Research Synthesis Methods (2026), **0**: 1–30
doi:10.1017/rsm.2026.10109

Research Synthesis Methods

TUTORIAL

A tutorial on fitting flexible meta-analytic models with structural equation modeling in R

Mike W.-L. Cheung

Department of Psychology, National University of Singapore, Singapore
Email: mikewlcheung@nus.edu.sg

Received: 21 September 2025 UTC; **Revised:** 1 May 2026 UTC; **Accepted:** 19 June 2026 UTC

Keywords: graphical models; meta-analysis; SEM-based meta-analysis; structural equation modeling

Abstract

There are various statistical models with different features and assumptions in meta-analysis. Popular models include the fixed-effect (or common effect), random-effects, and mixed-effects models. Apart from these models, several alternative models, such as the multiplicative error model (also known as the unrestricted weighted least squares model), the hybrid of additive and multiplicative error models, and the location-scale model, have been proposed in the literature. Understanding these models can be challenging for researchers without a solid mathematical background. Implementing or modifying these models is even more challenging for researchers unless they have advanced statistical and programming knowledge. This tutorial elucidates a structural equation modeling (SEM) framework to understand these models. Several R packages are introduced to facilitate the specification and generation of graphical models for these meta-analytic models. Researchers may fit these models with the full information maximum likelihood estimation method. This holds significant potential across two domains. First, it can be used as an educational tool in teaching and learning meta-analytic models. Second, the framework supports the development and evaluation of novel meta-analytical models, which have not yet been implemented in current meta-analysis software, thereby fostering advancements in meta-analytic methods. This tutorial also demonstrates, using two real datasets, how to extend it to address research questions involving complex meta-analytic data. Limitations and future directions to extend the SEM-based meta-analysis are discussed.

Highlights

What is already known?

- Several common models are used in meta-analysis, including the fixed-effect (or common effect), random-effects, and mixed-effects models.
- Structural equation modeling (SEM) has been proposed to represent these models, a concept referred to as SEM-based meta-analysis in the literature.

What is new?

- This tutorial extends SEM-based meta-analysis to accommodate complex meta-analytic models, such as the multiplicative error model (or unrestricted weighted least squares model), the hybrid of additive and multiplicative error models, and the location-scale model.
- It also demonstrates how to specify and analyze these complex meta-analytic models among effect sizes, which may include linear and nonlinear constraints, using SEM language in R.

This article was awarded Open Materials badge for transparent practices. See the Data availability statement for details.

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Society for Research Synthesis Methodology. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

Potential impact for RSM readers

- RSM readers will not only learn about meta-analytic models through equations and algorithms, but can also explore these models through graphical representations.
- RSM readers can specify novel meta-analytic models that may not yet be implemented in their preferred software, enabling them to effectively address their research questions.

1. Introduction

Meta-analysis is a powerful tool for synthesizing research findings in various disciplines, including psychology,[1] management,[2] public health,[3] and medical science,[4] among others. The significance of meta-analysis in scientific research is underscored by several dedicated journals, including *Psychological Bulletin*, *Review of Educational Research*, *Neuropsychology Review*, and *Cochrane Reviews*, which prioritize the publication of systematic reviews and meta-analyses. *Research Synthesis Methods* publishes papers related to systematic reviews and meta-analyses and contributes to methodological advancements in this field. Historically, the fixed-effect model, also known as the common- or equal-effects model, was widely employed in meta-analysis; however, its limitations have since been recognized. For instance, the findings are only generalized to the studies included in the meta-analysis. The estimated standard errors (*SE*s) are underestimated when the goal is to generalize beyond the studies included and the effect sizes are not homogeneous.

Today, the random-effects and mixed-effects models are the most commonly used in contemporary meta-analyses. The random-effects model posits that the variability in effect sizes between studies is due to some unknown random factors and sampling error. In the presence of excess heterogeneity, the mixed-effects model incorporates moderators or covariates to account for this heterogeneity. Recently, alternative meta-analytic models, such as the multiplicative error model (or unrestricted weighted least squares [UWLS] model), hybrid models combining additive and multiplicative error terms, location-scale models, and nonlinear models, have been proposed and increasingly applied in the literature.

These advanced models are typically presented using equations and algorithms to obtain parameter estimates and the *SE*s, requiring a strong statistical background to understand and implement. Researchers using these models often need to wait for these models to be implemented in programming languages, such as R, Python, or Stata. Although some popular meta-analysis tools, such as the `metafor` package[5] in R,[6] offer valuable functionalities, customizing these models for specific research purposes can still pose significant challenges.

Structural equation modeling (SEM) is a widely used framework to analyze latent and observed variables. It offers several attractive features. One of them is the ability to represent models graphically, which facilitates the clear communication of statistical models and their underlying assumptions. This graphical representation enables readers to quickly grasp the core concepts of the models without requiring advanced mathematical knowledge. Moreover, rather than mastering specific estimation methods or algorithms for each model, readers may learn a general estimation method, such as the full information maximum likelihood (FIML) estimation.

In a series of papers, Cheung[7–12] introduced the SEM-based meta-analysis to analyze univariate, multivariate, and three-level meta-analytic models as structural equation models. He demonstrated how to use SEM packages like `OpenMx`[13,14] in R and M*plus*[15] to perform meta-analyses. This SEM-based meta-analysis has been implemented in the `metaSEM` package[16] in R. The present tutorial builds on this work by extending the applications of SEM to more advanced meta-analytic models, providing a guide on specifying, visualizing, and analyzing various common and advanced meta-analytic models as structural equation models in R.

This tutorial addresses two significant research gaps in the meta-analysis literature. First, it aids researchers in learning and understanding various meta-analytic models and their underlying assumptions by using a graphical modeling approach. In these models, there is a direct correspondence

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

between graphical representations, equations, and matrix forms, simplifying the teaching and learning of meta-analysis without the need for complex equations. Second, it empowers researchers to fit a range of meta-analytic models in R. This approach also allows for the customization of novel models to include antecedents and consequences on effect sizes, offering researchers the flexibility to specify their own meta-analytic models. The SEM-based meta-analysis offers a straightforward yet powerful means to implement and test novel research questions in meta-analysis.

