Original Article

Correcting the Bias of the Root Mean Squared Error of Approximation Under Missing Data

Cailey E. Fitzgerald[1], Ryne Estabrook[2], Daniel P. Martin[1], Andreas M. Brandmaier[3],

Timo von Oertzen[1,4]

[1] *Department of Psychology, University of Virginia, Charlottesville, VA, USA.* **[2]** *Department of Psychology, Northwestern University, Evanston, IL, USA.* **[3]** *Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.* **[4]** *Department of Psychology, Germany University of the Federal Forces, Munich, Germany.*

Abstract

Missing data are ubiquitous in psychological research. They may come about as an unwanted result of coding or computer error, participants' non-response or absence, or missing values may be intentional, as in planned missing designs. We discuss the effects of missing data on χ^2-based goodness-of-fit indices in Structural Equation Modeling (SEM), specifically on the Root Mean Squared Error of Approximation (RMSEA). We use simulations to show that naive implementations of the RMSEA have a downward bias in the presence of missing data and, thus, overestimate model goodness-of-fit. Unfortunately, many state-of-the-art software packages report the biased form of RMSEA. As a consequence, the scientific community may have been accepting a much larger fraction of models with non-acceptable model fit. We propose a bias-correction for the RMSEA based on information-theoretic considerations that take into account the expected misfit of a person with fully observed data. The corrected RMSEA is asymptotically independent of the proportion of missing data for misspecified models. Importantly, results of the corrected RMSEA computation are identical to naive RMSEA if there are no missing data.

Keywords

Structural Equation Modeling (SEM; Bollen, 1989) is a widely used and powerful technique for the analysis of multivariate data. A host of common models can be specified and estimated within the SEM paradigm, including models based on the general linear model (GLM) like regression and ANOVA (see e.g., Miller, 1997), and more complex techniques ranging from factor analysis (Mulaik, 1972) growth mixture models with non-linear growth components (Grimm et al., 2010) to general multi-level models (Bauer, 2003; Curran, 2003). Before substantial interpretations are warranted, models are usually assessed using fit indices describing how well the model fits the data. Fit indices typically compare the proposed model to either a saturated model or both a saturated and an independence model. Saturated models include all possible means, variances and covariances, and represent the best possible fit of a covariance structure to the data, while independence models include only means and variances (i.e., assume all covariances are zero) and represent the worst reasonable fit of a covariance structure. Fit indices like Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980) describe absolute misfit of a model relative to the saturated model, while indices like the Comparative Fit Index and Tucker Lewis Index (Tucker & Lewis, 1973) place the proposed model on a continuum between the saturated and independence models.

Since its early days as a modeling framework of variances and covariances, the theory surrounding SEM has greatly expanded to include multi-level modeling (Heck, 2001), mixture modeling (Muthén & Shedden, 1999), non-linear growth modeling (Grimm & Ram, 2009), and even extensions to statistical learning approaches (Brandmaier et al., 2013). One important extension is the implementation of Full-Information Maximum Likelihood (FIML; cf. Finkbeiner, 1979) estimation based on raw data instead of covariance matrices. Early SEM packages relied solely on observed moment matrices (e.g., covariance matrices, mean vectors, SSCP matrices) as input, requiring researchers to convert their data into the appropriate matrices and handle missing data, ordinal data, and other complexities prior to model fitting. Modern SEM packages can use FIML to handle raw data, eschewing user-generated data reduction in favor of fitting the model to each row of the data individually. This allows SEM to handle missing values by applying the model to whatever values are observed for a particular row of data. FIML is robust to missingness under both the missing completely at random (MCAR) and missing at random (MAR) mechanisms (Little & Rubin, 1987; Rubin, 2002), yielding unbiased parameter estimates under ignorable missingness conditions and performs better than other missing data methods, such as similar response pattern imputation or listwise deletion (Enders & Bandalos, 2001).

While the methods by which we fit SEM have changed, the methods by which we assess the goodness-of-fit of these models largely have not. Most of the common fit indices used to assess model fit make certain assumptions about the models being compared and the data that were generated from them, namely that the models contain full rank covariance matrices with single values for sample size and degrees of freedom. To the

extent that a particular model uses raw data under FIML and is affected by missingness, these assumptions are decidedly false and thus bias is introduced into fit indices. For example, the RMSEA quantifies the amount of noncentrality per the product of sample size and degrees of freedom. When the data are complete, the product of adjusted sample size and degrees of freedom contain one measure of the total amount of information in the sample. When some portion of the data are missing, the χ^2 will decrease while the sample size and degrees of freedom values stay constant, which reduces the absolute value of the RMSEA statistic and make fit appear to be better than it would be with complete data. This artificial improvement in model fit has been noted in previous research (i.e., Davey, 2005; Hoyle, 2012) with no known solution to the problem. Zhang and Savalei (2020) have recently investigated this issue more closely and demonstrated that factors such as the type and degree of misfit in the hypothesized model, the type of missing data mechanism, and the number of missing data variables and patterns play an important role in this distortion.

