Secondary Publication



Gnambs, Timo; Sengewald, Marie-Ann

Meta-Analytic Structural Equation Modeling With Fallible Measurements

Date of secondary publication: 02.06.2023

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-596495

Primary publication

Gnambs, Timo; Sengewald, Marie-Ann: Meta-Analytic Structural Equation Modeling With Fallible Measurements. In: Zeitschrift für Psychologie. 231 (2023), 1, pp. 39-52.
DOI: 10.1027/2151-2604/a000511

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holder(s).

This document is made available under a Creative Commons license.



The license information is available online:
https://creativecommons.org/licenses/by/4.0/legalcode

Original Article

Meta-Analytic Structural Equation Modeling With Fallible Measurements

Timo Gnambs[1] and Marie-Ann Sengewald[1,2]

[1]Educational Measurement, Leibniz Institute for Educational Trajectories, Bamberg, Germany
[2]Institute for Psychology, University of Bamberg, Germany

Abstract: Meta-analytic structural equation modeling (MASEM) combines the strengths of meta-analysis with the flexibility of path models to address multivariate research questions using summary statistics. Because many research questions refer to latent constructs, measurement error can distort effect estimates in MASEMs if the unreliability of study variables is not properly acknowledged. Therefore, a comprehensive Monte Carlo simulation evaluated the impact of measurement error on MASEM results for different mediation models. These analyses showed that point estimates in MASEM were distorted by up to a third of the true effect, while confidence intervals exhibited undercoverage that were less than 10% in some situations. However, the use of adjustments for attenuation facilitated recovering largely undistorted point and interval estimates in MASEMs. These findings emphasize that MASEMs with fallible measurements can often yield highly distorted results. We encourage applied researchers to regularly adopt adjustment methods that account for attenuation in MASEMs.

Keywords: meta-analysis, structural equation modeling, mediation, reliability, measurement error

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

Meta-analytic structural equation models (MASEM; Cheung & Chan, 2005; Jak & Cheung, 2020; Ke et al., 2019; Lee & Beretvas, 2022) address multivariate hypotheses using sample statistics without requiring individual-level respondent data. In contrast to univariate meta-analyses that typically focus on bivariate effects (e.g., Cohen's *d*, *r*), MASEM allows for addressing more complex research questions involving three or more variables such as mediation or factor analyses (e.g., Marker et al., 2022; Schroeders et al., 2022; Wedderhoff et al., 2021). Because MASEM involves the pooling of correlation coefficients to examine the relations between variables of interest in a structural equation modeling (SEM) framework, established meta-analytic techniques such as artifact adjustments can be applied to study true score effects (see Wiernik & Dahlke, 2020, for an introduction). So far, adjustment techniques acknowledging, for example, measurement error are not regularly used in applications of MASEM (see Sheng et al., 2016). Some authors even suggested that respective adjustments do not provide substantial benefits for applied research but generally lead to comparable SEM results as MASEMs using unadjusted correlation coefficients (Michel et al., 2011). However, systematic methodological research on the impact of artifact adjustments in MASEM is currently missing.

Because of their importance in psychological, clinical, and epidemiological research, the present study focuses on mediation models with one or two intervening variables (see Figure 1). After summarizing prior work on biases resulting from unreliable measurements in single samples (see Savalei, 2019; Sengewald & Pohl, 2019), a comprehensive Monte Carlo study extends these results to MASEM and evaluates potential distortions when failing to account for measurement error in the involved variables. To increase the generalizability of our findings across different models and analytical choices, we study different types of mediation with either one or two mediators for different adjustment and MASEM methods.

Meta-Analytic Structural Equation Modeling

MASEM subsumes various statistical techniques that pool multiple correlation matrices to fit path or structural equation models to them. In the popular two-stage approach (TSSEM; Cheung & Chan, 2005), the $m \times m$ correlation matrices $\mathbf{R}_i$ with $i \in \{1, 2, \ldots, I\}$ for m variables observed in I individual studies are first pooled using a fixed- or random-effects model. To this end, the vectors $\mathbf{r}_i$ of lower diagonal correlations in $\mathbf{R}_i$ are subjected to a multivariate meta-analysis that decomposes the observed correlations in the ith study as

$$\mathbf{r}_i = \boldsymbol{\rho}_i + \mathbf{e}_i = \boldsymbol{\rho} + \mathbf{u}_i + \mathbf{e}_i \quad (1)$$

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

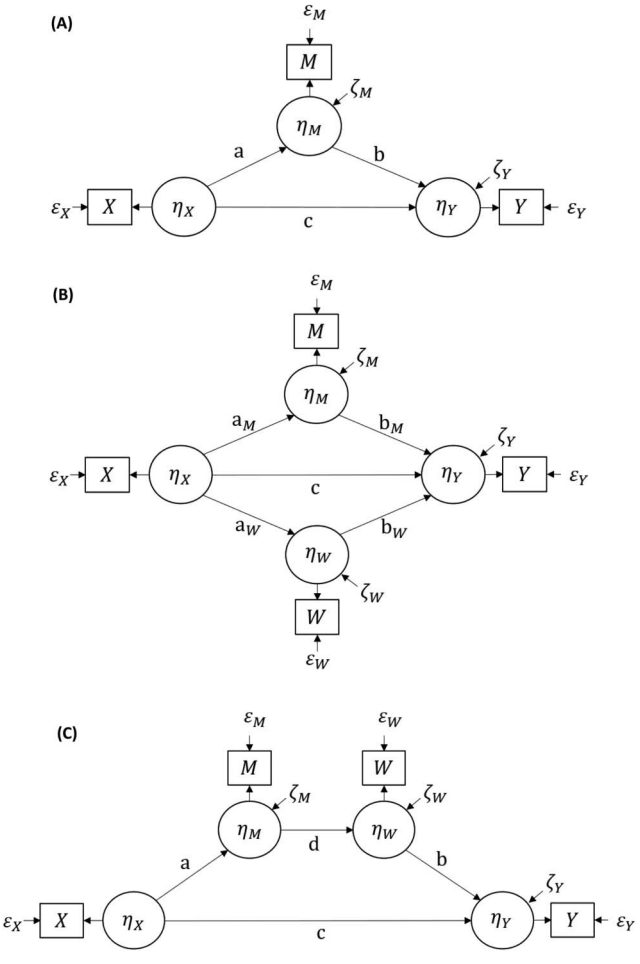


Figure 1. Mediation models with one or two intervening variables: (A) simple mediation, (B) parallel mediation, and (C) sequential mediation model.

where $\boldsymbol{\rho}_i$ are the population correlations in study i, $\boldsymbol{\rho}$ is the average population correlation across the I studies, $\mathbf{u}_i$ are the study-specific random effects (i.e., the deviations of $\boldsymbol{\rho}_i$ from $\boldsymbol{\rho}$), and $\mathbf{e}_i$ are the sampling errors (i.e., the deviations of $\mathbf{r}_i$ from $\boldsymbol{\rho}_i$). If $\mathbf{u}_i = 0$ for all i, the model reduces to a fixed-effects specification. The $\mathbf{r}_i$ are typically assumed to follow a multivariate normal distribution $\mathbf{r}_i \sim N(\boldsymbol{\rho}, \mathbf{T}^2 + \mathbf{V}_i)$ with $\mathbf{T}^2$ and $\mathbf{V}_i$ representing the $m \times m$ between-study (co) variance matrices for $\mathbf{u}_i$ and within-study (co)variance matrices for $\mathbf{e}_i$, respectively. In practice, $\mathbf{V}_i$ is rarely known but calculated from the observed sample statistics (see Cheung & Chan, 2004). Then, $\widehat{\mathbf{V}}_i$ can be used to jointly estimate $\widehat{\mathbf{T}}^2$ and $\widehat{\boldsymbol{\rho}}$ in a SEM framework using full information maximum likelihood estimation (Cheung, 2013). In the second stage of TSSEM, the hypothesized SEM is fitted to the pooled correlations $\widehat{\boldsymbol{\rho}}$ using a weighted least squares estimator with $\widehat{\mathbf{V}}_\rho^{-1}$, the inverse of the asymptotic (co) variances of $\widehat{\boldsymbol{\rho}}$, as the respective weight matrix (Cheung & Chan, 2005). An advantage of the multivariate approach is that the second stage does not require specifying an arbitrary sample size for the estimation of the SEM but directly incorporates the precision of the pooled estimates in the form of $\widehat{\mathbf{V}}_\rho$.