The tutorial begins with an overview of SEM. It then guides users through the specification of common meta-analytic models within the SEM framework, including the fixed-effect and random-effects models, multiplicative-error models, and hybrid models that combine additive and multiplicative errors. Next, it covers models with covariates and complex models, such as location-scale and nonlinear models. Then, it illustrates how this approach can tailor models that do not fit traditional meta-analysis frameworks, including regression models among effect sizes. The tutorial includes examples that demonstrate these procedures using the `metaSEM` package,[16] which uses the `OpenMx`,[14] `lavaan`,[17] and `semPlot`[18] packages in the backend, and the `symSEM` package.[19] Finally, it discusses the limitations and potential future extensions of this approach. While this tutorial includes the mathematical models and assumptions of the associated meta-analytic models, it is essential to note that understanding these mathematics is not compulsory. Readers can also grasp these concepts by studying the graphical representations.

2. Brief introduction of structural equation modeling

In SEM, observed variables are represented by rectangles (or squares), while latent variables are represented by ellipses (or circles). We may represent the means or intercepts with triangles with a "1" inside, indicating columns of ones. Once we have variables, we may hypothesize their relationship by linking them with single or double arrows. Single arrows represent direct paths, like regression coefficients or factor loadings; double arrows represent covariances or variances (see the $\boldsymbol{A}$ and $\boldsymbol{S}$ matrices in the next paragraph).

These graphical elements have a direct correspondence to model equations. Each single-headed arrow pointing into a variable contributes one term to its equation of the form: coefficient × source variable. For example, in Figure 1(a), two arrows enter an effect size y_i: one from the intercept triangle (constant "1") with coefficient μ, and one from the latent variable e_i with coefficient 1. Reading these arrows directly gives the model equation $y_i = \mu \cdot 1 + 1 \cdot e_i = \mu + e_i$. This graphical-to-equation correspondence is illustrated below using the fixed-effect model as a fully worked example; readers are then invited to apply the same logic when inspecting subsequent path diagrams throughout the tutorial.

Structural equation models are typically specified using equations due to their convenience and clarity. The `lavaan` syntax,[17] inspired by M*plus*, is a popular choice in R. Readers unfamiliar with the lavaan syntax may refer to `lavaan`'s website for an introduction. After a structural equation model is specified, most SEM packages automatically convert the equations to matrices to facilitate the computations. The Reticular Action Model (RAM)[20] is a powerful formulation for representing SEM models. Suppose that p_o is the number of observed variables, p_l is the number of latent variables in the model, and $p = p_o + p_l$ is the total number of variables. Four matrices are sufficient to specify most structural equation models in RAM:

1. The $\boldsymbol{A}$ $p \times p$ matrix represents asymmetric paths from variables in columns to variables in rows, such as regression coefficients and factor loadings.
2. The $\boldsymbol{S}$ $p \times p$ matrix represents symmetric paths, such as covariances and variances.
3. The $\boldsymbol{M}$ $p \times 1$ vector is used to represent the means and intercepts.
4. The $\boldsymbol{F}$ $p_o \times p$ filter matrix is used to select the observed variables.

Once the models are specified, SEM programs convert the proposed model into a model-implied covariance matrix $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$ and a model-implied mean vector $\boldsymbol{\mu}_i(\boldsymbol{\theta})$ in the ith subject (or study in a

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

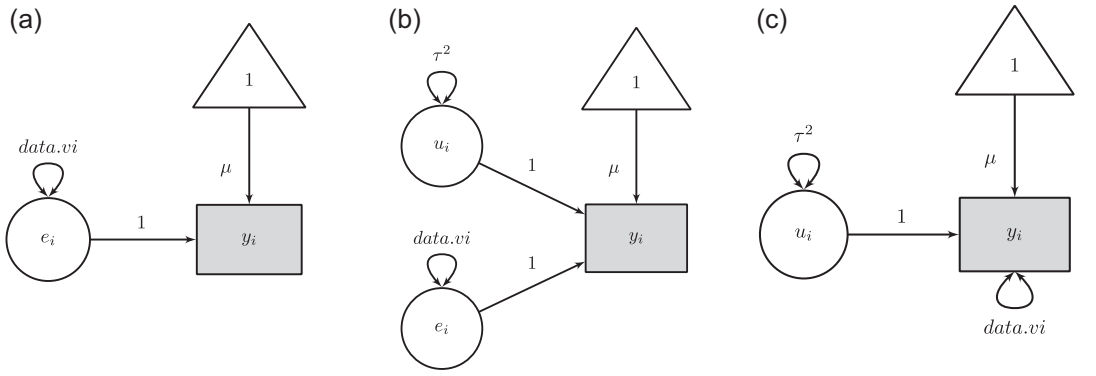


Figure 1. *Structural equation models of fixed- and random-effects models.*
Note: Figure (a) illustrates a fixed-effect (or common-effect) model. Panels (b) and (c) depict a random-effects model that incorporates additive heterogeneity error. The sampling error e_i is conceptualized as a latent variable in Figures (a) and (b) and as measurement error in Figure (c). Here, y_i represents the effect size for the ith study, while $data.vi$ indicates its known sampling error variance. Additionally, u_i stands for the random effect with τ^2 representing the heterogeneity variance. Furthermore, μ denotes the average effect (or common effect in the fixed-effect model).

meta-analysis):

$$\begin{aligned}\boldsymbol{\mu}_i(\boldsymbol{\theta}) &= \boldsymbol{F}_i(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{M},\\ \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) &= \boldsymbol{F}_i(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{S}((\boldsymbol{I}-\boldsymbol{A})^{-1})^{\mathrm{T}}\boldsymbol{F}_i^{\mathrm{T}},\end{aligned} \tag{2.1}$$

where $\boldsymbol{\theta}$ is a vector of the free parameters and I is a $p \times p$ identity matrix. It should be noted that there is a subscript i in $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$, $\boldsymbol{\mu}_i(\boldsymbol{\theta})$, and $\boldsymbol{F}_i$, indicating that the dimensions of $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$, $\boldsymbol{\mu}_i(\boldsymbol{\theta})$, and $\boldsymbol{F}_i$ may vary across subjects (or studies in meta-analysis). This feature is essential as it allows the model to filter out data with missing values before analysis. Assuming the data are multivariate normal, the −2*log-likelihood (−2*LL*) for all *K* subjects–or studies in a meta-analysis–is:

$$-2LL(\boldsymbol{\theta}) = \sum_{i=1}^{K}\left(k_i\log(2\pi) + \log|\boldsymbol{\Sigma}_i(\boldsymbol{\theta})| + (\boldsymbol{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\theta}))^{\mathrm{T}}\boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\theta}))\right), \tag{2.2}$$

where K is the number of subjects, k_i is the number of complete observed variables, and $\boldsymbol{y}_i$ is the vector of observed data in the ith subject. The parameter estimates, denoted as $\hat{\boldsymbol{\theta}}$, are derived by minimizing the $-2LL$ function. Once the model is fitted, the software provides a range of outputs, including parameter estimates, *SE*s, 95% Wald CIs, or likelihood-based confidence intervals (LBCIs),[21] and various fit statistics.