Despite the bias in fit statistics, using fit statistics in the presence of missing data remains common practice. Google Scholar search indicates approximately 47% of articles (or 37,400 of 79,931) that cited one of two popular articles with recommended "rules of thumb" for RMSEA (Browne & Cudeck, 1993; Hu & Bentler, 1999) also include either the terms "missing" or "full information".[1] While this search is imperfect and nowhere near exhaustive, it gives some background toward the widespread use of fit statistics in SEM and, more importantly, the widespread use of fit statistics in SEM with missing data.

Summary of This Article

In this article, we will first discuss the problem that missing data introduces into the realm of model fit indices in SEM, focusing on the RMSEA. We then introduce the mathematical basis for a bias correction of the RMESA. Next, we present data simulations and plots to illustrate the problem and demonstrate to what extent the proposed solutions are better suited to capture model misfit under missing data than the uncorrected RMSEA. We discuss our motivation for a correction, broader implications of the problem of missing data in SEM, limitations of our correction, and future directions in the research of model misfit indices.

---

1) As of January 28, 2021.

PsychOpen GOLD

Method

Mathematical Statement of the Problem

The RMSEA is an estimate based on the quantity $\sqrt{\frac{F}{df}}$, where df represents the difference of degrees of freedom between a fitted model and a saturated model, and F is a population quantity of the discrepancy between a normal population distribution and the best-fitting normal distribution under a given model. F is thus independent of sample size and, conventionally, $\frac{\chi^2}{N}$ is used as estimator of this discrepancy. The resulting population value, $\sqrt{\frac{F}{df}}$, is thus a measure of absolute misfit per person and per degree of freedom. The fraction $\sqrt{\frac{\chi^2}{N \cdot df}}$ is an estimate of this population quantity for model misfit, which forms the basis of RMSEA.

For complete data, the RMSEA for a proposed model is defined as

$$\text{RMSEA}_{\text{uncorrected}} = \sqrt{\max\left(0, \frac{\chi^2 - df}{(N-1)df}\right)} \tag{1}$$

where df is the difference in degrees of freedom between the proposed model and a saturated normal model, N the number of participants, and χ^2 the difference in the minus two log-likelihood between the proposed and saturated models for the given data set. Without model misspecification, the χ^2 index will be distributed as a central χ^2 distribution with df degrees of freedom; therefore, the fraction will have an expected value of zero, with decreasing sampling variance as N increases. If the model is misspecified, χ^2 follows a non-central χ^2 distribution with a non-centrality parameter proportional to N, and thus RMSEA will converge towards a fixed non-zero number in expectation. The degree to which RMSEA is larger than zero is therefore a measure of model misfit.

As a data set has an increasing proportion of missing data (e.g., assuming missing completely at random; MCAR), the minus two log likelihoods of both the proposed and saturated models will in expectation decrease proportionally to that amount. In consequence, the χ^2 – the difference between the two – will also decrease proportionally to the proportion of missing data, while df and N remain constant. Hence, RMSEA values will become smaller in value and, hence, more optimistic for increasing proportions of missing data. Note that when there is no model misspecification, the proportion of missing data does not bias RMSEA.

Defining the Correction

How can we best correct the computation of RMSEA such that it is invariant under the proportion of missing data for misspecified models? We require that the bias-correction must leave RMSEA unchanged if there are no missing data. Second, we require the bias-correction to yield asymptotically identical RMSEA values under misspecification and different levels of missing data under the assumption of MCAR and MAR. In the

following, we propose a method to correct the bias incurred in the RMSEA by missing data.

To account for the downward bias in the χ^2 as estimator of model misfit, the χ^2 value can be replaced by an estimate of its expected value if no missing data were present. We obtain this bias correction by computing the expected divergence of a single person with complete observations and the best-fitting model and multiply this by the effective sample size, that is, the sample size multiplied by the proportion of observed values, to obtain the expected discrepancy if there had been no missing data.

