that when using the incorrect formula, which is currently implemented in many SEM software packages (e.g., Mplus), the sample RMSEA was .207, and the 90% CI was (.189, .226) which suggested that the model fits much worse. For the SRMR, the unbiased sample estimate was .103, and the corresponding 90% CI was (.090, .012). When using the biased formula, the sample SRMR was .105, which was negligibly larger than the unbiased estimate. We can see that a one-factor model does not fit well the data of this example.

We also fitted a correlated two-factor model in which all positively worded items load on one factor and all negatively worded items load on another factor. Results indicated that the two-factor model yielded a close fit to the data. The unbiased sample RMSEA was .036, and the 90% CI was (.020, .052); the biased sample RMSEA was .053 and the 90% CI was (.029, .076), again the biased estimate suggesting that the model fits more poorly than what it does. The unbiased sample SRMR was .027 with a 90% CI (.016, .038). When using the biased formula, the sample SRMR was .034. The findings from this example were consistent with previous studies (Chang, D'Zurilla, & Maydeu-Olivares, 1994; Chang & McBride-Chang, 1996; Marshall, Wortman, Kusulas, Hervig, & Vickers, 1992; Robinson-Whelen, Kim, MacCallum, & Kiecolt-Glaser, 1997; Scheier & Carver, 1985). That is, a one-factor model does not fit well. Instead, a correlated two-factor model in which all positively worded items load on one factor and all negatively worded items load on another factor provides a substantially better fit to the data.

Discussion and Conclusions

Despite the fact that ordinal factor analysis models are widely applied in psychological and educational studies, little is known on how to assess the degree of model misfit in these models. In this study, we introduced statistical theory for using the SRMR to test close fit in ordinal factor analysis models. In addition, the performance of the SRMR was compared with that of the RMSEA in terms of the accuracy of the point estimates, confidence intervals (CIs), and tests of close fit when fitting ordinal factor analysis models using ULS estimation from tetrachoric and polychoric correlations.

Our simulation results show that for both the SRMR and RMSEA, when the unbiased/correct formulae are being used, the average sample estimates are generally close to their population values, even in small samples (e.g., $N = 100$). For the SRMR, the p values for testing close fit (i.e. $SRMR \leq SRMR_p$) produce acceptable 5% empirical rejection rates across all simulation conditions; the 90% CIs also generally yield adequate coverage rates, particularly when the sample size is large enough (e.g., $N \geq 500$). On the other hand, confidence intervals and tests of close fit based on the unbiased RMSEA can be inaccurate, especially when fitting large models (e.g., $p \geq 20$), or/and models with a high level of misspecification. Under these conditions, the coverage rates for the CIs are low, and the empirical rejection rates of the test of close fit are noticeably greater than the nominal level (i.e., inflated type I error rates), even in large samples (e.g., $N = 1{,}000$). The situation worsens considerably when the biased formula currently implemented in software is used to compute the sample RMSEA and the corresponding CIs. For binary data, the null hypothesis of close fit is almost never rejected, and for five category data, the null hypothesis of close fit is almost surely rejected in large samples (e.g., $N = 1{,}000$).

When comparing the performance of SRMR vs. RMSEA, the SRMR produced more accurate tests of close fit and confidence intervals across almost every condition tested. The confidence intervals for the unbiased RMSEA could only outperform those for the SRMR when the number of variables is small (e.g. $p = 10$), and the degree of misfit is small (e.g. a one-factor model is fitted to a model with two factors that correlate 0.9). In addition, the SRMR generally yields higher power rates to reject models that do not fit closely, especially in small samples (i.e., $N \leq 200$).

Why do the unbiased RMSEA produce relatively unbiased sample estimates, but fail to yield accurate confidence intervals and test of close fit? We think essentially our observations can be explained by findings in Shi, DiStefano, McDaniel and Jiang (2018): The robust test statistic with categorical data (e.g., T_{ULSMV}) is not well approximated by the reference chi-square distribution, especially as the number of observations (*N*) increases. Under correctly specified models, Shi et al. (2018) found that the empirical mean of T_{ULSMV} was generally close to the theoretical value (*df*), but that the empirical variance of T_{ULSMV} severely underestimated the theoretical value (i.e. smaller than 2*df*). As a result, the point estimate of RMSEA based on

T_{ULSMV} is generally close to its population value, but confidence intervals and statistical tests (*p* values) can be inaccurate.

We also note that for the SRMR, the empirical rejection rates for testing close fit (i.e. $SRMR \leq SRMR_p$) were more accurate than the CI coverage rates. That is, the 5% empirical rejection rates are acceptable across all simulation conditions but the 90% CI yield inadequate coverage rates, particularly under conditions where $N \leq 500$ and the degree of misfit is small (e.g., $\rho = .90$). The reason is as follows. According to its definition, the SRMR is bounded below by zero. Since the empirical distribution of the SRMR is left-truncated, a normal distribution approximation may fail to represent the observations in the left tail, particularly when the degree of model misfit is small (i.e., producing more zeros). A demonstration of this pattern is shown in Figure 1, where we plotted two empirical distributions of the sample unbiased SRMR against its reference normal distribution under two different levels of model misspecification. In the right panel of Figure 1, we plotted the SRMR values in a condition where the level of model misspecification is relatively large (i.e., $\rho = 0.7$, data with five categories, $p = 20$, and $N = 200$, population SRMR = 0.078). When the population value of the SRMR is not close to zero, the empirical distribution of the sample SRMR can be well approximated by a normal distribution, even in small samples.

On the other hand, as shown in the left panel of Figure 1, when the level of model misspecification is small and the sample size is small (i.e., $\rho = 0.9$, data with five categories, $p = 20$, and $N = 200$, population SRMR = 0.026), the left tail of the distribution is truncated by the value of zero. As a result, in these conditions, the CI, which is constructed as two-sided, does not perform adequately. However, the *p*-value is computed using only the right tail (

$H_0 : SRMR \leq SRMR_p$), and therefore, for the same condition, the 5% empirical rejection rate is still acceptable.