To make the RAM formulation concrete and accessible, we now work through the fixed-effect model in Figure 1(a) as a step-by-step illustration. This model contains $p_o = 1$ observed variable, y_i (the rectangle), and $p_l = 1$ latent variable, e_i (the ellipse), so there are $p = p_o + p_l = 2$ variables in total. Applying the `lavaan2RAM` function to the `lavaan` specification produces the following four RAM matrices, with variable names labeling all rows and columns:

$$\begin{aligned}\boldsymbol{A} &= \begin{array}{c} \\ y_i \\ e_i \end{array}\begin{array}{c}\begin{array}{cc} y_i & e_i \end{array}\\ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\end{array}, \quad \boldsymbol{S} = \begin{array}{c} \\ y_i \\ e_i \end{array}\begin{array}{c}\begin{array}{cc} y_i & e_i \end{array}\\ \begin{bmatrix} 0 & 0 \\ 0 & v_i \end{bmatrix}\end{array},\\ \boldsymbol{F} &= \begin{array}{c} \\ y_i \end{array}\begin{array}{c}\begin{array}{cc} y_i & e_i \end{array}\\ \begin{bmatrix} 1 & 0 \end{bmatrix}\end{array}, \quad \boldsymbol{M} = \begin{array}{c} \\ y_i \\ e_i \end{array}\begin{array}{c}\begin{array}{c} \\ \end{array}\\ \begin{bmatrix} \mu \\ 0 \end{bmatrix}\end{array}.\end{aligned} \tag{2.3}$$

Every non-zero entry has a direct counterpart in Figure 1(a):

- *A matrix (single-headed arrows).* The entry $\boldsymbol{A}[\text{row}, \text{col}]$ records the coefficient of the directional path pointing *from* the column variable *to* the row variable. The single non-zero entry $\boldsymbol{A}[y_i, e_i] = 1$

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

corresponds to the arrow from the ellipse e_i to the rectangle y_i with coefficient 1 in Figure 1(a). It encodes the contribution $1 \times e_i$ to the equation for y_i. All other entries are zero because no other directional arrows exist in this model.

- *S matrix (double-headed arrows: variances and covariances).* The entry $\boldsymbol{S}[\text{row}, \text{col}]$ records the covariance between the row and column variables, or the variance when row = col. The single non-zero entry $\boldsymbol{S}[e_i, e_i] = v_i$ corresponds to the curved double-headed loop on the ellipse e_i in Figure 1(a), encoding the sampling variance v_i of the ith study. Because v_i is a *known* constant that differs from study to study rather than a free parameter to be estimated, `metaSEM` reads it from the data via the definition variable `data.vi`, which will be introduced below.
- *F matrix (selecting observed variables).* The $\boldsymbol{F}$ matrix has one row for each observed variable and one column for every variable in the model. The entry $\boldsymbol{F}[y_i, y_i] = 1$ identifies y_i as the sole observed variable (the rectangle in Figure 1(a)). The entry $\boldsymbol{F}[y_i, e_i] = 0$ confirms that e_i (the ellipse) is latent and not directly measured.
- *M vector (means and intercepts).* The entry $\boldsymbol{M}[y_i] = \mu$ encodes the single-headed arrow from the small triangle labeled "1" (the constant term) to y_i with coefficient μ in Figure 1(a). This corresponds to the mean or intercept of y_i. The entry $\boldsymbol{M}[e_i] = 0$ reflects the standard SEM convention that latent variables are mean-zero.

In this tutorial, we will define meta-analytic models using `lavaan` syntax, and the `metaSEM` package will automatically translate these equations into RAM specifications. This process may also involve visualizing the models, computing the model-implied matrices, and conducting data analysis. Although a deep understanding of matrix representation is not necessary, having some familiarity with it can be beneficial. This knowledge will aid in verifying the models, ensuring they accurately represent the underlying theoretical constructs and hypotheses being tested.

Table 1 shows some common operators in `lavaan`. For observed variables (e.g., y and x), regression paths are specified using the ~ operator. For example, the formula `y ~ x` encodes the linear regression $y = \beta_0 + \beta_1 x + e$, where β_0 (intercept) and β_1 (slope) are the estimated parameters. To label the slope for later reference or constraints, users may write `y ~ b*x`, where `b` is the parameter name. This facilitates advanced specifications, such as imposing equality constraints (e.g., `y ~ b1*x1+b2*x2; b2==b1+1`). Latent variables, which represent unobserved constructs, are defined using the =~ operator. For instance, `f =~ x1+x2+x3` signifies that the latent variable `f` is measured by observed indicators x1, x2, and x3, implying the measurement model: $x_1 = \lambda_1 f + e_1$, $x_2 = \lambda_2 f + e_2$, and $x_3 = \lambda_3 f + e_3$, where λ_1, λ_2, and λ_3 are factor loadings, and e_1, e_2, and e_3 are measurement errors.

Covariances and variances are specified with ~~. The formula `x1 ~~ cov*x2` estimates the covariance between x1 and x2, labeled with `cov`, if they are exogenous, or their residual covariance if endogenous. Similarly, `x1 ~~ var*x1` labels the (residual) variance of `x1` as `var`. For derived parameters (e.g., indirect effects in mediation models), the `:=` operator is used. For example, `indirect := a*b` defines `indirect` as the product of paths a and b. Complex equality constraints can be imposed using the == operator. For example, `a == exp(b+c)` means that the relationship of $a = \exp(b + c)$ will be imposed among the parameters.