Under MCAR or MAR missingness, the model-implied covariance matrix from a model with missing data is an estimate of the population covariance matrix, and by extension, of the model-implied covariance matrix if no missing data were present. Using the Kullback-Leibler divergence (Kullback & Leibler, 1951), also known as relative entropy, between the model-implied distribution and the saturated model's distribution, we compute the expected value of the χ^2 value of a single, fully observed person. Hence, the negative two log-likelihood of the complete data can be estimated as $2N$ times the Kullback–Leibler (Kullback & Leibler, 1951) divergence between the two distributions, just as $2N$ is multiplied by the fit function in covariance modeling to generate χ^2 values. $2KL$ is given by

$$2KL = \text{Trace}\left(\Sigma^{-1}S\right) + \left(\mu - m\right)^T \Sigma^{-1}\left(\mu - m\right) - k - \ln\left(\frac{|S|}{|\Sigma|}\right) \tag{2}$$

where Σ and μ are the model distribution covariance matrices and mean, S and m the saturated model distribution covariance matrix and mean, and k the number of variables (i.e., the dimensionality of the distributions).

Note that $2 \cdot KL$ is identical to $\frac{\chi^2}{N}$ for complete data sets and, thus, the bias correction will not change RMSEA for complete data sets. To ensure an unbiased estimation of the average discrepancy for missingness with rate $1 - p$, we need to discount the sample size, N, by the proportion of observed cases, p. The resulting term $2pN \cdot KL$ has the same expectation as χ^2 for all rates of missingness if $S = \Sigma$, that is, if the model is correctly specified. So we obtain a bias correction if we replace our estimator of F with $2KL$. For $S \neq \Sigma$ in the population, $2pN \cdot KL - df$ has a non-zero expectation, that is, asymptotically proportional to pN. Adding the maximum operator to avoid negative values under the square root, we obtained the bias-corrected RMSEA:

$$
\begin{aligned}
\text{RMSEA}_{KL} &= \sqrt{\max\left(0, \frac{2 \cdot KL}{df} - \frac{1}{pN}\right)} \\
&= \sqrt{\max\left(0, \frac{2pN \cdot KL - df}{pNdf}\right)}
\end{aligned} \tag{3}
$$

Note that even though the fraction term is guaranteed to have a constant expectation for different missingness rates, the expression under the square root (which includes the max operator) may show an increasing expectation with p if the distribution of the fraction term includes negative results. However, the actual center of the distribution is not changed; the effect is just due to the negative values being moved to zero. The median and mode of the distribution stay constant.

Simulation Design

To demonstrate the effect of the proposed bias correction, we performed two simulations. The first uses a small SEM to demonstrate the effect best without interference, while the second uses a larger Latent Growth Curve Model taken from a real substantive study.

In the first simulation, the generating model was a bivariate normal distribution with zero mean, variance one, and a correlation which we varied in three steps, no correlation ($r = 0.0$), a typical weak association one may encounter in psychological research ($r = 0.125$), and a very strong correlation ($r = 0.9$) to investigate the behavior close to the limit. $N = 1000$ participants were simulated. We then analyzed the data with a model in which we fixed the covariance to be zero, using a saturated mean structure (this is akin to a covariance-only model). Note that this miss-specified model is identical to the null model if one would test against a zero covariance between the two manifest variables.

For a larger substantive example, we simulated misfit in a linear Latent Growth Curve Model (LGCM; McArdle, 1988). The linear LGCM models change across repeated observations by specifying an intercept and a slope component describing the overall level and change as well as individual differences in each. Our simulation was inspired by estimates that are typical for cognitive ageing research (Ghisletta et al., 2020). The model contained five manifest variables. We set the intercept mean to 20 and the intercept variance to 30, the slope mean to −4 and the slope variance to 5 and the residual error to 14. This corresponds to a moderate growth curve reliability (Figure 1). We used the same model for data generation, but to add some typical misspecification, we added a quadratic growth component with mean = −5 and variance = 5. $N = 9000$ participants were generated with this model.