In simulations reported in this article, only one type of model misspecification was considered (i.e., misspecified dimensionality). As a result, the conclusions are somewhat limited. To explore the generalizability of the findings to other types of misspecified models, we conducted additional simulations where the model misspecification was introduced by omitting cross-loadings. The additional simulations are reported as supplementary materials and the results are consistent with the findings presented earlier. Based on the findings, we conjecture that keeping other factors (i.e., the effect size of misfit, sample size, model size, and the number of response categories) constant, the accuracy of CIs for the SRMR and RMSEA does not depend on the types of model misspecifications. Additional studies are needed to further investigate this issue. In addition, we only studied ordinal data with two or five response categories with the same values of population thresholds for all items. The performance of using RMSEA and SRMR to test close fit under conditions in which the items have different population thresholds and different number of response categories should be examined in future work.

The findings of the current research also offer implications for empirical studies. First, when reporting and interpreting the RMSEA and SRMR, researchers should be aware of the formula used by their software to compute the sample estimates. To produce relatively unbiased estimates, the unbiased (correct) formula should be applied. It is worth mentioning that in the current version of most SEM software (e.g., Mplus v8), the sample RMSEA from ordinal factor analysis under the “ULSMV” estimation does not consistently estimate the population RMSEA, and it estimates a different parameter. In addition, the sample SRMR computed based on Eq. 2

can be upwardly biased in small samples, therefore suggesting that the model fit more poorly than it actually does.

Second, when fitting ordinal factor analysis (using ULSMV estimation), the confidence intervals and test of close fit for the RMSEA, are not accurate enough, except when the number of variables is small (e.g., $p = 10$), and the degree of misfit is small. In contrast, confidence intervals for the SRMR are generally accurate, especially when the sample size is large ($N \geq 500$). The close fit tests based on the SRMR also yield acceptable type I error rates across all simulated conditions in this study. In addition, compared to the RMSEA, the SRMR shows higher power to reject non-close fit models, especially in small samples (i.e., $N \leq 200$). Therefore, the degree of misfit of an ordinal factor analysis model can be safely assessed using the SRMR. We hope that this paper provides useful information to assist applied researchers when conducting the difficult task of assessing the degree of misfit of a model.

Footnotes

1. We distinguish between effect sizes of model misfit and goodness of fit indices. Effect sizes of model misfit are population parameters that capture the discrepancy between the fitted model and the data generating process for which statistical theory is available, thus enabling the construction of confidence intervals, and, if of interest, statistical tests (Maydeu-Olivares, 2017). We reserve the term goodness of fit index to refer to a sample statistic used to adjudge model fit disregarding its sampling variability and without referencing any population parameter. Overviews of the array of the goodness of fit indices that have been proposed in the literature are Bollen and Long (1993) and Marsh, Hau, and Grayson (2005).

2. We recognize that under the special case that the observed variables are binary, the correlation between two unobserved continuous variables is referred to as tetrachoric correlation. In this paper, we use the more general term polychoric correlation for variables with any number of ordered categories.

3. In this paper, we focus on correlation structure models in which the threshold structure is saturated.

4. We also examined the behavior of 95% CIs. In this case, coverage rates between 90% to 99% were considered acceptable. For both the SRMR and RMSEA, similar patterns were observed across 90% and 95% CIs, though the 95% CIs seemed to produce acceptable coverage rates in slightly more conditions.

References

Asparouhov, T., & Muthén, B. (2010). *Simple second order chi-square correction scaled chi-square statistics* (Technical appendix). Los Angeles, CA.

Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(1), 102–116. https://doi.org/10.1080/10705511.2014.859510

Bollen, K. A., & Long, J. S. (1993). *Testing Structural Equation Models*. Newbury Park, CA: Sage.

Box, G. E. P. (1979). Some Problems of Statistics and Everyday Life. *Journal of the American Statistical Association*, *74*(365), 1–4. https://doi.org/10.1080/01621459.1979.10481600

Brosseau-Liard, P. E., & Savalei, V. (2014). Adjusting Incremental Fit Indices for Nonnormality. *Multivariate Behavioral Research*, *49*(5), 460–470. https://doi.org/10.1080/00273171.2014.933697

Brosseau-Liard, P. E., Savalei, V., & Li, L. (2012). An Investigation of the Sample Performance of Two Nonnormality Corrections for RMSEA. *Multivariate Behavioral Research*, *47*(6), 904–930. https://doi.org/10.1080/00273171.2012.715252

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. s. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chang, E. C., D'Zurilla, T. J., & Maydeu-Olivares, A. (1994). Assessing the dimensionality of optimism and pessimism using a multimeasure approach. *Cognitive Therapy and Research*, *18*(2), 143–160. https://doi.org/10.1007/BF02357221

Chang, L., & McBride-Chang, C. (1996). The Factor Structure of the Life Orientation Test.

Educational and Psychological Measurement, *56*(2), 325–329. https://doi.org/10.1177/0013164496056002013

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An Empirical Evaluation of the Use of Fixed Cutoff Points in RMSEA Test Statistic in Structural Equation Models. *Sociological Methods & Research*, *36*(4), 462–494. https://doi.org/10.1177/0049124108314720

Curran, P. J., Bollen, K. A., Chen, F., Paxton, P., & Kirby, J. B. (2003). Finite Sampling Properties of the Point Estimates and Confidence Intervals of the RMSEA. *Sociological Methods & Research*, *32*(2), 208–252. https://doi.org/10.1177/0049124103256130

DiStefano, C., & Morgan, G. B. (2014). A Comparison of Diagonal Weighted Least Squares Robust Estimation Techniques for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 425–438. https://doi.org/10.1080/10705511.2014.915373

Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 56–83. https://doi.org/10.1080/10705519909540119

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor Analysis with Ordinal Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation. *Structural Equation Modeling*, *16*(4), 625–641. https://doi.org/10.1080/10705510903203573

Hancock, G. R., Stapleton, L. M., & Mueller, R. O. (2018). *The Reviewer' s Guide to Quantitative Methods* (2nd ed.). New York: Routledge.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453.

https://doi.org/10.1037//1082-989X.3.4.424

Jöreskog, K. G., & Sörbom, D. (1988). LISREL 7. A guide to the program and applications (2nd ed.). Chicago, IL: International Education Services.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The Performance of RMSEA in Models With Small Degrees of Freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Kenny, D. A., & McCoach, D. B. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(3), 333–351. https://doi.org/10.1207/S15328007SEM1003_1

MacCallum, R. C. (2003). Working with Imperfect Models. *Multivariate Behavioral Research*, *38*(1), 113–139. https://doi.org/10.1207/S15327906MBR3801_5

Marsh, H., Hau, K., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. Mcardle (Eds.), *Contemporary Psychometrics: A Festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah, NJ: Erlbaum.

Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers, R. R. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology*, *62*(6), 1067–1074. https://doi.org/10.1037/0022-3514.62.6.1067

Maydeu-Olivares, A. (2005). Linear item response theory, nonlinear item response theory, and factor analysis: A unified framework. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary Psychometrics. A Festschrift for Roderick P. McDonald* (pp. 73–100). Mahwah, NJ: Lawrence Erlbaum.

Maydeu-Olivares, A. (2017). Assessing the Size of Model Misfit in Structural Equation Models.

Psychometrika, *82*(3), 533–558. https://doi.org/10.1007/s11336-016-9552-7

Maydeu-Olivares, A., Cai, L., & Hernández, A. (2011). Comparing the Fit of Item Response Theory and Factor Analysis Models. *Structural Equation Modeling*, *18*(3), 333–356. https://doi.org/10.1080/10705511.2011.581993

Maydeu-Olivares, A., Fairchild, A. J., & Hall, A. G. (2017). Goodness of Fit in Item Factor Analysis: Effect of the Number of Response Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(4), 495–505. https://doi.org/10.1080/10705511.2017.1289816

Maydeu-Olivares, A., Shi, D., & Rosseel, Y. (2018). Assessing Fit in Structural Equation Models: A Monte-Carlo Evaluation of RMSEA Versus SRMR Confidence Intervals and Tests of Close Fit. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 389–402. https://doi.org/10.1080/10705511.2017.1389611

McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*(1), 64–82. https://doi.org/10.1037//1082-989X.7.1.64

Muthén, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika*, *43*(4), 551–560. https://doi.org/10.1007/BF02293813

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *49*(1), 115–132. https://doi.org/10.1007/BF02294210

Muthén, B. (1993). Goodness of fit with categorical and other nonnormal variables. In K. a. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 205–234). Newbury Park, CA: Sage.

Muthén, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of*

Mathematical and Statistical Psychology, *45*(1), 19–30. https://doi.org/10.1111/j.2044-8317.1992.tb00975.x

Muthén, B. O., Toit, S. H. C. du, & Spisic, D. (1997). *Robust Inference using Weighted Least Squares and Quadratic Estimating Equations in Latent Variable Modeling with Categorical and Continuous Outcomes*. Retrieved from http://pages.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf

Muthén, L. K., & Muthén, B. (2017). MPLUS 8 [Computer program]. Los Angeles, CA: Muthén & Muthén.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo Study Investigating the Impact of Item Parceling on Measures of Fit in Confirmatory Factor Analysis. *Educational and Psychological Measurement*, *63*(5), 729–757. https://doi.org/10.1177/0013164403258228

Nevitt, J., & Hancock, G. R. (2000). Improving the Root Mean Square Error of Approximation for Nonnormal Conditions in Structural Equation Modeling. *The Journal of Experimental Education*, *68*(3), 251–268. https://doi.org/10.1080/00220970009600095

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, *44*(4), 443–460. https://doi.org/10.1007/BF02296207

Pornprasertmanit, S., Miller, P., & Schoemann, A. (2013). simsem: Simulated structural equation modeling. *R Package Version 0.5-3*.

R Development Core Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria.

Rhemtulla, M., Brosseau-Liard, P. É. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–73.

https://doi.org/10.1037/a0029315

Robinson-Whelen, S., Kim, C., MacCallum, R. C., & Kiecolt-Glaser, J. K. (1997). Distinguishing optimism from pessimism in older adults: Is it more important to be optimistic or not to be pessimistic? *Journal of Personality and Social Psychology*, *73*(6), 1345–1353. https://doi.org/10.1037/0022-3514.73.6.1345

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

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing Structural Equation Models or Detection of Misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 561–582. https://doi.org/10.1080/10705510903203433

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. Von Eye & C. C. Clogg (Eds.), *Latent variable analysis. Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Savalei, V. (2012). The Relationship Between Root Mean Square Error of Approximation and Model Misspecification in Confirmatory Factor Analysis Models. *Educational and Psychological Measurement*, *72*(6), 910–932. https://doi.org/10.1177/0013164412452564

Savalei, V. (2017). On the computation of the RMSEA and its confidence interval to accompany the mean and variance corrected test statistic with nonnormal data. *Structural Equation Modeling*.

Savalei, V., & Rhemtulla, M. (2013). The performance of robust test statistics with categorical data. *British Journal of Mathematical and Statistical Psychology*, *66*(2), 201–223. https://doi.org/10.1111/j.2044-8317.2012.02049.x

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, *4*(3), 219–247. https://doi.org/10.1037/0278-6133.4.3.219

Shi, D., DiStefano, C., & McDaniel, H. (2017). Examining Chi-Square Fit under Conditions of Large Model Size and Ordinal Data. *Manuscript Submitted for Publication*.

Shi, D., Maydeu-Olivares, A., & DiStefano, C. (2018). The Relationship Between the Standardized Root Mean Square Residual and Model Misspecification in Factor Analysis Models. *Multivariate Behavioral Research*, 1–19. https://doi.org/10.1080/00273171.2018.1476221

Steiger, J. H. (1989). EzPATH: A supplementary module for SYSTAT and SYGRAPH. Evanston, IL: Systat, Inc.