3. Univariate meta-analyses

We utilize a dataset from Jaramillo et al.[22] to illustrate various univariate meta-analyses, thereby reinforcing our understanding and gaining practical insights. The effect size in the ith study is the correlation (r_i) between organizational commitment and salesperson job performance, which we refer to as y_i here. Its sampling variance v_i is calculated by $(1-r_i^2)^2/(n_i-1)$, where n_i is the sample size. The dataset comprises 61 studies with various study-level characteristics, and we have chosen to include the individualism score of the country (x_i) where the study was conducted as a covariate. Moreover, we manually searched the citations of studies (z_i) in Google Scholar as a potential outcome variable of the effect size.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

Table 1. *Examples of* `lavaan` *syntax.*

Symbol	Example	Meaning	Remarks
`~`	`y ~ x`	Regression: y is modeled by x.	Implied model: $y = b_0 + b_1 x$, where b_0 (intercept) and b_1 (slope) are estimated.
`~`	`y ~ a*x`	Regression with labeled coefficient.	Explicitly names the slope parameter a: $y = b_0 + a * x$. The parameter a can be used in the constraints or computations.
`~`	`y ~ m*1`	Mean/intercept of y.	If y is an independent variable, estimate its mean labeled m. If y is a dependent variable, estimate its intercept labeled m.
`~~`	`y ~~ cov*x`	Covariance between y and x.	Estimates covariance labeled cov if y and x are independent. For dependent variables, estimate residual covariance labeled cov.
`~~`	`y ~~ var*y`	Variance of y.	Estimates variance labeled var if y is an independent variable. For dependent variables, estimate its residual variance labeled var.
`=~`	`f =~ x1 + x2 + x3`	Latent variable definition.	f is a latent variable measured by $x1$, $x2$, and $x3$. Implied model: $x1 = a_1 * f + e_1$, $x2 = a_2 * f + e_2$, and $x3 = a_3 * f + e_3$, where e_1, e_2, and e_3 are the measurement errors. Factor loadings are freely estimated unless constrained.
`=~`	`f =~ 1*x1 + a*x2 + b*x3`	Latent variable with labeled parameters.	Fixes the loading of $x1$ to 1 (scaling), names the loading of $x2$ as a, and names the loading of $x3$ as b.
`:=`	`ind := a*b`	Parameter definition.	Creates a new function of parameters ind (e.g., indirect effect) from $(a * b)$.
`==`	`a == b+1`	Equality constraint.	Imposes constraints on parameters $a = b + 1$. Simpler constraints, such as `a == b` (equality of two parameters), are equally supported.

To help readers navigate between the SEM-based and the conventional meta-analysis frameworks, we note which models in this tutorial have equivalent implementations in the `metafor` package.[5] The fixed-effect and random-effects models are directly available via `rma(yi, vi, method = "FE")` and `rma(yi, vi, method = "ML")`, respectively. The multiplicative-error (UWLS) model is available via `rma(yi, vi, method = "EE", test = "knha")`. Fixed-x meta-regression is standard via the `mods` argument; the location-scale model is supported via the `scale` argument[23]; and the Hartung–Knapp–Sidik–Jonkman (HKSJ) correction is available via `test = "knha"`. The Baker–Jackson hybrid models, meta-regressions involving true effect sizes as predictors or outcomes, partially standardized regression, and correlation between true effect sizes are *not* currently available in `metafor` and represent novel contributions of the SEM-based framework. Complete `metafor` syntax for the applicable models is provided in the Supplementary Material.

3.1. Models without external variables

The most popular meta-analytic models involve only the effect sizes and their sampling variances. We will discuss several such models in this section.

3.1.1. Fixed-effect model

The fixed-effect or common-effect model usually assumes all studies share the same true or population effect size μ (cf. the studies[24,25] for an alternative fixed-effects model that does not assume equal true effect sizes). The fixed-effect model is

$$\begin{aligned} y_i &= \mu + e_i, \\ \mathrm{Var}\,(e_i) &= v_i. \end{aligned} \tag{3.1}$$

Figure 1(a) displays a graphical representation of the fixed-effect model. To analyze a meta-analysis as a structural equation model, we can consider the k studies in a meta-analysis as the k subjects in SEM. Since there are 61 studies in Jaramillo et al.,[22] we can consider this as having 61 subjects from the SEM perspective. In this perspective, y_i is an observed variable, while a latent variable represents the sampling error e_i. Because the sampling variance v_i is treated as known and varies across studies, we use a *definition variable* to assign each observed v_i to the variance of the ith study, which is represented by `data.vi` in `OpenMx`. The model-implied variance and mean of the ith study are

$$\begin{aligned} \mu_i(\theta) &= \mu, \\ \Sigma_i(\theta) &= v_i. \end{aligned} \tag{3.2}$$

This model has only one parameter ($\theta = \mu$). After fitting the model with FIML, we may obtain the parameter estimate, its SE, and construct a 95% Wald CI by $\hat{\theta} \pm 1.96SE$.

The typical analysis workflow involves specifying a model using `lavaan` syntax. Next, the proposed model is converted to the RAM specification using the `lavaan2RAM` function, and the data are fitted using the `sem` function. Optionally, the graphical model can be displayed using the `plot` function, and the model-implied matrices can be obtained using the `impliedS` function. The following R code specifies a fixed-effect model. The first model, `m1a` (lines 15–19), explicitly specifies e_i as a latent variable. The mean of y_i is labeled `mu` in the model (line 16), whereas the known variance v_i is assigned as the sampling variance of y_i using `data.vi` (line 18). Since the variances are assumed to be present on the observed variables in SEM, the second model, `m1b` (lines 26–28), fits the same model without specifying e_i. Because these two models are equivalent, it is up to researchers to decide how to present the model.

After specifying the model, we may convert the `lavaan` equations to RAM specification with the `lavaan2RAM` function (line 32) and compute the model-implied variance and mean (Equation (3.1)) symbolically with the `impliedS` function (line 38). We specify `std.lv=FALSE` (line 32) when converting the model to RAM specification with the `lavaan2RAM` function because the