In both simulations, we then created missing data by removing all values except for the first variable in some participants. The choice of participants with missing data was done either by an MCAR or MAR process: For the MCAR process, each participant had the same probability p_{miss} to miss all values but the first. For the MAR process, the probability $p_{miss;i}$ of the i th participant to show missingness depended on the first value x_i which was never missing and the overall missingness rate p_{miss} by a squashed sigmoid function:

$$p_{miss;i} = \frac{2 \cdot p_{miss}}{2 + e^{-x_i}} \tag{4}$$

Figure 1

Latent Growth Curve Model Used for Analyzing the Data in the Larger Simulation
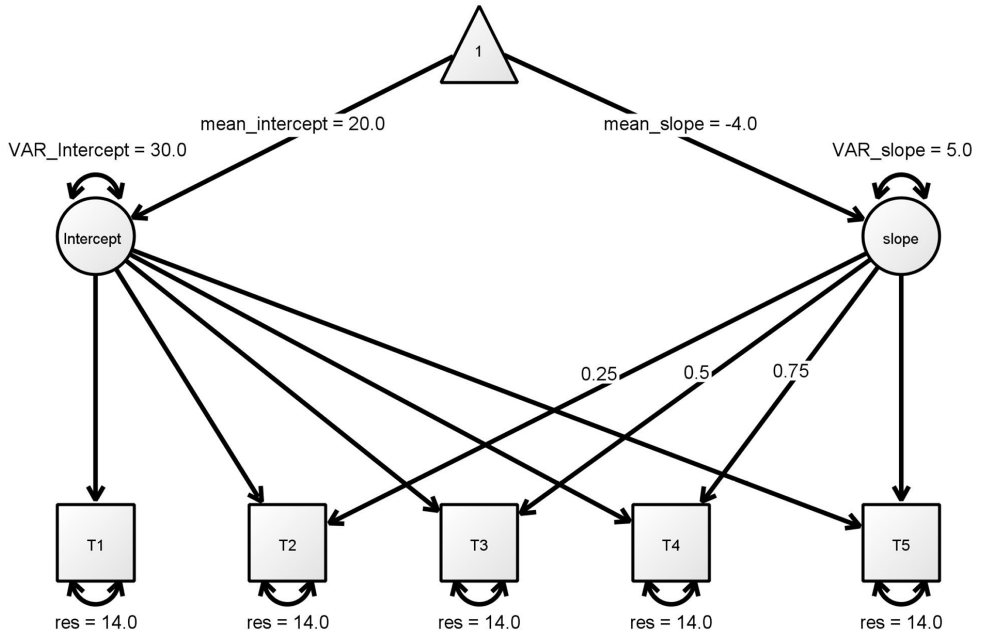
Note. Data was generate with the same model, but including a quadratic component with mean = −5 and variance = 5

Participants were chosen with this probability until exactly p_{miss} of the participant showed missingness. Throughout the simulations, a single data set was created in each condition, and missing data were then simulated in the same data set for each missingness scenario. This data was then analyzed with the misspecified model.

Overall, we simulated 1000 trials in both simulations and recorded the RMSEA with and without the correction. We used Onyx (von Oertzen, Brandmaier, & Tsang, 2015) and lavaan (Rosseel, 2012) for the simulations.

Results

Description of Plots

Results are presented visually by condition, with one plot per type of missingness (MCAR vs. MAR) and observed covariance between variables. The horizontal (x) axes of each plot represent the percent of data missing, while the vertical (y) axes of each plot

represent RMSEA. To aid in the interpretation of these simulations, we follow general guidelines in the literature to define RMSEA values less than 0.05 to indicate good fit, greater than 0.1 to indicate poor fit, and values in between 0.05 and 0.1 to indicate mediocre fit (Browne & Cudeck, 1993). A line for RMSEA of 0.05 is included where necessary.

Simulation results revealed consistent patterns for RMSEA calculations. RMSEA values were most strongly related to model misspecification (i.e., the strength of the correlation between variables that was constrained to zero in the misspecified model), but also showed effects of missingness patterns (MCAR vs. MAR). In the bivariate model, for both covariance values of 0.0 and 0.125, MAR and MCAR data generally yielded similar patterns.

Bivariate Normal Model Under Miss-Specification

For data MAR with a covariance of 0.125 between variables (Figure 2), the uncorrected RMSEA values scale linearly with missingness. For the data simulation conducted here, a missingness percentage of zero results in an RMSEA value of approximately 0.12, which indicates that the model has unacceptable fit. With a little more than 20% missingness, however, the RMSEA value drops below 0.1, indicating "poor" fit, and with about 70% missingness, it indicates "good" fit. *KL* corrected RMSEA values remain constant across all levels of missingness, with a slight upward bias starting at around 50% missingness.