In addition to the described TSSEM approach, several other MASEM techniques have been proposed (e.g., Jak & Cheung, 2020; Ke et al., 2019; Lee & Beretvas, 2022; Viswesvaran & Ones, 1995). For example, Jak and Cheung (2020) recently introduced a one-stage method (OSMASEM) that does not require estimating a pooled correlation matrix as an intermediate step. Rather, it restricts the pooled correlations in the multivariate meta-analysis to the model implied correlations given by the specified SEM (e.g., regression weights, covariances). In this way, OSMASEM can directly estimate the SEM parameters without having to estimate the pooled correlations first. Although OSMASEM and TSSEM result in highly comparable results (Jak & Cheung, 2022), OSMASEM is more versatile and can also include study-level moderators for each SEM parameter.

Adjustments for Measurement Error in MASEM

Psychological constructs can rarely be measured without error (Gnambs, 2015). Rather, measurement error variance distorts sample statistics including the correlation coefficient (see Wiernik & Dahlke, 2020). Therefore, various statistical procedures have been developed to counteract the distorting effects. In the context of MASEM, at least three approaches seem worthwhile to consider. On the one hand, MASEM could be conducted with latent correlations that are free from measurement error. Rather than pooling observed sample correlations, MASEM might make use of true score correlations from each sample, for example, estimated with respective latent variable SEMs. Then, the resulting meta-analytic path models should not be distorted by measurement error variance. However, this approach is often limited to specific applications (see, for example, Brunner et al., 2022) because latent correlation matrices are frequently not reported in primary studies. Alternatively, latent variables could be directly modeled in the MASEM. However, this would require access to item-level statistics for all involved constructs, and item-level correlation matrices are typically not reported in primary studies. Therefore, the specification of latent variables in MASEM, so far, is limited to factor analytic research on single instruments such as the *General Health Questionnaire* (Gnambs & Staufenbiel, 2018), *Rosenberg's Self-Esteem Scale* (Gnambs et al., 2018), or *Toronto Alexithymia Scale* (Schroeders et al., 2022). But it seems rarely conceivable to identify item-level correlation matrices for multiple instruments that would be required for respective multiconstruct MASEMs. Consequently, a third approach that

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

adjusts observed sample correlations for attenuation due to measurement error seems most applicable for meta-analyses based on the available information.

Following classical test theory, the degree of attenuation of the observed correlations $\mathbf{r}_i$ is a multiplicative function of the square roots of the reliabilities, $\mathbf{g}_{i,x}$ and $\mathbf{g}_{i,y}$, of the involved variables x and y such that $\mathbf{r}_i = (\mathbf{g}_{i,x} \odot \mathbf{g}_{i,y})^{\odot(1/2)} \odot \mathbf{r}_i^{(g)}$. Consequently, the disattenuated correlations $\mathbf{r}_i^{(g)}$ and their sampling (co)variances $\mathbf{V}_i^{(g)}$ can be derived as

$$\mathbf{r}_i^{(g)} = \left(\mathbf{g}_{i,x} \odot \mathbf{g}_{i,y}\right)^{\odot(-1/2)} \odot \mathbf{r}_i \quad (2)$$

$$\mathbf{V}_i^{(g)} = \left(\mathbf{g}_{i,x} \odot \mathbf{g}_{i,y}\right)^{\odot(-1/2)} \otimes \left(\mathbf{g}_{i,x} \odot \mathbf{g}_{i,y}\right)^{\odot(-1/2)\mathrm{T}} \odot \mathbf{V}_i. \quad (3)$$

In Equations 2 and 3, $\odot$, $\odot(-1)$, and $\odot(1/2)$ denote the Hadamard (element-wise) product, the Hadamard inverse, and the Hadamard square root, respectively, while $\otimes$ gives the Kronecker product. If the reliabilities are known for the m variables in the I studies, each correlation coefficient $\mathbf{r}_i$ can be individually adjusted for unreliability. Then, MASEM can pool the $\mathbf{r}_i^{(g)}$ as outlined above into the average disattentuated population correlations $\widehat{\boldsymbol{\rho}}^{(g)}$, thus leading to estimates of true score effects in the hypothesized SEM. Unfortunately, poor reporting practices in primary studies often lead to unknown reliabilities for selected (or even most) correlations. In this case, the available reliabilities are first averaged into $\boldsymbol{\gamma}_{i,X}$ and $\boldsymbol{\gamma}_{i,Y}$, either by taking their mean values across studies or adopting more sophisticated techniques of reliability generalization (e.g., Scherer & Teo, 2020) or imputed from other sources (e.g., Gnambs, 2014, 2015). Then, TSSEM adjusts the average population correlations $\widehat{\boldsymbol{\rho}}$ and their asymptotic (co)variances $\widehat{\mathbf{V}}_\rho$ as in Equations 2 and 3 using the $\boldsymbol{\gamma}_i$, thus giving the disattenuated average population correlations $\widehat{\boldsymbol{\rho}}^{(g)}$ and their sampling variances $\widehat{\mathbf{V}}_\rho^{(g)}$. For OSMASEM, these adjustments are slightly more complicated because the model-implied correlations need to be adjusted for unreliability (see Appendix A). Generally, adjustments using artifact distributions should lead to comparable results as individual adjustments as long as systematically missing reliabilities[1] do not distort the pooled reliabilities in the artifact-adjustment approach. However, a potentially serious disadvantage of these adjustments is their reliance on reliability estimates reported in the primary studies. If these represent biased indicators of a measure's measurement error variance, for example, because important assumptions were violated (e.g., essentially τ-equivalent measurements for coefficient alpha; Dunn et al., 2014), these reliabilities might lead to overadjustments or underadjustments of the observed correlations and, consequently, to distorted path estimates in the MASEM (see Rhemtulla et al., 2020).

The use of artifact adjustments in MASEM is, so far, not standard in meta-analytic practice. While MASEMs in industrial-organizational psychology following the tradition of psychometric meta-analysis (Viswesvaran & Ones, 1995) often apply adjustments for various measurement artifacts, it is less common in other domains such as educational or health psychology that more frequently adopt the described TSSEM or OSMASEM. A review of 160 MASEMs published between 1995 and 2015 (Sheng et al., 2016) reported that about three-quarters of them applied some form of adjustment for unreliability. Surprisingly, some research suggested that these adjustments might be superfluous because substantial conclusions are generally not affected by them (Michel et al., 2011). Reanalyses of five published MASEMs with and without adjustments for attenuation showed negligible changes in SEM path estimates between $M(\Delta\beta)$ = .01 and .05 and comparable fit indices. Therefore, these authors argued that respective adjustments are inconsequential for MASEMs as long as the included variables exhibit satisfactory levels of reliability (i.e., about .70–.90). However, these results are at odds with related work on meta-analyses of bivariate effects (e.g., Le et al., 2016) and also a plethora of methodological studies on measurement error in primary research (e.g., Aiken & West, 1991; Cole & Preacher, 2014; Fritz et al., 2016; Savalei, 2019; Sengewald & Pohl, 2019; Steiner et al., 2011; Westfall & Yarkoni, 2016).

The Impact of Measurement Error in Primary Mediation Research

Growing methodological evidence shows the consequences of measurement error on regression results for different contexts. Recently, Savalei (2019) derived the impact of measurement error in a single mediation model as shown in Figure 1A, with a latent treatment variable η_X, a latent outcome η_Y, and a latent mediator η_M. Based on this model, the process by which the treatment affects the outcome can be investigated by separating the direct effect c and the indirect effect that is equal to the product ab, using the two regression equations,

$$\eta_M = \mathrm{a} \cdot \eta_X + \zeta_M, \quad (4)$$

[1] The pooled reliability estimates might be systematically distorted, if studies administering, for example, less reliable measures have a higher propensity of reporting omissions (i.e., do not inform about the reliability). Then, the pooled reliabilities will be overestimated and, consequently, the adjusted correlations will lead to underestimations of the SEM true score effects.