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

default is `std.lv=TRUE`, which indicates that the variances of the latent variables are standardized at 1. However, we want to assign the known sampling variance `data.vi` in our model (lines 18 and 27) rather than fixing it at 1. Then, we may fit the model using the `sem` function (lines 41 and 42). Moreover, we may also plot the conceptual model (line 23) and the fitted model (line 46) with the parameter estimates using the `plot` function. After running the analysis, the parameter estimate and its 95% Wald CIs are $\hat{\mu} = 0.1938, [0.1782, 0.2093]$. We will discuss and compare the results after introducing all models.

```
1  ## Load the required libraries
2  library(metaSEM)
3  library(symSEM)
4
5  ## Prepare a sample data
6  ## xi: a predictor
7  ## zi: an outcome variable
8  dat1 <- data.frame(yi=Jaramillo05$r,
9                     vi=Jaramillo05$r_v,
10                    xi=scale(Jaramillo05$IDV, scale=FALSE),
11                    zi=scale(Jaramillo05$Citations, scale=TRUE))
12
13 ## Fixed-effect model with a latent variable ei
14 ## data.vi means that vi in the data are fixed at each case in the model.
15 m1a <- "ei =~ 1*yi           ## e is a latent variable and yi is its indicator
16         yi ~ mu*1            ## Mean(yi) = mu
17         yi ~~ 0*yi           ## Set the measurement error variance of yi at 0
18         ei ~~ data.vi*ei     ## Var(yi) = vi, which is assumed known in meta-analysis
19 "
20
21 ## Display the conceptual model using the semPlot package
22 ## sizeInt=7: Set 7 as the size of the intercept, see help(semPaths) in R
23 plot(m1a, color = "yellow", sizeInt = 7)
24
25 ## Fixed-effect model without specifying ei
26 m1b <- "yi ~ mu*1            ## Mean(yi) = mu
27         yi ~~ data.vi*yi     ## Var(yi) = vi, the measurement error variance is assigned from
28                                         the data
29 "
30
31 ## Convert the lavaan model to RAM specification
32 ## obs.variables = "yi": "yi" is the observed variable
33 ram1 <- lavaan2RAM(m1a, obs.variables = "yi", std.lv = FALSE)
34
35 ## Display the RAM model
36 ram1
37
38 ## Get the model-implied mean and variance structures
39 impliedS(ram1)
40
41 ## Fit the model to the data
42 fit1 <- sem("Fixed", RAM = ram1, data = dat1)
43 summary(fit1)
44
45 ## Display the model with results using the semPlot package
46 ## nDigits=4: Display 4 digits
47 plot(fit1, color = "green", sizeInt = 7, nDigits = 4)
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

3.1.2. Random-effects model

Studies can differ in various characteristics, such as samples, research designs, and measurement techniques. These differences can lead to variability in the effect sizes observed across different studies beyond sampling error. The random-effects model addresses this issue by allowing the possibility that each study estimates its own unique population effect size. Specifically, the random-effects model assumes that the observed effect size in each study (y_i) is influenced by two sources of variability: true heterogeneity (τ^2), which represents the actual differences in effect sizes across studies, and sampling error (v_i). This approach allows researchers to generalize the findings beyond the individual studies in the meta-analysis, making the results more applicable to a broader context.[26] The random-effects model of the ith study is written as

$$\begin{aligned} y_i &= \mu + u_i + e_i, \\ \mathrm{Var}\,(e_i) &= v_i, \\ \mathrm{Var}\,(u_i) &= \tau^2, \end{aligned} \tag{3.3}$$

where μ is the average effect, and u_i is the study-specific true effect deviated from the average effect. The model-implied variance and mean of the ith study are

$$\begin{aligned} \mu_i(\boldsymbol{\theta}) &= \mu, \\ \Sigma_i(\boldsymbol{\theta}) &= \tau^2 + v_i. \end{aligned} \tag{3.4}$$

Now, there are two parameters in the model ($\boldsymbol{\theta} = [\mu, \tau^2]^{\mathrm{T}}$). The random-effects model becomes the fixed-effect model when $\tau^2 = 0$. Several estimators can be used to fit the model and make statistical inferences, such as the unweighted and weighted methods of moments, ML, and restricted maximum likelihood (REML) estimation. Readers can refer to, for example, Viechtbauer[27] and Veroniki et al.,[28] for comparing these estimators. The current tutorial uses FIML estimation, though REML estimation can also be implemented in SEM.[29]

Figure 1(b) displays the random-effects model. The study-specific true effect deviating from the average effect is conceptualized as a latent variable u_i. Its variance τ^2 represents an additive heterogeneity variance (lines 3 and 4 in the following R code). Therefore, there are two latent variables u_i and e_i in model `m2a`. Figure 1(c) shows the same model by simplifying it with a single latent variable u_i, whereas e_i is conceptualized as the measurement error. The model is shown in `m2b` in R. To save space, we would not repeat the steps similar to those in the fixed-effect model. Readers can refer to the Supplementary Material for the complete R code and output. After fitting the model, the estimated parameters and their 95% Wald CIs are $\hat{\mu} = 0.1866\ [0.1487, 0.2245]$ and $\hat{\tau}^2 = 0.0170 [0.0089, 0.0251]$.

```
## ei is conceptualized as a latent variable.
## data.vi means that vi in the data are fixed at each case in the model.
m2a <- "ui =~ 1*yi          ## ui is the random effect
        ui ~~ tau2*ui       ## Var(ui) = tau2
        ei =~ 1*yi          ## ei, the error term, is conceptualized as a latent variable
        ei ~~ data.vi*ei    ## Var(ei) = vi, which is assigned from data
        yi ~ mu*1           ## Mean(yi) = mu
        yi ~~ 0*yi          ## Set the measurement error variance of yi at 0
"

## ei is conceptualized as the measurement error of yi
## data.vi means that vi in the data are fixed at each case in the model.
m2b <- "ui =~ 1*yi          ## ui is the random effect
        ui ~~ tau2*ui       ## Var(ui) = tau2
        yi ~ mu*1           ## Mean(yi) = mu
        yi ~~ data.vi*yi    ## Var(yi) = vi, the measurement error variance is assigned from
                                       the data
"
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

3.1.3. Multiplicative-error model

As defined in Equation (3.3), the random-effects model conceptualizes heterogeneity additively through the u_i. However, several alternative models have been proposed to account for heterogeneity in different ways. Notably, Thompson and Sharp[30] introduced a multiplicative heterogeneity model to meta-analysis. The key idea behind this model is that the sampling error variance v_i is not treated as a fixed and known quantity but rather as something that is estimated up to a proportion. This approach acknowledges the potential uncertainty in the precision of the effect size estimates across studies, offering a different perspective on handling heterogeneity in meta-analysis. Several recent studies, for example,[31–33] suggest this model performs well, especially in handling potential publication bias. The model is

$$\begin{aligned} y_i &= \mu + \sqrt{\phi} e_i, \\ \mathsf{Var}\,(e_i) &= v_i, \end{aligned} \tag{3.5}$$

where μ is the common effect and ϕ is a scaling or dispersion factor. The model-implied variance and mean are

$$\begin{aligned} \mu_i(\boldsymbol{\theta}) &= \mu, \\ \Sigma_i(\boldsymbol{\theta}) &= \phi v_i. \end{aligned} \tag{3.6}$$

Figure 2(a) displays the graphical model. The *factor loading* $\sqrt{\phi}$ scales the impact of the latent factor e_i on y_i. The multiplicative-error model simplifies to the fixed-effect model when $\phi = 1$. However, when the studies exhibit heterogeneity, the estimated $\hat{\phi}$ is typically greater than 1, a situation known as the over-dispersion model.[34] In this context, the estimated *SE*s are larger than those in the fixed-effect model, reflecting the additional variability due to heterogeneity. It is also possible for the estimated $\hat{\phi}$ to be less than 1, which is referred to as the under-dispersion model. In this case, the estimated *SE*s are smaller than those in the fixed-effect model, indicating the estimated variability among the studies is even smaller than that assumed by the fixed-effect model. Some researchers opt to truncate $\hat{\phi}$ to 1 to ensure that the estimated *SE*s of the multiplicative-error model are not smaller than those in the fixed-effect model. This truncation helps maintain a conservative estimate of variability and prevents underestimating *SE*s. Readers may refer to van Aert and Jackson,[35] which discusses the implications and methodologies regarding the adjustment of $\hat{\phi}$ in the multiplicative-error model.