Figure 2

Simulation Results for Data Generated With a Covariance of 0.125
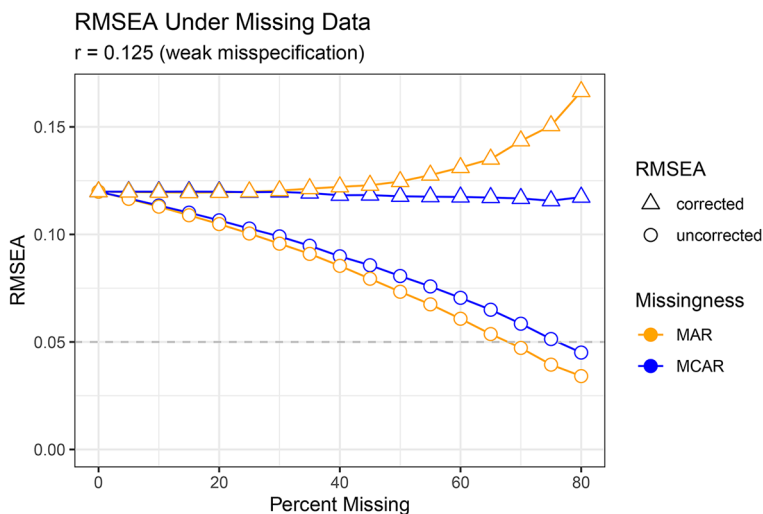
Note. RMSEA values are given for data completely missing at random (MCAR; blue) and data missing at random (MAR; orange). While *KL*-corrected RMSEA values (shown as triangles) remain mostly constant across

levels of missingness (except extreme conditions of missingness), uncorrected RMSEA (shown as circles) yield artificially improved model fit.

For data MCAR with a covariance value of 0.125 between variables (Figure 2), we see a similar pattern as was observed with the MAR data condition above. Uncorrected RMSEA values continue to decline linearly as missingness percentage increases. Values for the *KL* corrected RMSEA stay mostly constant in the simulation, as is expected from the mathematical results above.

As model misspecification increases, the resulting patterns observed become more pronounced. In the condition with a covariance of 0.9, we observe a similar trend, with uncorrected values scaling downward as the percentage of missingness increases, and *KL*-corrected values being almost completely independent of missingness. Figure 3 display results for data simulation when model covariance *cov* = 0.9 under MAR and MCAR data conditions, respectively. The mean of the *KL* corrected RMSEA increases slightly for high missingness values. This is due to the floor effect in which *KL*-corrected values are found using the maximum operator, because the sample error increases as missingness increases creates higher variance in the RMSEA values. This effect causes more RMSEA values to become zero by chance. In addition, high missingness also decreases the asymptotic properties of the parameter estimates, which also explains a part of the increase in the RMSEA values.

Figure 3

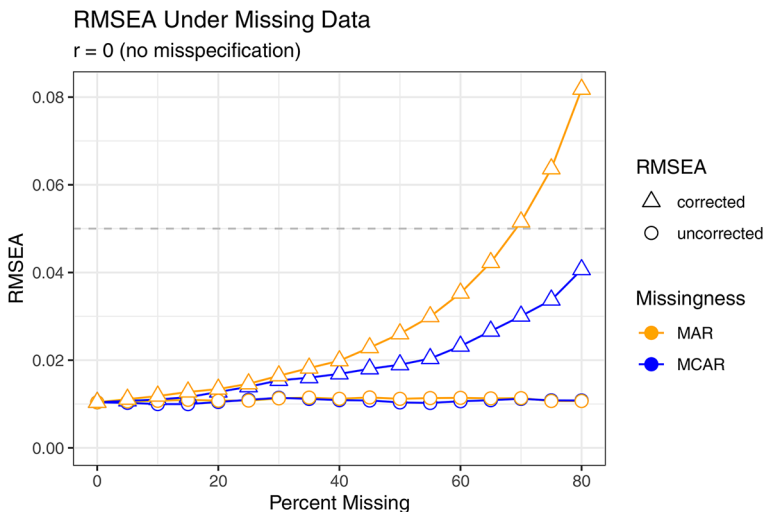*Simulation Results for Data Generated With a Covariance of 0.9*

Note. RMSEA values are given for data completely missing at random (MCAR; blue) and data missing at random (MAR; orange). Again, *KL*-corrected RMSEA values (shown as triangles) remain mostly constant across levels of missingness and uncorrected RMSEA values (shown as circles) yield artificially improved model fit.

Bivariate Normal Model With No Miss-Specification

For models with no model misspecification (i.e., those with no covariance between variables), we observe no decline in RMSEA values for any of the RMSEA values. Instead, uncorrected RMSEA values remain constant at a level of almost zero, even as missingness increase. Our proposed bias-corrected RMSEA, shows slight increases as missingness increases, that is, it has a slight tendency to overpessim in adjuging model fit. When missing percentage values reach 20%, both MAR and MCAR *KL* corrected values show an increase in terms of RMSEA value. See Figure 4 for these results.