Steiger, J. H. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*, *25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4

Steiger, J. H., & Lind, J. C. (1980). Statistically-based tests for the number of common factors. Iowa: Paper presented at the Annual Meeting of the Psychometric Society.

Xia, Y. (2016). *Investigating the chi-square-based model-fit indexes for WLSMV and ULSMV estimators*. *Unpublished doctoral dissertation.* Florida State University.

Table 1: Population RMSEA and SRMR and average of sample RMSEA and SRMR across replications

`*p*	ρ	*N*	RMSEA					SRMR				
			POP	Sample. B		Sample. U		POP	Sample. B		Sample. U	
				CAT=2	CAT=5	CAT=2	CAT=5		CAT=2	CAT=5	CAT=2	CAT=5
10	0.9	100	*0.028*	**0.024**	**0.033**	0.028	**0.024**	*0.022*	**0.087**	**0.055**	0.024	0.020
		200	*0.028*	**0.021**	**0.036**	**0.024**	**0.025**	*0.022*	**0.063**	**0.042**	0.020	0.020
		500	*0.028*	**0.023**	**0.041**	**0.025**	0.027	*0.022*	**0.044**	**0.031**	0.020	0.022
		1000	*0.028*	**0.024**	**0.042**	0.026	0.027	*0.022*	**0.035**	**0.027**	0.021	0.022
	0.8	100	*0.055*	**0.037**	**0.066**	**0.047**	0.050	*0.044*	**0.096**	**0.068**	**0.039**	0.041
		200	*0.055*	**0.042**	**0.074**	**0.049**	0.054	*0.044*	**0.075**	**0.057**	0.040	0.043
		500	*0.055*	**0.047**	**0.079**	0.054	0.055	*0.044*	**0.058**	**0.050**	0.043	0.044
		1000	*0.055*	**0.048**	**0.080**	0.054	0.055	*0.044*	**0.052**	0.047	0.044	0.044
	0.7	100	*0.083*	**0.058**	**0.099**	0.075	0.079	*0.066*	**0.110**	**0.085**	0.062	0.064
		200	*0.083*	**0.065**	**0.107**	0.080	0.082	*0.066*	**0.091**	**0.076**	0.065	0.066
		500	*0.083*	**0.069**	**0.111**	0.082	0.083	*0.066*	**0.077**	0.070	0.066	0.066
		1000	*0.083*	**0.070**	**0.113**	0.083	0.083	*0.066*	0.072	0.068	0.066	0.066
	0.6	100	*0.110*	**0.078**	**0.128**	0.104	0.107	*0.088*	**0.126**	**0.103**	0.085	0.087
		200	*0.110*	**0.084**	**0.135**	0.107	0.109	*0.088*	**0.108**	**0.096**	0.087	0.088
		500	*0.110*	**0.089**	**0.139**	0.110	0.110	*0.088*	**0.097**	0.091	0.088	0.088
		1000	*0.110*	**0.090**	**0.141**	0.110	0.110	*0.088*	0.093	0.090	0.088	0.088
	0.5	100	*0.138*	**0.094**	**0.154**	0.129	0.135	*0.110*	**0.142**	**0.123**	0.105	0.109
		200	*0.138*	**0.103**	**0.160**	0.135	0.137	*0.110*	**0.127**	0.117	0.109	0.110
		500	*0.138*	**0.106**	**0.164**	0.136	0.138	*0.110*	0.117	0.113	0.110	0.110
		1000	*0.138*	**0.108**	**0.165**	0.137	0.138	*0.110*	0.114	0.112	0.110	0.110
20	0.9	100	*0.026*	**0.017**	0.026	0.025	**0.022**	*0.023*	**0.098**	**0.062**	0.024	0.021
		200	*0.026*	**0.018**	**0.032**	0.024	0.024	*0.023*	**0.071**	**0.046**	0.022	0.022
		500	*0.026*	**0.021**	**0.036**	0.025	0.025	*0.023*	**0.049**	**0.034**	0.022	0.023
		1000	*0.026*	**0.022**	**0.038**	0.025	0.026	*0.023*	**0.038**	**0.029**	0.023	0.023
	0.8	100	*0.052*	**0.032**	0.056	0.048	0.051	*0.047*	**0.108**	**0.076**	0.045	0.046
		200	*0.052*	**0.036**	**0.062**	0.050	0.051	*0.047*	**0.083**	**0.062**	0.045	0.046
		500	*0.052*	**0.041**	**0.068**	0.051	0.052	*0.047*	**0.064**	**0.053**	0.046	0.047
		1000	*0.052*	**0.043**	**0.071**	0.051	0.052	*0.047*	**0.056**	0.050	0.046	0.046
	0.7	100	*0.078*	**0.048**	0.081	0.075	0.077	*0.070*	**0.122**	**0.093**	0.069	0.070
		200	*0.078*	**0.054**	**0.088**	0.076	0.077	*0.070*	**0.099**	**0.082**	0.069	0.069
		500	*0.078*	**0.060**	**0.095**	0.077	0.077	*0.070*	**0.083**	0.075	0.069	0.070
		1000	*0.078*	**0.062**	**0.098**	0.077	0.077	*0.070*	0.076	0.072	0.070	0.070
	0.6	100	*0.103*	**0.062**	0.102	0.100	0.102	*0.093*	**0.138**	**0.112**	0.091	0.093
		200	*0.103*	**0.069**	0.109	0.101	0.102	*0.093*	**0.117**	**0.103**	0.092	0.093
		500	*0.103*	**0.076**	**0.117**	0.102	0.103	*0.093*	**0.103**	0.097	0.092	0.093
		1000	*0.103*	**0.079**	**0.119**	0.103	0.103	*0.093*	0.098	0.095	0.093	0.093
	0.5	100	*0.129*	**0.076**	0.121	0.126	0.128	*0.116*	**0.156**	**0.133**	0.115	0.117
		200	*0.129*	**0.084**	0.128	0.127	0.128	*0.116*	**0.137**	0.124	0.115	0.116
		500	*0.129*	**0.091**	0.135	0.129	0.129	*0.116*	0.125	0.120	0.116	0.116