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

$$\eta_Y = b \cdot \eta_M + c \cdot \eta_Y + \zeta_Y. \quad (5)$$

However, using fallible measures X, Y, and M in a path analysis and ignoring measurement error ε_X, ε_Y, and ε_M can affect the efficacy and accuracy of the effect estimates, depending on the specific parameter constellation.

According to Savalei (2019), measurement error in the outcome Y simply adds to the error term ζ_Y and decreases the efficacy of effect estimates, but not the accuracy (see also Aiken & West, 1991; Wiernik & Dahlke, 2020). Thus, unstandardized regression coefficients will not be affected, but statistical power and standardized effect estimates decrease, in relation to the reliability γ_Y of the dependent variable Y.

As opposed, when predictor variables are fallible, this will also affect the accuracy of effect estimates (e.g., Aiken & West, 1991; Sengewald & Pohl, 2019; Steiner et al., 2011). Savalei (2019) derived how the regression coefficients with fallible variables (i.e., a^*, b^*, c^*) change, when ignoring measurement error. For simplicity, we assume standardized variables with variances $\mathrm{Var}(\eta_X) = \mathrm{Var}(\eta_M) = \mathrm{Var}(\eta_Y) = 1$. With fallible variables, the *a*-path always decreases in relation to the reliability γ_X of the predictor variable X, that is,

$$a^* = a \cdot \gamma_X. \quad (6)$$

Instead, partial regression coefficients can show bias in different directions, depending on the relation of all variables. This holds for the *b*- and *c*-path, that are, in case of fallible variables,

$$b^* = \gamma_M \cdot (b + a \cdot c - \gamma_X \cdot a \cdot (a \cdot b + c)) \cdot \left(1 - \gamma_X \cdot \gamma_M \cdot a^2\right)^{-1}, \quad (7)$$

$$c^* = \gamma_X \cdot (c + a \cdot b - \gamma_M \cdot a \cdot (b + a \cdot c)) \cdot \left(1 - \gamma_X \cdot \gamma_M \cdot a^2\right)^{-1}. \quad (8)$$

Thus, the reliabilities γ_X and γ_M of both predictors X and M and the specific constellation of all variables are relevant for obtaining the consequences of measurement error in the coefficients. Differences to the true effects (i.e., $b^* - b$ and $c^* - c$) will occur, in relation to the predictor's own reliability and impact on the outcome (i.e., b with γ_M and c with γ_X, respectively), but also depend on the reliability of the other predictor and its partial impact on the respective coefficient (i.e., for the *b*-path: ac with γ_X and for the *c*-path: ab with γ_M). In addition, the *a*-path has the potential to amplify bias in relation to the reliabilities, as $(1 - \gamma_X \cdot \gamma_M \cdot a^2)^{-1}$ shows (see also Sengewald & Pohl, 2019). Especially for large effects, low reliabilities, and substantial *a*-paths, the consequences of measurement error can be serious but depend on the partial relations (i.e., whether the predictors increase or reduce each other's impact).

Yet, these implications for the consequences of measurement error only hold for a single mediation model. When model complexity increases, for instance, when more variables are involved, the impact of measurement error can be even more serious and less tractable because additional variables can compensate but also amplify the bias (e.g., Cole & Preacher, 2014; Fritz et al., 2016; Savalei, 2019; Sengewald & Pohl, 2019). Corrections for measurement error can help to prevent for the bias as well as for its amplification due to other variables.

Objectives of the Present Study

Measurement error is a pervasive problem when studying psychological phenomena referring to latent constructs. Fallible measures can distort, among others, path and variance estimates in SEM and, thus, lead to invalid conclusions (e.g., Aiken & West, 1991; Cole & Preacher, 2014; Sengewald & Pohl, 2019; Westfall & Yarkoni, 2016). In mediation models, bias depends on a complex interplay between the reliability of the intervening variables and the size of the involved effects (Savalei, 2019). Despite the well-established consequences of measurement error in primary research, adjustments for unreliability are currently used rather unsystematically in MASEM (Sheng et al., 2016). Some authors even suggest that they might not provide notable benefits (Michel et al., 2011). Therefore, the present simulation study evaluated the consequences of ignoring measurement error in different MASEMs with one or two mediating variables (see Figure 1). We demonstrate how adjustments for unreliability that were initially developed for bivariate meta-analysis (Wiernik & Dahlke, 2020) allow recovering unbiased parameter estimates in MASEM and improve the precision of the estimated effects.

Method

Evaluated SEMs and MASEM Procedures

The simulation evaluated 18 different mediation models (see Table 1). The simple mediation models examined the indirect effect of an independent variable η_X on an outcome η_Y via a single mediator η_M, while the parallel mediation models specified two conditionally independent mediators η_M and η_W, thus modeling additive indirect effects (see Figure 1). Moreover, the sequential mediation models used two mediators η_M and η_W in consecutive order. All models assumed a zero effect of η_X on η_Y (*c*-path), thus reflecting full mediation. Based on the true models, we evaluated the impact of using imperfect measures (i.e., $X = \eta_X + \varepsilon_X$, $M = \eta_M + \varepsilon_M$, $W = \eta_W + \varepsilon_W$, $Y = \eta_Y + \varepsilon_Y$) on all path coefficients. In the special case of full mediation, the indirect effect will not be affected by the *c*-path, and thus, bias due to measurement error will be in the same direction (i.e.,

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

decreases the indirect effect toward zero). Accordingly, we regard the impact of measurement error in indirect effect estimates without the partial impact of the direct effect on the respective coefficient (i.e., for the *b*-path: ac with γ_X). To demonstrate the impact of measurement error as discussed before (Savalei, 2019; Sengewald & Pohl, 2019), we independently varied three design factors, that is, the size of (a) the average reliabilities of the mediators in the samples (.70 vs. .90), (b) the total indirect effect (.01–.04 vs. .09–.40), and (c) the *a*-paths (.15 vs. .60). For each factor, we considered a small or high value, while keeping the other factors constant at a medium size. For each of the 18 models, six MASEMs were estimated that differed regarding the adjustment method (without reliability adjustments vs. with adjustments of individual correlations vs. adjustments of pooled correlations) and the MASEM type (TSSEM vs. OSMASEM). This resulted in a total of 108 simulation conditions.

Data Generation and Simulation Procedure

The simulation was estimated in *R* (version 4.1.2; R Core Team, 2021) using *metaSEM* (version 1.2.5.1; Cheung, 2015) with *OpenMx* (version 2.20.6; Neale et al., 2016) and included 1,000 replications of each condition. Replications that resulted in nonpositive definite correlation matrices or for which a MASEM returned an improper solution were discarded and replaced with a valid case. The simulation of a given replication proceeded as follows:

First, the population values of the mediation model (see Table 1) were used to calculate the average population correlation matrix as $\mathbf{P} = (\mathbf{I} - \mathbf{A}) \cdot \mathbf{S} \cdot (\mathbf{I} - \mathbf{A})^{'}$ with **I**, **S**, and **A** representing a $m \times m$ identity matrix, diagonal matrix with the (residual) variances, and square matrix with the directed paths of the mediation model, respectively (see Appendix B for details). Consequently, $\boldsymbol{\rho}$ denotes the row vector including the lower diagonal elements of **P**. The random variances τ^2 for $\boldsymbol{\rho}$ as given by the diagonal matrix $\mathbf{T}^2$ were constrained, thus adopting a common between-study heterogeneity for all correlations. For a given replication, τ was randomly drawn from a half-normal distribution, $\tau \sim$ half-$N(0, 0.20)$, that was truncated at 0.50 to prevent excessively large heterogeneity estimates. This resulted in random variances that are typically observed in psychological meta-analyses (see Van Erp et al., 2017). Finally, the number of samples included in a meta-analysis was randomly drawn from a uniform distribution, $I \sim U(5, 50)$, to cover typical conditions in MASEM (see Table 2).