Figure 4

Simulation Results for Data Generated With no Model Misspecification


Note. RMSEA values are given for data completely missing at random (MCAR; blue) and data missing at random (MAR; orange). As expected, results stay close to zero for both corrected (shown as triangles) and uncorrected (shown as circles) RMSEA values. Corrected RMSEA values start increasing for high missingness rates. When evaluating RMSEA values in this condition, keep note of the small scale of the *y*-axis.

Results in the Latent Growth Curve Model

Results of the LGCM simulation are shown in Figure 5. The corrected RMSEA value stays mostly constant up to a rate where 80% of the participants show missingness both under MAR and under MCAR. In contrast, both for MAR and MCAR the uncorrected RMSEA reduces from 0.32 initially to 0.15 for very high missingness rates.

PsychOpen GOLD

Figure 5

Simulation Results for Data Generated From a Latent Growth Curve Model With N = 9000 for Different Missing Ratios and Missing Mechanisms
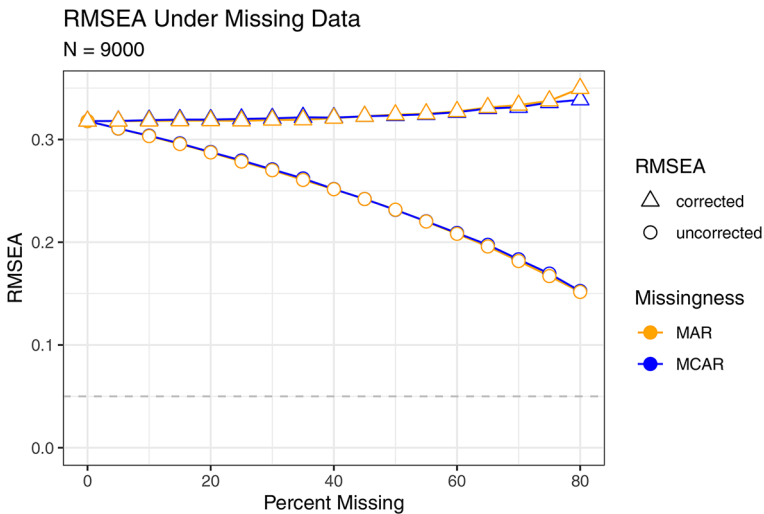
Note. Missing at random shown in orange, missing completely at random in blue. Uncorrected RMSEA values are shown as circles, corrected RMSEA values as triangles. Data were generated including a quadratic component, but analyzed using a linear Latent Growth Curve Model.

Results Summary

Among models with some misspecification, we see a a downward bias with the uncorrected RMSEA calculation and largely constant values for the *KL* corrected RMSEA. This shows up among both MCAR and MAR data simulations among models with both low (cov = 0.125) and high (cov = 0.9) model misspecification.

For the models having no misspecification, a pessimistic trend for *KL*-corrected values is observed when the amount of missing data is larger but not when it is small. In this case, the uncorrected RMSEA is unbiased and the *KL*-correction shows a slight pessimistic bias.

Discussion

The RMSEA, as defined by Steiger and Lind (1980) may be regarded as standardization of the χ^2 index of model misfit. Under the null hypothesis of no misfit, that is, the population distribution can be represented by the hypothesized model, the RMSEA follows a χ^2 distribution for any proportion of missing data (MAR and MCAR). Hence, we find in our simulation that the uncorrected RMSEA is constant in expectation for every

proportion of missing data when there is no model misspecification. We conclude that, under the null hypothesis, the uncorrected RMSEA behaves correct and according to our expectations. We also find that both the uncorrected RMSEA underestimates misfit and is thus an overoptimistic measure of goodness-of-fit when missing data are present. The overoptimism may severly bias our conclusions about correct model misspecification. In our simulations, the suggested bias correction decreases this overoptimism such that in many instances, models would be rejected as not well fitting even though the uncorrected RMSEA remains below typically used cut-off criteria for "good" model fit.

RMSEA was created to detect model misspecification. In fact, if we want a test whether a model is perfectly specified (to be precise, testing and possibly rejecting the null hypothesis of perfect fit), we could directly use the likelihood ratio test. The logic of the RMSEA is that models are never perfectly specified, and in fact would be useless if they were, since simplification is an integral part of statistical modeling. The RMSEA is a quantification of the degree of misspecification that allows us to work with "slightly" misspecified models that still are deemed to work as desired in most cases. We showed that with misspecification, the uncorrected RMSEA is no longer guaranteed to be constant over differing levels of missingness. This is evident from our simulations, where the uncorrected RMSEA decreases with higher proportions of missing data. We proposed a bias correction for the RMSEA adjusting the estimate of F via the Kullback-Leibler divergence such that RMSEA remains an appropriate fit index when data is only partially observed.