		1000	*0.129*	**0.094**	0.137	0.129	0.129	*0.116*	0.121	0.118	0.116	0.116
30	0.9	100	*0.025*	**0.015**	**0.023**	0.026	0.023	*0.024*	**0.102**	**0.064**	0.026	0.023
		200	*0.025*	**0.016**	**0.028**	0.024	0.025	*0.024*	**0.074**	**0.048**	0.023	0.023
		500	*0.025*	**0.019**	**0.033**	0.025	0.025	*0.024*	**0.050**	**0.036**	0.023	0.024
		1000	*0.025*	**0.021**	**0.036**	0.025	0.025	*0.024*	**0.039**	**0.030**	0.024	0.024
	0.8	100	*0.051*	**0.028**	0.048	0.050	0.050	*0.047*	**0.113**	**0.078**	0.049	0.048
		200	*0.051*	**0.032**	0.055	0.050	0.051	*0.047*	**0.086**	**0.064**	0.048	0.048
		500	*0.051*	**0.037**	**0.062**	0.051	0.051	*0.047*	**0.066**	**0.055**	0.048	0.048
		1000	*0.051*	**0.040**	**0.066**	0.051	0.051	*0.047*	**0.057**	0.051	0.048	0.048
	0.7	100	*0.076*	**0.041**	**0.069**	0.076	0.076	*0.071*	**0.127**	**0.096**	0.072	0.072
		200	*0.076*	**0.047**	**0.077**	0.076	0.076	*0.071*	**0.103**	**0.084**	0.071	0.071
		500	*0.076*	**0.054**	**0.085**	0.076	0.076	*0.071*	**0.085**	0.077	0.071	0.072
		1000	*0.076*	**0.057**	**0.089**	0.076	0.076	*0.071*	0.078	0.074	0.071	0.071
	0.6	100	*0.102*	**0.053**	**0.087**	0.100	0.101	*0.095*	**0.143**	**0.115**	0.095	0.095
		200	*0.102*	**0.060**	0.095	0.101	0.101	*0.095*	**0.121**	**0.106**	0.095	0.095
		500	*0.102*	**0.068**	0.104	0.102	0.102	*0.095*	**0.106**	0.099	0.095	0.095
		1000	*0.102*	**0.072**	0.107	0.102	0.102	*0.095*	0.101	0.097	0.095	0.095
	0.5	100	*0.127*	**0.064**	**0.103**	0.125	0.126	*0.118*	**0.161**	**0.136**	0.118	0.119
		200	*0.127*	**0.072**	**0.111**	0.126	0.127	*0.118*	**0.141**	0.128	0.118	0.119
		500	*0.127*	**0.081**	0.118	0.127	0.127	*0.118*	0.129	0.122	0.119	0.119
		1000	*0.127*	**0.085**	0.121	0.127	0.127	*0.118*	0.124	0.120	0.119	0.119

Notes: *p* = # observed variables; *N* = sample size; *pop.* = population value; *RMSEA B* = biased RMSEA, as implemented in current software, *RMSEA U* = unbiased RMSEA; *SRMR B* = biased SRMR, as implemented in current software, *SRMR U* = unbiased SRMR.

Table 2: Coverage rates for 90% confidence intervals around the population RMSEA and SRMR

p	ρ	*N*	CAT = 2			CAT= 5		
			RSMEA B	**RMSEA U**	**SRMR**	**RSMEA B**	**RMSEA U**	**SRMR**
10	0.9	100	**0.98**	**0.98**	0.91	0.93	0.95	**0.78**
		200	**0.97**	**0.96**	0.89	**0.84**	0.94	**0.75**
		500	**0.96**	0.95	**0.74**	**0.57**	0.85	**0.84**
		1000	0.88	0.89	0.85	**0.26**	**0.81**	0.86
	0.8	100	**0.99**	**0.97**	**0.71**	**0.81**	0.87	**0.80**
		200	0.85	0.93	**0.79**	**0.56**	**0.79**	0.85
		500	**0.73**	**0.81**	0.85	**0.18**	**0.74**	0.87
		1000	**0.64**	**0.79**	0.86	**0.02**	**0.71**	0.88
	0.7	100	**0.80**	0.92	**0.82**	**0.72**	**0.76**	**0.83**
		200	**0.67**	**0.81**	0.86	**0.44**	**0.70**	0.88
		500	**0.51**	**0.76**	0.88	**0.08**	**0.70**	0.89
		1000	**0.35**	**0.75**	0.89	**0.01**	**0.63**	0.88
	0.6	100	**0.59**	**0.81**	**0.82**	**0.66**	**0.70**	0.85
		200	**0.42**	**0.75**	0.86	**0.42**	**0.68**	0.88
		500	**0.24**	**0.71**	0.88	**0.08**	**0.64**	0.89
		1000	**0.07**	**0.72**	0.88	**0.00**	**0.61**	0.88
	0.5	100	**0.37**	**0.76**	0.84	**0.64**	**0.65**	0.85
		200	**0.23**	**0.70**	0.87	**0.44**	**0.64**	0.89
		500	**0.05**	**0.67**	0.87	**0.13**	**0.62**	0.89
		1000	**0.01**	**0.66**	0.90	**0.01**	**0.58**	0.89
20	0.9	100	**1.00**	**1.00**	0.92	**0.96**	**0.98**	**0.72**
		200	**1.00**	**0.99**	**0.74**	**0.84**	0.92	**0.80**
		500	**0.83**	0.90	**0.79**	**0.34**	**0.79**	0.85
		1000	**0.72**	**0.84**	**0.82**	**0.03**	**0.74**	0.87
	0.8	100	**0.77**	**0.97**	**0.84**	**0.84**	**0.83**	**0.82**
		200	**0.46**	0.88	**0.83**	**0.54**	**0.72**	0.86
		500	**0.28**	**0.74**	**0.83**	**0.05**	**0.59**	0.85
		1000	**0.17**	**0.69**	0.86	**0.00**	**0.57**	0.88
	0.7	100	**0.22**	0.88	0.86	**0.73**	**0.68**	**0.83**
		200	**0.08**	**0.77**	0.86	**0.47**	**0.58**	0.86
		500	**0.04**	**0.63**	0.87	**0.04**	**0.48**	0.87
		1000	**0.01**	**0.59**	0.87	**0.00**	**0.50**	0.89
	0.6	100	**0.03**	**0.79**	0.85	**0.67**	**0.60**	**0.83**
		200	**0.00**	**0.62**	0.86	**0.54**	**0.50**	0.87
		500	**0.00**	**0.57**	0.86	**0.15**	**0.45**	0.86
		1000	**0.00**	**0.53**	0.88	**0.01**	**0.46**	0.89
	0.5	100	**0.01**	**0.70**	0.85	**0.56**	**0.52**	**0.84**
		200	**0.00**	**0.59**	0.88	**0.59**	**0.45**	0.88