The following R code specifies the multiplicative-error model. The factor loading from the latent factor e_i to y_i is represented by `phi_sqrt` ($\sqrt{\phi}$) in line 2 of the code. To derive the estimate and CIs on ϕ, we introduce a new function of parameters `phi`, which is defined as the square of `phi_sqrt` using the `:=` operator (line 6). This transformation allows us to request a Wald CI using the Delta method or an LBCI on `phi`. After fitting the model, the estimated parameters and their 95% Wald CIs are as follows: the point estimate for μ is 0.1938, with a 95% Wald CI of $[0.1571, 0.2304]$, and the point estimate for $\sqrt{\phi}$ is 2.3588, with a 95% Wald CI of $[1.9402, 2.7773]$. Consequently, the point estimate for ϕ is 5.5637, with a 95% Wald CI of $[3.5891, 7.5384]$. Since the 95% Wald CI for ϕ does not include 1, this indicates that the data exhibit over-dispersion with excess heterogeneity, suggesting that the variability in the data is greater than the sampling error.

```
1 ## Multiplicative error, same as the UWLS model
2 m3 <- "ei =~ phi_sqrt*yi     ## phi (phi_sqrt^2) is the heterogeneity scaling factor
3        ei ~~ data.vi*ei      ## Var(ei) = vi
4        yi ~ mu*1             ## Mean(yi) = mu
5        yi ~~ 0*yi            ## Set the default error variance of yi at 0
6        phi := phi_sqrt^2     ## Define phi as a function of phi_sqrt
7 "
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

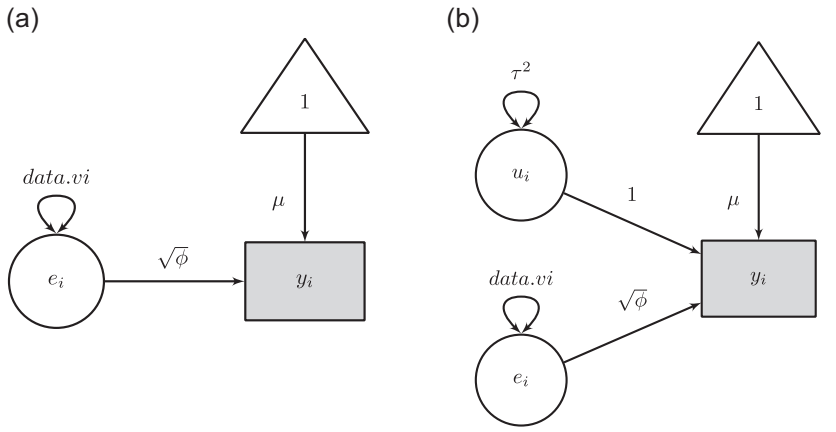


Figure 2. *Structural equation models of meta-analytic models with multiplicative and additive heterogeneity.*

Note: Figure (a) illustrates a meta-analytic model characterized by multiplicative heterogeneity, while Figure (b) depicts a model that incorporates both multiplicative and additive heterogeneity. y_i represents the effect size for the *i*th study, and $data.vi$ indicates the known sampling error variance associated with that effect size. The symbol μ denotes the average effect across studies. Additionally, u_i represents the additive random effect, with τ^2 indicating the variance associated with heterogeneity. The parameter ϕ serves as the scaling factor for the multiplicative heterogeneity.

This multiplicative-error model has been applied in large-scale empirical meta-analyses. For example, Stanley et al.[33] compared the UWLS and random-effects models across 67,308 Cochrane meta-analyses, finding that the UWLS model often yielded superior representation of the heterogeneity structure in medical research. In addition, Stanley et al.[36] presented evidence that UWLS can help reduce bias in meta-analyses of correlations, particularly in situations where the effect sizes and their sampling variances are correlated.

3.1.4. Hybrid models of additive and multiplicative errors

Several researchers[37,38] have integrated additive and multiplicative error meta-analysis into hybrid models within a Bayesian framework. More recently, Baker and Jackson[39] proposed a hybrid model by combining two traditions of meta-analysis—the multiplicative model in particle physics and the additive model in the meta-analysis literature. Given the complexity of their approach, it is instructive first to present a simplified version suggested by Schmid.[40]

Hybrid model 1: Essentially, the simplest hybrid model combines two sources of variability: study-specific variance τ^2 and scaled sampling variance ϕv_i[40]:

$$\begin{aligned} y_i &= \mu + u_i + \sqrt{\phi} e_i, \\ \mathrm{Var}\,(e_i) &= v_i, \\ \mathrm{Var}\,(u_i) &= \tau^2. \end{aligned} \tag{3.7}$$

The model-implied variance and mean of the *i*th study are

$$\begin{aligned} \mu_i(\boldsymbol{\theta}) &= \mu, \\ \Sigma_i(\boldsymbol{\theta}) &= \tau^2 + \phi v_i. \end{aligned} \tag{3.8}$$

Figure 2(b) displays the hybrid model. Conceptually, the two sources of variability are represented as two latent factors. The square root of the scaling factor $\sqrt{\phi}$ (`phi_sqrt` in the R code) is depicted by the factor loading of the latent factor e_i to y_i. The hybrid model becomes the random-effects model in Equation (3.3) when $\phi = 1$, whereas it is equivalent to the multiplicative model in Equation (3.5) when

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

$\tau^2 = 0$. The following R code specifies the hybrid model. The latent factors `ui` (lines 5 and 6) and `ei` (lines 8 and 9) represent the study-specific effect and the sampling error, respectively. Similar to the R code for the multiplicative error model, we define a new function to compute the original ϕ (line 10).

After fitting the model, the parameter estimates and their 95% Wald CIs are as follows: $\hat{\mu} = 0.1874$ with a 95% Wald CI of $[0.1497, 0.2251]$, $\sqrt{\hat{\phi}} = 1.2870$ with a 95% Wald CI of $[0.3825, 2.1916]$, and $\hat{\tau}^2 = 0.0137$ with a 95% Wald CI of $[0.0007, 0.0267]$. Additionally, the estimated value of ϕ and its 95% Wald CI is 1.6565 with a 95% Wald CI of $[-0.6720, 3.9849]$. Since the Wald CI for ϕ (or $\sqrt{\hat{\phi}}$) includes 1, it is not statistically significant at the 0.05 level for the hypothesis test $H_0 : \phi = 1$.