Second, for each simulated sample *i*, the respective sample size was randomly drawn from a gamma distribution, $n_i \sim \Gamma(1.3, 250)$, to reflect the positively skewed distributions frequently observed in psychological meta-analyses (cf. Sánchez-Meca & Marín-Martínez, 2008). To avoid extreme sample sizes, the distribution was truncated at 50 and 900, respectively. Reliability generalization studies showed that, on average, reliabilities of psychological measures fall around .80 and rarely drop below .65 (Gnambs, 2015). Therefore, in each sample, the $1 \times m$ row vector of reliabilities $\mathbf{g}_i$ for the included variables was randomly drawn from uniform distributions, $\mathbf{g}_i \sim U(.65, .95)$. In conditions that manipulated the average reliability of the mediators in the samples (see Table 1), we used draws from $g_i \sim U(.65, .75)$ or $g_i \sim U(.85, .95)$ to represent low or high reliabilities, respectively. Note that the generated reliabilities refer to the measurement precision in the specific sample. The approach of random draws allows for generalizing to different reliabilities that might occur in specific samples, but does not generalize to the measurement precision in an unobserved population.

Third, the sample correlations between the *m* variables were generated by first drawing the population correlations $\boldsymbol{\rho}_i$ for sample *i* from a multivariate normal distribution $N(\boldsymbol{\rho}, T_i^2)$. Then, the vector of disattenuated

Table 1. Evaluated MASEM models

	SEM paths			Average reliabilities			
Model	a	b	Indirect	*X*	*M*	*W*	*Y*
Low reliability							
Simple mediation	.30	.30	.09	.80	.70	—	.80
Parallel mediation	.30	.30	.18	.80	.70	.70	.80
Sequential mediation	.30	.30	.04	.80	.70	.70	.80
High reliability							
Simple mediation	.30	.30	.09	.80	.90	—	.80
Parallel mediation	.30	.30	.18	.80	.90	.90	.80
Sequential mediation	.30	.30	.04	.80	.90	.90	.80
Small indirect effect							
Simple mediation	.30	.08	.02	.80	.80	—	.80
Parallel mediation	.30	.08	.04	.80	.80	.80	.80
Sequential mediation	.30	.08	.01	.80	.80	.80	.80
Large indirect effect							
Simple mediation	.30	.68	.20	.80	.80	—	.80
Parallel mediation	.30	.68	.40	.80	.80	.80	.80
Sequential mediation	.30	.68	.09	.80	.80	.80	.80
Small a-path							
Simple mediation	.15	.60	.09	.80	.80	—	.80
Parallel mediation	.15	.60	.18	.80	.80	.80	.80
Sequential mediation	.15	.60	.04	.80	.80	.80	.80
Large a-path							
Simple mediation	.60	.15	.09	.80	.80	—	.80
Parallel mediation	.60	.15	.18	.80	.80	.80	.80
Sequential mediation	.60	.15	.04	.80	.80	.80	.80

Note. In all models, the direct effect of η_X on η_Y (c-path) was set to 0. For the parallel mediation model, $a = a_M = a_W$ and $b = b_M = b_W$ was used, while for the sequential mediation, we used d = .44. Indirect = total indirect effect of η_X on η_Y.

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

correlations r_i^g for sample i was calculated based on random draws of size n_i from a multivariate normal distribution $N(0, \boldsymbol{\rho}_i)$. Finally, the attenuated correlations $\boldsymbol{r}_i$ were created by multiplying $\boldsymbol{r}_i^g$ with the lower diagonal of $\boldsymbol{G}_i = \mathbf{g}_i^{0.5} \cdot (\mathbf{g}_i^{0.5})^T$. This resulted in the sample correlation matrix $\mathbf{R}_i$ being analyzed in the MASEMs.

Steps 2 and 3 were repeated for each of the I samples included in a given meta-analysis. Then, the I inverse variance weighted correlation matrices were pooled with a diagonal random effects structure using maximum likelihood estimation, either using a two-stage (Cheung & Chan, 2005) or a one-stage approach (Jak & Cheung, 2020). The respective asymptotic sampling (co)variances for $\mathbf{R}_i$ were derived following Cheung and Chan (2004). Depending on the examined condition (see Table 1), either each sample correlation matrix $\mathbf{R}_i$ or the pooled correlation matrix $\widehat{\mathbf{P}}$ with their respective (co)variances were adjusted for unreliability. The precision of the parameter estimates in the mediation model was quantified using 95% likelihood-based confidence intervals that are superior for parameters with non-normal sampling distributions such as indirect effects (Cheung, 2009; Neale & Miller, 1997).

Performance Criteria

The performances of the different estimators were studied using the population bias, the root mean squared error (RMSE), and the coverage rates for each parameter. The raw population bias is the average difference between the estimated parameter for each condition and the data-generating true value across all replications, while the percent bias is the raw bias divided by its true value times 100. Percent biases less than 5% are frequently considered negligible (Hoogland & Boomsma, 1998). The RMSE is the square root of the average squared difference between each estimated parameter and its true value. Because it is a mixture of the bias and efficiency of an estimator, a biased method can be preferred if it is more accurate and, thus, on average, is closer to the true value. Finally, the coverage rates of the 95% confidence intervals represent the percentage of replications for which the true effect fell within the interval. Coverage rates for which the nominal coverage probability of 95% fell within approximately two *SD*s around the observed rates were deemed acceptable. The Monte Carlo error for each performance criterion was estimated using the jackknife method (Koehler et al., 2009).

Transparency and Openness

The computer code, simulated data, and full results are provided in the online material available in PsychArchives at https://doi.org/10.23668/psycharchives.8537.

This study's design and its analyses were not preregistered.

Table 2. Settings for simulations

Parameter	Distribution
Reliabilities in sample i ($\mathbf{g}_i$)	$\mathbf{g}_i \sim U(.65, .75)$ for $M(\mathbf{g}) = .70$
	$\mathbf{g}_i \sim U(.65, .95)$ for $M(\mathbf{g}) = .80$
	$\mathbf{g}_i \sim U(.85, .95)$ for $M(\mathbf{g}) = .90$
Random effect (τ_ρ)	$\tau_\rho \sim \min\{\text{half-}N(0, 0.20), 0.50\}$
	$T^2 = \text{diag}(\tau_\rho)$
Number of samples (I)	$I \sim U(5, 50)$
Sample size (n)	$n \sim \min\{\Gamma(1.3, 250) + 50, 900\}$
Population correlations in sample i ($\boldsymbol{\rho}_i$)	$\boldsymbol{\rho}_i \sim N(\boldsymbol{\rho}, \mathbf{c})$
Disattenuated correlations in sample i (r_i^g)	$r_i^g = Cor\,[{}_{n_i}\mathbf{X}_p \sim N(0, \boldsymbol{\rho}_i)]$
Attenuated correlations in sample i (r_i)	$r_i = r_i^g \cdot \mathbf{g}_i^{0.5} \cdot (\mathbf{g}_i^{0.5})^T$

Note. half-N = half-normal distribution; Γ = gamma distribution with shape and scale parameters, U = uniform distribution; $\boldsymbol{\rho}$ = average population correlations, ${}_{n_i}\mathbf{X}_p$ = variable scores for n_i respondents on p variables.

Results

The simulation results are summarized separately for the different performance criteria. Because the different adjustment methods and MASEM types yielded highly comparable results, only the findings for TSSEM with adjustments of the pooled correlations will be reported in detail. Full results are available in our online material present in PsychArchives.

Convergence Rates

The TSSEMs converged at both stages for all replications and simulation conditions. In contrast, 0.0%–0.6% (*Mdn* = 0.1%) OSMASEMs exhibited convergence problems and had to be replaced with a valid run. The convergence rates were not systematically related to the adjustment method (none vs. individual vs. pooled correlations) or specific simulation conditions (see online material). Thus, in general, the examined MASEM methods and adjustment techniques did not introduce pronounced estimation problems.