		500	**0.00**	**0.51**	0.87	**0.49**	**0.42**	0.87
		1000	**0.00**	**0.50**	0.89	**0.20**	**0.42**	0.89
30	0.9	100	**1.00**	**1.00**	**0.82**	**0.99**	**0.99**	**0.78**
		200	0.90	**0.99**	**0.77**	0.88	0.91	**0.81**
		500	**0.56**	0.92	**0.81**	**0.25**	**0.78**	0.85
		1000	**0.38**	**0.83**	**0.83**	**0.00**	**0.69**	0.87
	0.8	100	**0.25**	**0.98**	0.89	**0.84**	**0.81**	**0.82**
		200	**0.07**	0.88	0.85	**0.66**	**0.64**	**0.83**
		500	**0.02**	**0.70**	**0.83**	**0.06**	**0.55**	0.88
		1000	**0.01**	**0.64**	0.87	**0.00**	**0.50**	0.88
	0.7	100	**0.01**	0.91	0.88	**0.58**	**0.60**	**0.81**
		200	**0.00**	**0.71**	**0.84**	**0.61**	**0.52**	**0.83**
		500	**0.00**	**0.56**	0.88	**0.15**	**0.43**	0.89
		1000	**0.00**	**0.50**	0.90	**0.00**	**0.44**	0.89
	0.6	100	**0.00**	**0.77**	0.86	**0.33**	**0.51**	**0.81**
		200	**0.00**	**0.61**	0.85	**0.42**	**0.42**	0.86
		500	**0.00**	**0.49**	0.88	**0.50**	**0.39**	0.89
		1000	**0.00**	**0.49**	0.88	**0.19**	**0.40**	0.90
	0.5	100	**0.00**	**0.65**	**0.84**	**0.14**	**0.43**	**0.82**
		200	**0.00**	**0.53**	0.85	**0.15**	**0.38**	0.86
		500	**0.00**	**0.44**	0.89	**0.17**	**0.38**	0.90
		1000	**0.00**	**0.42**	0.89	**0.21**	**0.36**	0.89

Notes: *p* = # observed variables; *CAT* = # of response categories; *N* = sample size; *pop.* = population value; *RMSEA B* = biased RMSEA, as implemented in current software, *RMSEA U* = unbiased RMSEA; *SRMR B* = biased SRMR, as implemented in current software, *SRMR U* = unbiased SRMR.

Table 3: Test of close fit results. Empirical rejection rates at a 5% significance level of a test that the RMSEA and SRMR less or equal to their population values