```
1  ## Hybrid model 1: multiplicative and additive error
2  m4 <- "yi ~ mu*1                ## Mean(yi) = mu
3         yi ~~ 0*yi               ## Set the measurement error variance of yi at 0
4         ## additive error
5         ui =~ 1*yi
6         ui ~~ tau2*ui            ## Var(ui) = tau2
7         ## multiplicative error
8         ei =~ phi_sqrt*yi        ## phi is a scaling factor
9         ei ~~ data.vi*ei         ## Var(ei) = vi
10        phi := phi_sqrt^2        ## Define phi as a function of phi_sqrt
11 "
```

Hybrid model 2: Now, we discuss the original hybrid model proposed by Baker and Jackson.[39] The model is

$$\begin{aligned} y_i &= \mu + \tilde{e}_i, \\ \text{Var}\,(\tilde{e}_i) &= v_i + \tau^2 v_i^{\omega}, \end{aligned} \tag{3.9}$$

where $0 < \omega < 1$. The model-implied variance and mean in the ith study are

$$\begin{aligned} \mu_i(\boldsymbol{\theta}) &= \mu, \\ \Sigma_i(\boldsymbol{\theta}) &= v_i + \tau^2 v_i^{\omega}. \end{aligned} \tag{3.10}$$

This hybrid model simplifies to the random-effects model (Equation (3.3)) when $\omega = 0$. When $\omega = 1$, the model-implied variance is $\Sigma_i(\boldsymbol{\theta}) = (1 + \tau^2)v_i$, where $(1 + \tau^2)$ acts as a non-negative scaling factor. This contrasts with the hybrid model in Equation (3.7), where the scaling factor ϕ can be smaller than 1. Please also note that τ^2 is no longer the additive heterogeneity variance because it is multiplied by v_i. For values of $0 < \omega < 1$, the model exhibits intermediate behavior between additive and multiplicative models.

It is not straightforward to represent this hybrid model graphically. Here, we introduce a method to specify a structural equation model with arbitrary functions of parameters. In lavaan syntax, the == operator ensures a linear constraint on the parameters. For example, `a1==a2*a3` enforces that the parameter `a1` will equal the product of the parameters `a2` and `a3` after the optimization. It should be noted that all three parameters (`a1`, `a2`, and `a3`) must have been defined before imposing the constraint. Moreover, there are only two free parameters here, as the third one is constrained by the equation.

The following R code specifies a generic model with `mu` as the mean and `Vary` as the variance of y_i (lines 2 and 3). The line `Vary==data.vi+tau2*data.vi^omega` (line 5) implements $\Sigma_i(\boldsymbol{\theta}) = v_i + \tau^2 v_i^{\omega}$. However, the parameters (`tau2` and `omega`) have not yet been defined in the model. To resolve this, we need to add the `replace.constraints=TRUE` argument (lines 11 and 14) to tell the functions to replace the left-hand side parameter (`Vary`) with the right-hand side function and automatically create the two parameters (`tau2` and `omega`). With this trick, we may fit models with arbitrary functions on the mean and variance structure (see the next section). After fitting the model, the parameter estimates

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

and their 95% Wald CIs are $\hat{\mu} = 0.1869, [0.1491, 0.2248]$, $\hat{\tau}^2 = 0.0331, [-0.0733, 0.1394]$, and $\hat{\omega} = 0.1204, [-0.4562, 0.6971]$. The $\hat{\omega}$ appears closer to 0 (indicating an additive error model) than 1 (indicating a multiplicative error model).

```
## Hybrid model 2: multiplicative and additive error
m5 <- "yi ~ mu*1          ## Mean(yi) = mu
       yi ~~ Vary*yi      ## Var(yi) = vary
       ## Define Vary as a function of vi, tau2, and omega
       Vary == data.vi + tau2*data.vi^omega
"

ram5 <- lavaan2RAM(m5, obs.variables = "yi", std.lv = FALSE)

## replace.constraints = TRUE: Replace the constraint "Vary" with the new parameters
impliedS(ram5, replace.constraints = TRUE)