Biases in Parameter Estimates

Measurement error resulted in substantial parameter bias for all examined models (see Tables 3 and 4). Percent biases in the *a*- and *b*-paths varied across the different mediation models and simulation conditions between −17% and −28% (*Mdn* = −22%) and −18% and −33% (*Mdn* = −24%), respectively, indicating a substantial underestimation of the true path coefficients. The distortions in the respective indirect effects, integrating the distortion of both paths, were even more pronounced and fell between −32% and −63% (*Mdn* = −44%). As expected from prior research on simple mediation models in primary studies (Savalei, 2019), the biases in indirect effects were larger for less as compared to more reliable measurements (49% vs. 32%), for larger as

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

compared to smaller indirect effects (42% vs. 38%), and larger as compared to smaller *a*-paths (46% vs. 40%). Thus, depending on the examined population model, measurement error had an unequally strong impact. To some degree, similar patterns were also observed for the parallel and sequential mediation models showing especially larger percent biases in indirect effects for less as compared to more reliable measurements (−49%/−63% vs. −33%/−41%).

After adjusting the pooled correlations for unreliability, the percent biases in all path coefficients substantially dropped and fell for the *a*- and *b*-paths across the different simulation conditions at medians of −3% (Min = −6%, Max = 2%) and −5% (Min = −17%, Max = 0%), respectively. Although few estimates exhibited percent biases falling strictly below 5% which is often considered negligible (Hoogland & Boomsma, 1998), most of them closely bordered this threshold (see Table 3). Only in conditions involving rather large correlations between selected variables (e.g., for large indirect effects), biases in the *b*-paths remained slightly elevated. Percent biases for the indirect effects were also substantially smaller in all conditions after adjustments for unreliability but generally slightly larger as compared to the direct effects (*Mdn* = −10%, Min = −2%, Max = −15%).

The percent biases of the unadjusted effects translated into raw biases between −.22 and −.01 (*Mdn* = −.06) and, thus, spanned a rather broad range from inconsequential to substantial (see Table 4). In contrast, after adjustments for unreliability, the remaining raw biases fell at a median of −.01 across the simulation conditions (Min = −.11, Max = .01). Again, the remaining biases were largest for conditions with highly correlated variables, for example, reflecting large indirect effects. The raw bias shows that also direct effects estimates were distorted without adjustments in our generated data with a full mediation (see Table 4). For this case, biases in percentage are not defined, as bias does not depend on the true *c*-path, that is zero, but on the partial impact of the fallible mediators and the fallible outcome. In line with the indirect effect estimates, the direct effect estimates were close to the true effects when reliability adjustments were implemented.

Root Mean Squared Error of Parameter Estimates

The RMSE mirrored the results of the bias analyses and emphasized the benefits of adjustments for unreliability (see online material). While the median RMSEs of the unadjusted *a*- and *b*-paths across mediation models and simulation conditions fell at .08 (Min = .04, Max = .15) and .08 (Min = .04, Max = .23), respectively, they reduced to .05 (Min = .04, Max = .07) and .06 (Min = .04, Max = .14) after adjusting the pooled correlations for unreliability. For the indirect effects, a similar pattern was observed with larger RMSEs for the unadjusted (*Mdn* = .04, Min = .01, Max = .19) as compared to the adjusted effects (*Mdn* = .02, Min = .01, Max = .08). The benefits of adjustments for unreliability seemed more pronounced for larger effects such as conditions with large *a*-paths as compared to population models with smaller effects. For example, for the parallel mediation model, the simulation condition with a large indirect effect resulted in RMSEs of .19 and .08 for the unadjusted and adjusted models, respectively, whereas the respective estimates were .02 and .02 for a small indirect effect. Detailed results are given in the online material present in PsychArchives. Together, these results show that adjustments for unreliability did not lead to substantially less efficient estimators that might offset their reduced bias.

Coverage Rates of Confidence Intervals

The 95% confidence intervals of the parameter estimates were substantially distorted by unaccounted measurement error (see Table 5). The respective coverage rates for the *a*- and *b*-paths varied across the different mediation models and simulation conditions between 2% and 64% (*Mdn* = 27%) and 0% and 68% (*Mdn* = 29%), respectively, indicating a substantial undercoverage. The respective distortions for the indirect effects were even more pronounced and fell between 0% and 56% (*Mdn* = 10%). In line with the bias analyses, the undercoverage in indirect effects of the simple mediation model was larger for less as compared to more reliable measurements (8% vs. 26%) and for larger as compared to smaller indirect effects (4% vs. 59%). Thus, depending on the examined population model, measurement error affected the confidence intervals to a different degree. After adjusting the pooled correlations for unreliability, the coverage rates approached the nominal level and fell for the *a*- and *b*-paths across the different simulation conditions at medians of 93% (Min = 81%, Max = 95%) and 93% (Min = 25%, Max = 95%), respectively. Although few estimates resulted in coverage rates of exactly 95%, most of them were in close range (see Table 5). Only in conditions involving rather large correlations between selected variables (e.g., for large indirect effects), coverage rates for the *b*-paths remained poor, even after adjustments for unreliability. Coverage rates for the indirect effects also improved in all conditions after adjustments for unreliability but remained slightly below the nominal level (*Mdn* = 88%, Min = 57%, Max = 95%).

Discussion

Most research questions in psychology and related domains refer to unobservable constructs that cannot be measured without error (see Gnambs, 2014, 2015). Because analyses of fallible measurements can lead to seriously distorted estimates of true score effects (e.g., Savalei, 2019; Sengewald & Pohl, 2019), hypotheses

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

Table 3. Percent biases in regression coefficients with and without adjustments for unreliability

	Simple mediation		Parallel mediation		Sequential mediation	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Low reliability						
a^*_M	−28%	**−3%**	−28%	−4%	−27%	**−4%**
a^*_W			−27%	**−3%**		
b^*_M	−30%	−4%	−30%	−6%		
b^*_W			−30%	−6%	−27%	**−4%**
Indirect*	−49%	−7%	−49%	−9%	−63%	−10%
High reliability						
a^*_M	−18%	**−3%**	−17%	**−3%**	−18%	**−4%**
a^*_W			−17%	**−3%**		
b^*_M	−18%	**−3%**	−19%	−5%		
b^*_W			−19%	−5%	−18%	−4%
Indirect*	−32%	−5%	−33%	−8%	−41%	−10%
Small indirect effect						
a^*_M	−23%	**−3%**	−23%	−4%	−21%	**−3%**
a^*_W			−23%	−4%		
b^*_M	−21%	**0%**	−24%	−5%		
b^*_W			−25%	−6%	−20%	**−1%**
Indirect*	−38%	−2%	−42%	−9%	−51%	−8%
Large indirect effect						
a^*_M	−22%	**−3%**	−19%	**2%**	−22%	**−4%**
a^*_W			−18%	**2%**		
b^*_M	−26%	−6%	−33%	−16%		
b^*_W			−33%	−17%	−24%	−6%
Indirect*	−42%	−8%	−45%	−15%	−52%	−11%
Small a-path						
a^*_M	−22%	**−2%**	−20%	**1%**	−21%	−3%
a^*_W			−21%	**−1%**		
b^*_M	−24%	−5%	−28%	−10%		
b^*_W			−28%	−11%	−23%	−6%
Indirect*	−40%	−6%	−43%	−10%	−51%	−10%
Large a-path						
a^*_M	−24%	−5%	−22%	−5%	−24%	−6%
a^*_W			−22%	−5%		
b^*_M	−29%	−5%	−24%	−6%		
b^*_W			−24%	−6%	−24%	−5%
Indirect*	−46%	−10%	−41%	−10%	−53%	−12%