We conclude that, under the null hypothesis of no misfit, the uncorrected RMSEA is unbiased. We also find the uncorrected RMSEA underestimates misfit and is thus an overoptimistic measure of goodness-of-fit when missing data are present. The overoptimism may severely bias our conclusions about correct model misspecification. Because we rarely operate with correctly-specified models, we believe that for the advancement of knowledge, it is more beneficial to risk a slight pessimism in the rare case of correctly-specified models than overoptimism with miss-specified models when missing data are present. The mathematical derivation, backed up by the simulations, show that the suggested, show that the KL corrected RMSEA index provides a better indicator of model fit in the presence of missingness.

Limitations

Although the simulations strongly support the notion that KL-corrected RMSEA values are unaffected by missing data, further investigation is needed for more general models. The model used in our simulation was minimal and was designed as a proof of concept that demonstrates the behavior of our proposed corrections In most cases, such models are oversimplified and do not adequately represent more complicated datasets. Increasing the complexity of data simulation models and reassessing RMSEA behavior will allow

PsychOpen GOLD

for broader confirmation of the effectiveness of the *KL* correction as a modification for the RMSEA fit index.

A possible objection to implementing the *KL* correction for RMSEA lies in the realization that, in presence of missing data, corrected RMSEA values indicate worse model fit than uncorrected values. Though our rationale for an RMSEA correction is rooted in theory and practice, it may be hard to accept that RMSEA values that have been reported thus far have been impling artificially inflated goodness-of-fit for models estimated from missing data. Accepting the *KL* correction will pressure researchers to reassess model fit, and there may be resistance if *KL* corrected RMSEA values point to a different conclusion than uncorrected RMSEA values had indicated previously. Considering that the typically used cut-off values are based on rules-of-thumb derived from observed RMSEA values without using the correction, it may be a viable strategy to re-think the typically employed cut-off values. Changing the RMSEA calculation under missing data at large, with higher RMSEA values considered acceptable, may reduce resistance to such a change without necessarily accepting considerably worse models as acceptable but rather restore fairness among models derived from completely and partially observed data. RMSEA values may then become comparable across different levels of missingness, and preference will no longer be given to models where missing data rates are high.

Summary and Future Directions

The original formulation of RMSEA was derived under the assumption of complete observations. We found that a naive computation of the RMSEA under missing data results in overoptimistic model fit (i.e., RMSEA values that are too low). Hence, we must conclude that most reported model fits using RMSEA are too optimistic. We propose that the RMSEA can be corrected by replacing the χ^2 with an estimator based on the *KL* divergence in the RMSEA equation. This leads to an intuitive extension of the original idea behind the RMSEA to settings of missing data.

This paradigm can and should be extended to other fit indices as well, and be incorporated into the updating of SEM fit statistics to account for novel estimation approaches. Other fit indices are affected by missing data, and the *KL* correction should be explored and applied to CFI, TLI, and other indices. As the *KL* correction is an estimate, estimation accuracy needs to be considered when creating confidence intervals around *KL*-corrected RMSEA. Beyond these corrections, we must accept and further explore ways of evaluating SEM model fit that do not depend on the assumption of a complete saturated model, especially as we extend SEM into planned missingness, design and definition variables, and other important advances that take SEM from a covariance method to flexible GLM-based modeling.

Informing researchers about the benefits of an RMSEA correction and its ability to handle missing data may encourage the use of SEM in the realistic scenario of missing data. In order to do this, it is important to provide the scientific community easy access

to the correction. A feasible solution is writing a package for R which includes wrapper tools that will extract data from existing functions and calculate RMSEA corrected values. In this line, we attach a small R script to this article that derives *KL*-corrected RMSEA values for a given estimation result. The script is available in the Supplementary Materials. Ω nyx is currently the only software providing both uncorrected and corrected RMSEA. We hope that eventually other software developer teams will follow and integrate the *KL* correction into existing data software tools to encourage and support widespread use of corrected RMSEA estimates.

Supplementary Materials

For this article R-code for computing the corrected RMSEA is available via PsychArchives (for access see Index of Supplementary Materials below).