`*p*	ρ	*N*		RMSEA B		RMSEA U		SRMR		
			POP	*CAT*= 2	*CAT*= 5	*CAT*= 2	*CAT*= 5	POP	*CAT*= 2	*CAT* = 5
10	0.9	100	*0.028*	0.02	0.08	0.03	0.05	*0.022*	0.09	0.08
		200	*0.028*	0.03	**0.16**	0.05	0.06	*0.022*	0.09	0.07
		500	*0.028*	0.04	**0.43**	0.05	0.08	*0.022*	0.08	0.06
		1000	*0.028*	0.03	**0.74**	0.06	0.10	*0.022*	0.06	0.07
	0.8	100	*0.055*	0.02	**0.20**	0.04	0.08	*0.044*	0.05	0.06
		200	*0.055*	0.02	**0.44**	0.07	0.10	*0.044*	0.05	0.04
		500	*0.055*	0.02	**0.82**	0.10	**0.13**	*0.044*	0.05	0.06
		1000	*0.055*	0.02	**0.98**	0.09	**0.14**	*0.044*	0.04	0.06
	0.7	100	*0.083*	0.02	**0.28**	0.07	**0.11**	*0.066*	0.04	0.06
		200	*0.083*	0.01	**0.57**	0.10	**0.15**	*0.066*	0.04	0.04
		500	*0.083*	0.01	**0.92**	**0.12**	**0.15**	*0.066*	0.04	0.05
		1000	*0.083*	**0.00**	**1.00**	**0.13**	**0.18**	*0.066*	0.05	0.05
	0.6	100	*0.110*	0.01	**0.32**	0.10	**0.14**	*0.088*	0.05	0.04
		200	*0.110*	**0.00**	**0.58**	**0.11**	**0.14**	*0.088*	0.04	0.04
		500	*0.110*	**0.00**	**0.92**	**0.13**	**0.18**	*0.088*	0.04	0.05
		1000	*0.110*	**0.00**	**1.00**	**0.13**	**0.19**	*0.088*	0.05	0.05
	0.5	100	*0.138*	**0.00**	**0.33**	**0.11**	**0.16**	*0.110*	0.04	0.04
		200	*0.138*	**0.00**	**0.56**	**0.13**	**0.17**	*0.110*	0.04	0.04
		500	*0.138*	**0.00**	**0.87**	**0.14**	**0.19**	*0.110*	0.05	0.05
		1000	*0.138*	**0.00**	**0.99**	**0.16**	**0.21**	*0.110*	0.05	0.06
20	0.9	100	*0.026*	**0.00**	**0.04**	**0.00**	0.02	*0.023*	0.08	0.08
		200	*0.026*	**0.00**	**0.16**	0.01	0.05	*0.023*	0.08	0.06
		500	*0.026*	0.01	**0.66**	0.05	0.10	*0.023*	0.07	0.05
		1000	*0.026*	0.02	**0.97**	0.09	**0.13**	*0.023*	0.07	0.04
	0.8	100	*0.052*	**0.00**	**0.15**	0.03	0.10	*0.047*	0.05	0.06
		200	*0.052*	**0.00**	**0.46**	0.07	**0.14**	*0.047*	0.05	0.04
		500	*0.052*	**0.00**	**0.95**	**0.12**	**0.20**	*0.047*	0.06	0.06
		1000	*0.052*	**0.00**	**1.00**	**0.14**	**0.21**	*0.047*	0.05	0.05
	0.7	100	*0.078*	**0.00**	**0.20**	0.08	**0.15**	*0.070*	0.05	0.05
		200	*0.078*	**0.00**	**0.51**	**0.11**	**0.19**	*0.070*	0.04	0.04
		500	*0.078*	**0.00**	**0.96**	**0.17**	**0.24**	*0.070*	0.05	0.06
		1000	*0.078*	**0.00**	**1.00**	**0.19**	**0.23**	*0.070*	0.05	0.04
	0.6	100	*0.103*	**0.00**	**0.15**	0.10	**0.18**	*0.093*	0.04	0.06
		200	*0.103*	**0.00**	**0.38**	**0.16**	**0.23**	*0.093*	0.04	0.04
		500	*0.103*	**0.00**	**0.85**	**0.19**	**0.26**	*0.093*	0.05	0.06
		1000	*0.103*	**0.00**	**0.99**	**0.21**	**0.25**	*0.093*	0.04	0.05
	0.5	100	*0.129*	**0.00**	**0.11**	**0.14**	**0.23**	*0.116*	0.04	0.05
		200	*0.129*	**0.00**	**0.20**	**0.18**	**0.25**	*0.116*	0.03	0.04
		500	*0.129*	**0.00**	**0.47**	**0.23**	**0.27**	*0.116*	0.05	0.05
		1000	*0.129*	**0.00**	**0.79**	**0.23**	**0.27**	*0.116*	0.04	0.04
30	0.9	100	*0.025*	**0.00**	0.01	**0.00**	0.01	*0.024*	0.07	0.07
		200	*0.025*	**0.00**	**0.13**	0.01	0.05	*0.024*	0.08	0.08
		500	*0.025*	**0.00**	**0.75**	0.05	**0.12**	*0.024*	0.07	0.07
		1000	*0.025*	0.01	**1.00**	0.10	**0.15**	*0.024*	0.08	0.06

	0.8	100	*0.051*	**0.00**	**0.08**	0.02	**0.12**	*0.047*	0.04	0.08
		200	*0.051*	**0.00**	**0.30**	0.08	**0.18**	*0.047*	0.07	0.07
		500	*0.051*	**0.00**	**0.94**	**0.16**	**0.24**	*0.047*	0.07	0.06
		1000	*0.051*	**0.00**	**1.00**	**0.18**	**0.26**	*0.047*	0.06	0.05
	0.7	100	*0.076*	**0.00**	**0.08**	0.07	**0.20**	*0.071*	0.05	0.08
		200	*0.076*	**0.00**	**0.22**	**0.15**	**0.24**	*0.071*	0.07	0.07
		500	*0.076*	**0.00**	**0.85**	**0.23**	**0.30**	*0.071*	0.05	0.05
		1000	*0.076*	**0.00**	**1.00**	**0.25**	**0.29**	*0.071*	0.05	0.05
	0.6	100	*0.102*	**0.00**	0.03	**0.13**	**0.24**	*0.095*	0.05	0.07
		200	*0.102*	**0.00**	0.09	**0.19**	**0.28**	*0.095*	0.05	0.05
		500	*0.102*	**0.00**	**0.38**	**0.26**	**0.32**	*0.095*	0.06	0.06
		1000	*0.102*	**0.00**	**0.81**	**0.25**	**0.31**	*0.095*	0.05	0.04
	0.5	100	*0.127*	**0.00**	**0.00**	**0.17**	**0.27**	*0.118*	0.05	0.07
		200	*0.127*	**0.00**	0.01	**0.22**	**0.30**	*0.118*	0.05	0.05
		500	*0.127*	**0.00**	0.01	**0.29**	**0.32**	*0.118*	0.05	0.05
		1000	*0.127*	**0.00**	0.01	**0.30**	**0.33**	*0.118*	0.05	0.05