Note. a^*_M = effect of *X* on *M*, a^*_W = effect of *X* on *W*, b^*_M = effect of *M* on *Y*, b^*_W = effect of *W* on *Y*, Indirect* = total indirect effect of *X* on *Y*. The percentage bias is not defined for the direct effect of *X* on *Y* in a full mediation model and, thus, not included. Bold values indicate estimates for which the absolute larger bound of the 95% jackknife interval is smaller than 5%. Results are for TSSEM with adjustments of pooled correlations.

involving these variables are often examined with latent variable models or, in the context of bivariate meta-analyses, with adjustments for attenuation (see Wiernik & Dahlke, 2020). Surprisingly, awareness of this problem has not yet widely diffused to applications of multivariate MASEM. Adjustments for unreliability that correct for the biasing influence of measurement error is currently used rather unsystematically (Sheng et al., 2016). Some authors even argued that measurement error is no serious threat to the validity of MASEMs and, thus, can be generally neglected (Michel et al., 2011). Therefore, the present study evaluated this claim using a comprehensive simulation of

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

Table 4. Raw biases in regression coefficients with and without adjustments for unreliability

	Simple mediation		Parallel mediation		Sequential mediation	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Low reliability						
a^*_M	−.08	−.01	−.08	−.01	−.08	−.01
a^*_W			−.08	−.01		
b^*_M	−.09	−.01	−.09	−.02		
b^*_W			−.09	−.02	−.08	−.01
c*	.03	.01	.05	.01	.02	.01
Indirect*	−.04	−.01	−.09	−.02	−.02	.00
High reliability						
a^*_M	−.05	−.01	−.05	−.01	−.05	−.01
a^*_W			−.05	−.01		
b^*_M	−.05	−.01	−.06	−.02		
b^*_W			−.06	−.01	−.06	−.01
c*	.01	.01	.02	.01	.01	.01
Indirect*	−.03	.00	−.06	−.01	−.02	.00
Small indirect effect						
a^*_M	−.07	−.01	−.07	−.01	−.06	−.01
a^*_W			−.07	−.01		
b^*_M	−.02	.00	−.02	.00		
b^*_W			−.02	.00	−.02	.00
c*	.00	.00	.01	.00	.00	.00
Indirect*	−.01	.00	−.02	.00	−.01	.00
Large indirect effect						
a^*_M	−.07	−.01	−.06	.01	−.07	−.01
a^*_W			−.06	.01		
b^*_M	−.18	−.04	−.22	−.11		
b^*_W			−.22	−.11	−.16	−.04
c*	.05	.02	.09	.05	.03	.01
Indirect*	−.09	−.02	−.18	−.06	−.05	−.01
Small a-path						
a^*_M	−.03	.00	−.03	.00	−.03	−.01
a^*_W			−.03	.00		
b^*_M	−.14	−.03	−.17	−.06		
b^*_W			−.17	−.06	−.14	−.03
c*	.02	.00	.04	.02	.01	.00
Indirect*	−.04	−.01	−.08	−.02	−.02	.00
Large a-path						
a^*_M	−.15	−.03	−.13	−.03	−.14	−.03
a^*_W			−.13	−.03		
b^*_M	−.04	−.01	−.04	−.01		
b^*_W			−.04	−.01	−.04	−.01
c*	.02	.01	.04	.02	.01	.00
Indirect*	−.04	−.01	−.07	−.02	−.02	.00

Note. a^*_M = effect of *X* on *M*, a^*_W = effect of *X* on *W*, b^*_M = effect of *M* on *Y*, b^*_W = effect of *W* on *Y*, c* = effect of *X* on *Y*, Indirect* = total indirect effect of *X* on *Y*. Results are for TSSEM with adjustments of pooled correlations.

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

Table 5. Coverage rates of 95% confidence intervals with and without adjustments for unreliability

	Simple mediation		Parallel mediation		Sequential mediation	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Low reliability						
a^*_M	18%	**94%**	17%	93%	15%	93%
a^*_W			19%	**94%**		
b^*_M	19%	93%	15%	91%		
b^*_W			15%	93%	16%	93%
c*	71%	92%	48%	92%	85%	92%
Indirect*	8%	93%	1%	86%	1%	86%
High reliability						
a^*_M	39%	92%	45%	93%	39%	93%
a^*_W			44%	93%		
b^*_M	45%	93%	43%	92%		
b^*_W			46%	91%	40%	92%
c*	91%	**94%**	87%	**94%**	**94%**	**95%**
Indirect*	26%	91%	12%	88%	12%	88%
Small indirect effect						
a^*_M	30%	92%	27%	93%	27%	**94%**
a^*_W			26%	93%		
b^*_M	86%	**95%**	85%	**95%**		
b^*_W			85%	**95%**	84%	**94%**
c*	**94%**	**94%**	93%	**94%**	**95%**	**95%**
Indirect*	59%	**95%**	45%	**94%**	33%	93%
Large indirect effect						
a^*_M	27%	93%	34%	93%	27%	92%
a^*_W			35%	92%		
b^*_M	1%	79%	0%	26%		
b^*_W			0%	25%	1%	79%
c*	54%	92%	18%	80%	69%	92%
Indirect*	4%	88%	0%	57%	0%	83%
Small a-path						
a^*_M	61%	**94%**	64%	**95%**	64%	92%
a^*_W			63%	**95%**		
b^*_M	3%	87%	0%	64%		
b^*_W			1%	64%	3%	83%
c*	85%	**94%**	61%	**95%**	91%	**95%**
Indirect*	23%	93%	4%	87%	8%	87%
Large a-path						
a^*_M	3%	84%	3%	84%	2%	81%
a^*_W			2%	84%		
b^*_M	60%	**95%**	69%	**95%**		
b^*_W			70%	**95%**	58%	93%
c*	85%	93%	79%	**94%**	89%	**94%**
Indirect*	29%	93%	23%	93%	9%	87%

Note. a^*_M = effect of *X* on *M*, a^*_W = effect of *X* on *W*, b^*_M = effect of *M* on *Y*, b^*_W = effect of *W* on *Y*, c* = effect of *X* on *Y*, Indirect* = total indirect effect of *X* on *Y*. Bold values indicate estimates for which the 95% jackknife interval includes 95%. Results are for TSSEM with adjustments of pooled correlations.

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

MASEMs for different mediation models. These analyses led to three main conclusions:

First, measurement error biases effect estimates in MASEMs. In our simulation, the true path coefficients were underestimated by up to a third in some cases. For indirect effects, this problem was even more severe and reached distortions up to 50%. Moreover, biases depended on the specifics of the examined model. For example, larger biases were observed when variables were highly correlated or reliabilities were low. These results clearly undermine the claim that measurement error can generally be ignored in MASEMs as long as the variables exhibit levels of reliability typically observed in psychological research (Michel et al., 2011). Rather, the difference between observed and true effects was often substantial and, thus, potentially undermines the validity of interpretations for MASEMs. Second, measurement error affected the interval estimates of the effects, more so than the respective point estimates. As a result, the coverage rates of the confidence intervals often resembled a game of chance, resulting in intervals that, in some instances, included the true effects less than 1% of all times. Again, the undercoverage depended to some degree on the studied population model and, among others, was larger for less reliable measures. Finally, MASEMs with adjustments for unreliability in many situations led to largely unbiased point estimates and confidence intervals that were close to the nominal level. Although these adjustments generally improved the results for all examined models in all simulation conditions, they were less effective for models with highly correlated variables (e.g., $r > .60$ or .70). On the other hand, the performance of the adjusted models was not affected by the chosen meta-analytic method, that is, whether TSSEM (Cheung & Chan, 2005) or OSMASEM (Jak & Cheung, 2020) was used or whether the correlations were adjusted individually or using an artifact distribution.

These results demonstrate that measurement error has a detrimental impact on MASEM results that can threaten the validity of interpretations. Acknowledging fallible measurements by using adjustments for attenuation improves point and interval estimates and, thus, can help study true score effects in MASEMs.