Index of Supplementary Materials

Fitzgerald, C. E., Estabrook, R., Martin, D. P., Brandmaier, A. M., & von Oerzen, T. (2021). *Supplementary materials to: Correcting the bias of the Root Mean Squared Error of Approximation under missing data* [Code]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.5135

References

Bauer, D. J. (2003). Estimating Multilevel Linear Models as Structural Equation Models. *Journal of Educational and Behavioral Statistics, 28*(2). http://www.jstor.org/stable/3701259

Bollen, K. (1989). *Structural equations with latent variables.* John Wiley.

Brandmaier, A. M., von Oertzen, T., McArdle, J. J., & Lindenberger, U. (2013). Structural Equation Model Trees. *Psychological Methods, 18*(1), 71-86. https://doi.org/10.1037/a0030001

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). SAGE.

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research, 38*(4), 529-569. https://doi.org/10.1207/s15327906mbr3804_5

Davey, A. (2005). Issues in evaluating model fit with missing data. *Structural Equation Modeling, 12*(4), 578-597. https://doi.org/10.1207/s15328007sem1204_4

PsychOpen GOLD

Enders, C. K., & Bandalos, D. (2001, July.). The relative performance of full information maximum
    likelihood estimation for missing data in structural equation models. *Structural Equation
    Modeling, 8*(3), 430-457. https://doi.org/10.1207/S15328007SEM0803_5

Finkbeiner, C. (1979). Estimation for the multiple factor model when data are missing.
    Psychometrika, 44(4), 409-420. https://doi.org/10.1007/BF02296204

Ghisletta, P., Mason, F., von Oertzen, T., Hertzog, C., Nilson, L.-G., & Lindenberger, U. (2020, June.).
    On the use of growth models to study normal cognitive aging. *International Journal of
    Behavioral Development, 1*(144), 88-96. https://doi.org/10.1177/0165025419851576

Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in MPlus and SAS. *Structural Equation
    Modeling, 16*(4), 676-701. https://doi.org/10.1080/10705510903206055

Grimm, K. J., Ram, N., & Estabrook, R. (2010). Nonlinear Structured Growth Mixture Models in M
    plus and OpenMx. *Multivariate Behavioral Research, 45*(6), 887-909.
    https://doi.org/10.1080/00273171.2010.531230

Heck, R. (2001). *Multilevel modeling in SEM.* In G. A.Marcoulides & R. E.Schumacker (Eds.), *New
    developments and techniques in structural equation modeling* (pp. 89–127). Lawrence Erlbaum
    Associates.

Hoyle, R. H. (Ed.).(2012). *Handbook of structural equation modeling.* Guilford Press.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis:
    Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.
    https://doi.org/10.1080/10705519909540118

Kullback, S., & Leibler, R. (1951). On infromation and sufficiency. *Annals of Mathematical Statistics,
    22*(1), 79-86. https://doi.org/10.1214/aoms/1177729694

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* John Wiley & Sons.

McArdle, J. J. (1988). *Dynamic but structural equation modeling of repeated measures data.* In J.
    R.Nesselroade & R. B.Cattell (Eds.), *Handbook of multivariate experiment psychology* (pp. 561–
    614). Plenum Press.

Miller, R. G. (1997). *Beyond ANOVA: Basics of applied statistics.* Chapman and Hall.

Mulaik, S. (1972). *Foundations of factor analysis* (2nd ed.). Chapman and Hall.

Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM
    algorithm. *Biometrics, 55*(2), 463-469. https://doi.org/10.1111/j.0006-341x.1999.00463.x

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

Rubin, D. B. (2002). Inference with missing data. *Biometrika, 63*(3), 581-592.
    https://doi.org/10.2307/2335739

Steiger, J. H., & Lind, J. M. (1980). *Statistically based tests for the number of factors.* In *Paper
    Presented at the Annual Meeting of the Psychometric Society.* Iowa City, IA, USA.

Tucker, L., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis.
    Psychometrika, 38(1), 1-10. https://doi.org/10.1007/BF02291170

von Oertzen, T., Brandmaier, A. M., & Tsang, S. (2015). Structural Equation Modeling with Onyx.
    Structural Equation Modeling: An Interdisciplinary Journal, 22, 148-161.

PsychOpen GOLD

Zhang, X., & Savalei, V. (2020). Examining the effect of missing data on RMSEA and CFI under
        normal theory full-information maximum likelihood. *Structural Equation Modeling: A
        Multidisciplinary Journal, 27*(2)–. https://doi.org/10.1080/10705511.2019.1642111