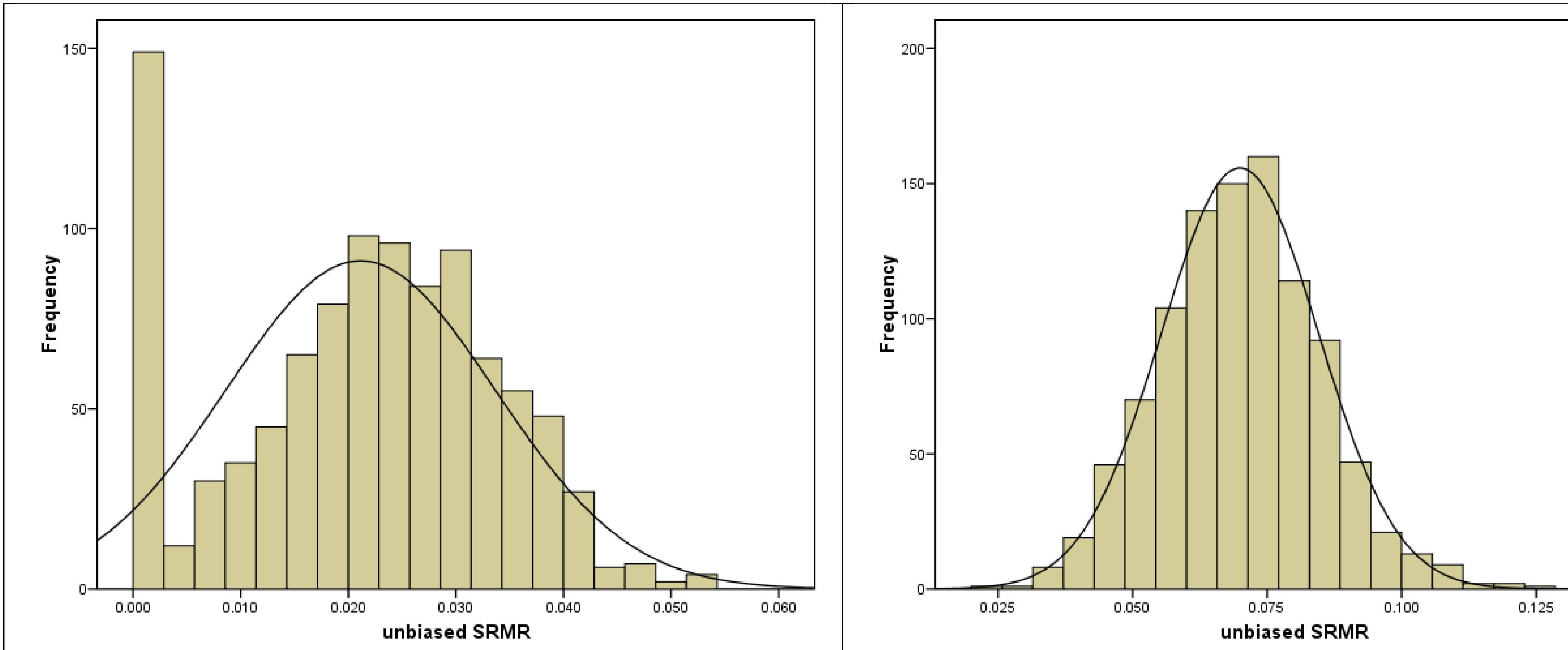
Notes: *p* = # observed variables; *CAT* = # of response categories; *N* = sample size; *pop.* = population value; *RMSEA B* = biased RMSEA, as implemented in current software, *RMSEA U* = unbiased RMSEA; *SRMR B* = biased SRMR, as implemented in current software, *SRMR U* = unbiased SRMR.

Table 4: Power for Test of close fit. Empirical rejection rates at a 5% significance level of a test that the RMSEA and SRMR less or equal to a fixed cutoff (c)

p	ρ	*N*	RMSEA						SRMR			
			pop	*c*	Power B		Power U		pop	*c*	Power	
					CAT=2	CAT=5	CAT=2	CAT=5			CAT=2	CAT=5
10	0.8	100	0.055	0.028	0.10	0.44	0.12	0.36	0.044	0.022	0.25	0.44
		200	0.055	0.028	0.23	0.84	0.29	0.71	0.044	0.022	0.40	0.71
		500	0.055	0.028	0.63	1.00	0.74	0.98	0.044	0.022	0.75	0.97
		1000	0.055	0.028	0.91	1.00	0.95	1.00	0.044	0.022	0.95	1.00
	0.7	100	0.083	0.028	0.31	0.85	0.35	0.81	0.066	0.022	0.55	0.85
		200	0.083	0.028	0.69	1.00	0.75	0.99	0.066	0.022	0.82	0.99
		500	0.083	0.028	0.99	1.00	0.99	1.00	0.066	0.022	0.99	1.00
		1000	0.083	0.028	1.00	1.00	1.00	1.00	0.066	0.022	1.00	1.00
20	0.8	100	0.052	0.026	0.07	0.64	0.13	0.59	0.047	0.023	0.44	0.72
		200	0.052	0.026	0.32	0.98	0.51	0.95	0.047	0.023	0.70	0.95
		500	0.052	0.026	0.89	1.00	0.97	1.00	0.047	0.023	0.98	1.00
		1000	0.052	0.026	1.00	1.00	1.00	1.00	0.047	0.023	1.00	1.00
	0.7	100	0.078	0.026	0.43	0.97	0.56	0.96	0.070	0.023	0.83	0.98
		200	0.078	0.026	0.91	1.00	0.96	1.00	0.070	0.023	0.99	1.00
		500	0.078	0.026	1.00	1.00	1.00	1.00	0.070	0.023	1.00	1.00
		1000	0.078	0.026	1.00	1.00	1.00	1.00	0.070	0.023	1.00	1.00
30	0.8	100	0.051	0.025	0.04	0.67	0.14	0.69	0.047	0.024	0.54	0.84
		200	0.051	0.025	0.27	1.00	0.65	0.99	0.047	0.024	0.85	0.99
		500	0.051	0.025	0.93	1.00	1.00	1.00	0.047	0.024	1.00	1.00
		1000	0.051	0.025	1.00	1.00	1.00	1.00	0.047	0.024	1.00	1.00
	0.7	100	0.076	0.025	0.41	0.98	0.65	0.99	0.071	0.024	0.92	1.00
		200	0.076	0.025	0.93	1.00	1.00	1.00	0.071	0.024	1.00	1.00
		500	0.076	0.025	1.00	1.00	1.00	1.00	0.071	0.024	1.00	1.00
		1000	0.076	0.025	1.00	1.00	1.00	1.00	0.071	0.024	1.00	1.00

Notes: *p* = # observed variables; *CAT* = # of response categories; *N* = sample size; *pop.* = population value; *RMSEA B* = biased RMSEA, as implemented in current software, *RMSEA U* = unbiased RMSEA; *SRMR B* = biased SRMR, as implemented in current software, *SRMR U* = unbiased SRMR.

Figure 1: Empirical Distributions of the unbiased SRMR against its reference normal distribution



a. Five categories, $p = 20$, $N = 200$, and $\rho = 0.9$

b. Five categories, $p = 20$, $N = 200$, and $\rho = 0.7$