Limitations and Directions for Future Research

The presented findings provide ample opportunity for follow-up research. For example, our simulations were limited to MASEMs of specific mediation models because these address popular multivariate hypotheses in psychological research. Especially, with more complex interrelations among the analysis variables (e.g., partial mediation, correlations between parallel mediators, sequential mediation with additional paths) or in more complex models (e.g., multivariable SEMs), the consequences of measurement error can be more or less serious as bias amplifying and compensating mechanisms can be present (e.g., Savalei, 2019; Sengewald & Pohl, 2019). Therefore, it might be particularly helpful for applied research to identify specific conditions under which adjustments for attenuation are more or less useful. In this respect, we encourage further research on analytic bias estimation (e.g., Savalei, 2019; Sengewald & Pohl, 2019), which could give insights into different factors that affect biases in different analysis models. This line of research could also be extended to the analysis of moderating effects. So far, adjustments for unreliability in MASEMs focus on the pooled correlations to estimate true score SEM parameters. How to best incorporate potentially fallible moderators in these analyses is, however, a currently still unresolved challenge. Similarly, our analyses were limited to the impact of measurement error on the SEM parameter estimates because these are typically the primary focus in MASEM research. Therefore, future research should explore in what way fallible measurements might also affect model fit statistics such as the chi-squared value or related goodness-of-fit indices. Finally, our research was limited to the effect of measurement error in MASEMs. However, psychometric meta-analyses emphasized several artifacts that might distort bivariate meta-analytic estimates (Wiernik & Dahlke, 2020). Therefore, follow-up research should evaluate how artifacts such as range restriction or artificial dichotomization might also affect MASEM results (e.g., De Jonge et al., 2020) and, more importantly, develop respective adjustment techniques.

Conclusion

Measurement errors in study variables can distort point and interval estimates of parameters in MASEMs. Although the size of these biases depends on the specific characteristics of the studied population model, the size of the unreliability, and the studied parameter, MASEM results are generally biased to some degree when measurement error is present. Individual adjustments for attenuation or adjustments using an artifact distribution allow recovering effect estimates in TSSEMs and OSMASEM that often are largely unbiased with confidence intervals close to the nominal level. Therefore, we encourage applied meta-analysts to regularly adopt adjustments for attenuation in MASEM research.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Brunner, M., Keller, L., Stallasch, S. E., Kretschmann, J., Hasl, A., Preckel, F., Lüdtke, O., & Hedges, L. V. (2022). Meta-analyzing individual participant data from studies with complex survey designs: A tutorial on using the two-stage approach for data from educational large-scale assessments. *Research Synthesis Methods*. Advance online publication. https://doi.org/10.1002/jrsm.1584

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

Cheung, M. W. L. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods*, *41*(2), 425–438. https://doi.org/10.3758/BRM.41.2.425

Cheung, M. W. L. (2013). Multivariate meta-analysis as structural equation models. *Structural Equation Modeling*, *20*(3), 429–454. https://doi.org/10.1080/10705511.2013.797827

Cheung, M. W.-L. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*, Article 1521. https://doi.org/10.3389/fpsyg.2014.01521

Cheung, M. W.-L., & Chan, W. (2004). Testing dependent correlation coefficients via structural equation modeling. *Organizational Research Methods*, *7*(2), 206–223. https://doi.org/10.1177/1094428104264024

Cheung, M. W.-L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, *10*(1), 40–64. https://doi.org/10.1037/1082-989X.10.1.40

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–315. https://doi.org/10.1037/a0033805

De Jonge, H., Jak, S., & Kan, K. J. (2020). Dealing with artificially dichotomized variables in meta-analytic structural equation modeling. *Zeitschrift für Psychologie*, *228*(1), 25–35. https://doi.org/10.1027/2151-2604/a000395

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

Fritz, M. S., Kenny, D. A., & MacKinnon, D. P. (2016). The combined effects of measurement error and omitted confounding in the single-mediator model. *Multivariate Behavioral Research*, *51*(5), 681–697. https://doi.org/10.1080/00273171.2016.1224154

Gnambs, T. (2014). A meta-analysis of dependability coefficients (test-retest reliabilities) for measures of the Big Five. *Journal of Research in Personality*, *52*, 20–28. https://doi.org/10.1016/j.jrp.2014.06.003

Gnambs, T. (2015). Facets of measurement error for scores of the Big Five: Three reliability generalizations. *Personality and Individual Differences*, *84*, 84–89. https://doi.org/10.1016/j.paid.2014.08.019

Gnambs, T., Scharl, A., & Schroeders, U. (2018). The structure of the Rosenberg Self-Esteem Scale: A cross-cultural meta-analysis. *Zeitschrift für Psychologie*, *226*(1), 14–29. https://doi.org/10.1027/2151-2604/a000317

Gnambs, T., & Sengewald, M.-A. (2023). *Supplemental materials to "Meta-analytic structural equation modeling with fallible measurements"*. https://doi.org/10.23668/psycharchives.8537

Gnambs, T., & Staufenbiel, T. (2018). The structure of the General Health Questionnaire (GHQ-12): Two meta-analytic factor analyses. *Health Psychology Review*, *12*(2), 179–194. https://doi.org/10.1080/17437199.2018.1426484

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, *26*(3), 329–367. https://doi.org/10.1177/0049124198026003003

Jak, S., & Cheung, M. W. L. (2020). Meta-analytic structural equation modeling with moderating effects on SEM parameters. *Psychological Methods*, *25*(4), 430–455. https://doi.org/10.1037/met0000245

Jak, S., & Cheung, M. W. L. (2022). Can findings from meta-analytic structural equation modeling in management and organizational psychology be trusted?. *PsyArxiv Preprints*. https://doi.org/10.31234/osf.io/b3qvn

Ke, Z., Zhang, Q., & Tong, X. (2019). Bayesian meta-analytic SEM: A one-stage approach to modeling between-studies heterogeneity in structural parameters. *Structural Equation Modeling*, *26*(3), 348–370. https://doi.org/10.1080/10705511.2018.1530059

Koehler, E., Brown, E., & Haneuse, S. J. P. (2009). On the assessment of Monte Carlo error in simulation-based statistical analyses. *The American Statistician*, *63*(2), 155–162. https://doi.org/10.1198/tast.2009.0030

Le, H., Oh, I. S., Schmidt, F. L., & Wooldridge, C. D. (2016). Correction for range restriction in meta-analysis revisited: Improvements and implications for organizational research. *Personnel Psychology*, *69*(4), 975–1008. https://doi.org/10.1111/peps.12122

Lee, K., & Beretvas, S. N. (2022). An evaluation of methods for meta-analytic structural equation modeling. *Structural Equation Modeling*, *29*(5), 1–13. https://doi.org/10.1080/10705511.2022.2047976

Marker, C., Gnambs, T., & Appel, M. (2022). Exploring the myth of the chubby gamer: A meta-analysis on sedentary video gaming and body mass. *Social Science & Medicine*, *301*, Article 112325. https://doi.org/10.1016/j.socscimed.2019.05.030

Michel, J. S., Viswesvaran, C., & Thomas, J. (2011). Conclusions from meta-analytic structural equation models generally do not change due to corrections for study artifacts. *Research Synthesis Methods*, *2*(3), 174–187. https://doi.org/10.1002/jrsm.47

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8

Neale, M. C., & Miller, M. B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics*, *27*(2), 113–120. https://doi.org/10.1023/A:1025681223921

R Core Team (2021). *R: A language and environment for statistical computing* (Version 4.1.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Rhemtulla, M., van Bork, R., & Borsboom, D. (2020). Worse than measurement error: Consequences of inappropriate latent variable measurement models. *Psychological Methods*, *25*(1), 30–45. https://doi.org/10.1037/met0000220

Sánchez-Meca, J., & Marín-Martínez, F. (2008). Confidence intervals for the overall effect size in random-effects meta-analysis. *Psychological Methods*, *13*(1), 31–48. https://doi.org/10.1037/1082-989X.13.1.31

Savalei, V. (2019). A comparison of several approaches for controlling measurement error in small samples. *Psychological Methods*, *24*(3), 352–370. https://doi.org/10.1037/met0000181

Scherer, R., & Teo, T. (2020). A tutorial on the meta-analytic structural equation modeling of reliability coefficients. *Psychological Methods*, *25*(6), 747–775. https://doi.org/10.1037/met0000261

Schroeders, U., Kubera, F., & Gnambs, T. (2022). The structure of the Toronto Alexithymia Scale (TAS-20): A meta-analytic confirmatory factor analysis. *Assessment*, *29*(8), 1806–1823. https://doi.org/10.1177/10731911211033894

Sengewald, M. A., & Pohl, S. (2019). Compensation and amplification of attenuation bias in causal effect estimates. *Psychometrika*, *84*(2), 589–610. https://doi.org/10.1007/s11336-019-09665-6

Sheng, Z., Kong, W., Cortina, J. M., & Hou, S. (2016). Analyzing matrices of meta-analytic correlations: Current practices and recommendations. *Research Synthesis Methods*, *7*(2), 187–208. https://doi.org/10.1002/jrsm.1206

Steiner, P. M., Cook, T. D., & Shadish, W. R. (2011). On the importance of reliable covariate measurement in selection bias adjustments using propensity scores. *Journal of Educational and Behavioral Statistics*, *36*(2), 213–236. https://doi.org/10.3102/1076998610375835

Van Erp, S., Verhagen, J., Grasman, R. P., & Wagenmakers, E. J. (2017). Estimates of between-study heterogeneity for 705 meta-analyses reported in *Psychological Bulletin* from 1990–2013. *Journal of Open Psychology Data*, *5*(1), Article 4. https://doi.org/10.5334/jopd.33

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, *48*(4), 865–885. https://doi.org/10.1111/j.1744-6570.1995.tb01784.x

Wedderhoff, N., Gnambs, T., Wedderhoff, O., Burgard, T., & Bosnjak, M. (2021). On the structure of affect: A meta-analytic investigation of the dimensionality and the cross-national applicability of the Positive and Negative Affect Schedule (PANAS). *Zeitschrift für Psychologie*, *229*(1), 24–37. https://doi.org/10.1027/2151-2604/a000434

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS ONE*, *11*(3), Article e0152719. https://doi.org/10.1371/journal.pone.0152719

Wiernik, B. M., & Dahlke, J. A. (2020). Obtaining unbiased results in meta-analysis: The importance of correcting for statistical artifacts. *Advances in Methods and Practices in Psychological Science*, *3*(1), 94–123. https://doi.org/10.1177/2515245919885611

History
Received June 1, 2022
Revision received July 25, 2022
Accepted September 16, 2022
Published online February 22, 2023

Acknowledgments
All authors approved the final version of the article.

Appendix A: Adjustments for Measurement Error in OSMASEM

One-stage MASEM (Jak & Cheung, 2020) adopts the same decomposition of the observed sample correlations $\mathbf{r}_i$ as TSSEM (see Equation 1). However, instead of fitting the hypothesized structural model to the pooled correlations $\widehat{\boldsymbol{\rho}}$, OSMASM constrains the population correlations $\boldsymbol{\rho}_i$ in Equation 1 to the correlations implied by the SEM as

$$\boldsymbol{\rho}_i = \text{vechs}\left(\mathbf{F}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{S}(\mathbf{I}-\mathbf{A})^{-1\text{T}}\mathbf{F}^{\text{T}}\right). \quad \text{(A1)}$$

If the SEM includes m manifest and l latent variables, $\mathbf{I}$ is a $m \times l$ identity matrix, $\mathbf{F}$ is a $m \times (m + l)$ selection matrix distinguishing manifest from latent variables, $\mathbf{A}$ is a $(m + l) \times (m + l)$ square matrix with asymmetric paths (e.g., regression weights, factor loadings), $\mathbf{S}$ is a $(m + l) \times (m + l)$ symmetrical matrix with (co)variances, and vechs() returns the lower diagonal elements of its arguments. The model can be estimated using full maximum likelihood with the *metaSEM* package (Cheung, 2015).

To estimate structural effects adjusted for measurement error, the implied correlations in A1 need to be attenuated using the average reliabilities across samples, $\boldsymbol{\gamma}_{i,x}$ and $\boldsymbol{\gamma}_{i,y}$, as

$$\boldsymbol{\rho}_i = \text{vechs}\left(\mathbf{F}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{S}(\mathbf{I}-\mathbf{A})^{-1\text{T}}\mathbf{F}^{\text{T}}\right)\left(\boldsymbol{\gamma}_{i,x}\boldsymbol{\gamma}_{i,y}\right)^{1/2}. \quad \text{(A2)}$$

Consequently, the structural parameters given in $\mathbf{A}$ and $\mathbf{S}$ represent unbiased estimates adjusted for measurement error. Although not directly implemented in *metaSEM*, the respective functions can be easily adapted to accommodate this additional constraint. The respective code is available in the online material.

Appendix B: Generation of Average Population Correlations

The data-generating average population correlations were calculated from the regression weights given in Table 1 as $\mathbf{P} = (\mathbf{I}-\mathbf{A})\mathbf{S}(\mathbf{I}-\mathbf{A})^{T}$. The size and values of the three matrices $\mathbf{I}$, $\mathbf{A}$, and $\mathbf{S}$ depended on the simulated mediation model (see Figure 1).

Simple Mediation

The three regression weights for the $m = 3$ variables (η_X, η_M, η_Y) referred to the effects of η_X on η_M (a), η_M on η_Y (b), and η_X on η_Y (c). Therefore, $\mathbf{I}$, $\mathbf{A}$, and $\mathbf{S}$ were given as

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ c & b & 0 \end{bmatrix}, \text{ and}$$

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-a^2 & 0 \\ 0 & 0 & 1-b^2-c^2-2abc \end{bmatrix}.$$

Conflict of Interest
We have no conflicts of interest to disclose.

Open Data
The study was not pre-registered. The information needed to reproduce all of the reported results and methodology is available in PsychArchives at https://doi.org/10.23668/psycharchives.8537 (Gnambs & Sengewald, 2023).

Funding
Open access publication enabled by the Leibniz Institute for Educational Trajectories, Bamberg, Germany.

ORCID
Timo Gnambs
https://orcid.org/0000-0002-6984-1276
Marie-Ann Sengewald
https://orcid.org/0000-0003-4155-394X

Timo Gnambs
Educational Measurement
Leibniz Institute for Educational Trajectories
Wilhelmsplatz 3
96047 Bamberg
Germany
timo.gnambs@lifbi.de

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000511 - Friday, June 02, 2023 1:38:39 AM - Universitätsbibliothek Bamberg IP Address:141.13.150.32

Parallel Mediation

The five regression weights for the $m = 4$ variables (η_X, η_M, η_W, η_Y) referred to the effects of η_X on η_M (a_M), η_M on η_Y (b_M), η_X on η_W (a_W), η_W on η_Y (b_W), and η_X on η_Y (c). Therefore, **I**, **A**, and **S** were given as

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a_M & 0 & 0 & 0 \\ a_W & 0 & 0 & 0 \\ c & b_M & b_W & 0 \end{bmatrix}, \text{ and}$$

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - a_M^2 & 0 & 0 \\ 0 & 0 & 1 - a_W^2 & 0 \\ 0 & 0 & 0 & 1 - b_M^2 - b_W^2 - c^2 - 2a_M b_M c - 2a_W b_W c - 2a_M b_M a_W b_W \end{bmatrix}.$$

Sequential Mediation

The four regression weights for the $m = 4$ variables (η_X, η_M, η_W, η_Y) referred to the effects of η_X on η_M (a), η_M on η_W (d), η_W on η_Y (b), and η_X on η_Y (c), resulting in

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ c & 0 & b & 0 \end{bmatrix}, \text{ and}$$

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - a^2 & 0 & 0 \\ 0 & 0 & 1 - d^2 & 0 \\ 0 & 0 & 0 & 1 - b^2 - c^2 - 2abcd \end{bmatrix}.$$

© 2023 The Author(s). Distributed as a Hogrefe OpenMind article under the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)