## replace.constraints = TRUE: Replace the constraint "Vary" with the new parameters
fit5 <- sem("Hybrid2", RAM = ram5, data = dat1, replace.constraints = TRUE)
summary(fit5)
```

Hybrid model 3: The third version of the hybrid model is the HKSJ model.[41–43] These authors recognized that the average effect size's *SE* does not perform well under the random-effects model when the number of studies is small, primarily because the heterogeneity variance τ^2 and sampling variance v_i are estimated rather than known. They proposed adjusting the calculated *SE*, which has been shown to perform better than the random-effects model when there are only a small number of studies.[44–46] At first glance, it may not be apparent how their adjustment can be considered a hybrid model of multiplicative and additive error. To understand this, we can compare the multiplicative error model with the fixed-effect model. In the multiplicative error model, the term ϕv_i in Equation (3.5) is proportional to v_i in the fixed-effect model (Equation (3.1)). This proportionality means that the estimated common effects μ are identical in both models, but they differ in their *SE*s due to the scaling factor ϕ. Conventionally, the multiplicative error models are estimated with UWLS, which does not impose $\phi = 1$.[31,32]

van Aert and Jackson[35] noted the similar relationship between the HKSJ and the random-effects models and suggested estimating the HKSJ model with UWLS. Implementing the HKSJ method involves two steps. First, a random-effects model is fitted to the data. This step produces an estimate of the heterogeneity variance, $\tilde{\tau}^2$, which, combined with the known sampling variance v_i, results in the total error variance $\tilde{v}_i = \tilde{\tau}^2 + v_i$ in the ith study. The HKSJ model is

$$\begin{aligned} y_i &= \mu + \tilde{e}_i, \\ \mathrm{Var}\,(\tilde{e}_i) &= \phi(\tilde{\tau}^2 + v_i) = \phi\tilde{v}_i. \end{aligned} \tag{3.11}$$

And the model-implied variance and mean in the ith study are

$$\begin{aligned} \mu_i(\boldsymbol{\theta}) &= \mu, \\ \Sigma_i(\boldsymbol{\theta}) &= \phi(\tilde{\tau}^2 + v_i). \end{aligned} \tag{3.12}$$

The random-effects model assumes that the y_i is distributed as $\tilde{v}_i$. On the other hand, the HKSJ model considers that y_i is only distributed proportional to $\tilde{v}_i$, that is, $\phi\tilde{v}_i$. Thus, the average effect is identical under the random-effects and HKSJ models, whereas the *SE*s differ unless $\phi = 1$. The HKSJ-like multiplicative adjustment can be implemented as a structural equation model with the following R code. This implementation represents the heterogeneity effect and sampling error as two latent variables, u_i and e_i. Their impact on the observed effect size y_i is scaled by the factor `phi_sqrt`,

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

Table 2. *Results of six meta-analytic models.*

Name	Fixed	Random	Multiplicative	Hybrid1	Hybrid2	Hybrid3
`mu`	0.1938	0.1866	0.1938	0.1874	0.1869	0.1866
`tau2`	NA	0.0170	NA	0.0137	0.0331	0.0170
`phi_sqrt`	NA	NA	2.3588	1.2870	NA	1.0086
`omega`	NA	NA	NA	NA	0.1204	NA
`mu (SE)`	0.0079	0.0193	0.0187	0.0192	0.0193	0.0195
`tau2 (SE)`	NA	0.0041	NA	0.0066	0.0543	0.0041
`phi_sqrt (SE)`	NA	NA	0.2136	0.4615	NA	0.0913
`omega (SE)`	NA	NA	NA	NA	0.2942	NA
No. of parameters	1	2	2	3	3	2
−2LL	129.0697	−55.4423	−44.6262	−55.7740	−55.6036	−55.4513
AIC	131.0697	−51.4423	−40.6262	−49.7740	−49.6036	−51.4513
BIC	133.1806	−47.2205	−36.4045	−43.4414	−43.2710	−47.2296

Note: Fixed, Random, Multiplicative, Hybrid1, Hybrid2, and Hybrid3 represent the models in Equations (3.1), (3.3), (3.5), (3.7), (3.9), and (3.11), respectively. Names with `(SE)` are the standard errors of the parameter estimates. *−2LL* = −2*loglikelihood. AIC = Akaike information criterion. BIC = Bayesian information criterion.

as shown in lines 5 and 8 of the R code. The heterogeneity variance, obtained from the output of the random-effects model, is treated as a fixed value (0.01703407 in this example). After fitting the model, the estimated average effect $\hat{\mu}$ and its 95% Wald CI are 0.1866, with an interval of [0.1484, 0.2248]. The estimated scaling factor $\hat{\phi}$ and its 95% Wald CI are 1.0173, with an interval of [0.6563, 1.3784]. It is important to note that this model has no estimate for τ^2, as it is treated as a fixed value. Researchers can use the value and confidence interval from the random-effects model output for completeness.

```
## Hybrid model 3: Hartung-Knapp-Sidik-Jonkman method
m6 <- "yi ~ mu*1              ## Mean(yi) = mu
       yi ~~ 0*yi             ## Set the measurement error variance of yi at 0
       ## additive error
       ui =~ phi_sqrt*yi      ## phi is a scaling factor
       ui ~~ 0.01703407*ui    ## Estimated tau2 is obtained from the random-effects model
       ## multiplicative error
       ei =~ phi_sqrt*yi      ## phi is a scaling factor
       ei ~~ data.vi*ei       ## Var(ei) = vi
       phi := phi_sqrt^2      ## Define phi as a function of phi_sqrt
"
```

Table 2 summarizes the results of all six models. The estimate of heterogeneity variance ($\hat{\tau}^2$) and its *SE* for the hybrid 3 (HKSJ) model were copied from those of the random-effects model. Although the hybrid 3 model has the lowest AIC and BIC values, these values are likely incorrect because they do not account for the estimation of τ^2. For instance, these values were calculated based on two parameters (`mu` and `phi_sqrt`), while τ^2 was not considered in the model complexity.

By inspecting the table, we can confirm that the estimated average effect ($\hat{\mu}$) is identical in both the fixed-effect and multiplicative error models. However, the *SE* of the multiplicative error model is larger than that of the fixed-effect model because $\sqrt{\hat{\phi}} > 1$. The same pattern is observed when comparing the random-effects and HKSJ models. The *SE* is larger in the HKSJ model due to the adjustment factor ϕ. When we compare the AIC values, it becomes evident that the fixed-effect model is the worst fit for the data, unlike the other models. The random-effects model and the three hybrid models all show similar good fits. However, it is crucial to remember that this is not a definitive assessment of their statistical

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

performance, as there is only one sample. Future research may explore the advantages and limitations of these models.

3.2. Models with external variables

This section examines cases involving external variables, including study characteristics. These external variables are usually considered predictors of the effect sizes. However, in some cases, external variables may also be the outcome variables predicted by the effect size. In the illustrations, we use the individualism score of the countries where the studies were conducted, denoted as x_i, as the predictor of the effect size, and the number of citations of the studies, denoted as z_i, as the outcome predicted by the effect size.

3.2.1. A fixed-X meta-regression

The most common model is to use a covariate to predict the effect size in cases of excess heterogeneity among studies. This approach is called a mixed-effects meta-analysis or meta-regression. Similar to regression analysis, predictors in this context can be seen as either fixed or random. In a fixed-X approach, predictors are considered fixed values and are part of the design matrix. This approach is widely used in meta-analysis, as seen in works in the literature.[47–49] On the other hand, the SEM literature typically assumes a random-X approach, treating predictors as random variables. Both approaches have their own strengths and limitations. Treating x_i as fixed values has the advantage of not requiring any distributional assumptions about x_i. This makes the method versatile, as it can handle binary, categorical, and continuous covariates equally well. However, the primary limitation of this approach is that it does not accommodate missing values in the covariates; any missing data must be removed before analysis.

To better understand these two approaches, let us consider the fixed-X meta-regression, where the individualism score of the countries where the studies were conducted, x_i, serves as the covariate. The model is

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + u_i + e_i, \\ \mathsf{Var}\,(u_i) &= \tau^2, \\ \mathsf{Var}\,(e_i) &= v_i. \end{aligned} \tag{3.13}$$

The model-implied variance and mean of the ith study are

$$\begin{aligned} \mu_i(\theta) &= \beta_0 + \beta_1 x_i, \\ \Sigma_i(\theta) &= \tau^2 + v_i. \end{aligned} \tag{3.14}$$

Figure 3(a) presents the graphical model in SEM, which is implemented in the following R code. This model introduces a phantom variable η_i, a latent variable with a zero variance[50] to represent the covariate. Specifically, the value of x_i in the ith study is assigned to the specific path by the definition variable `data.xi` (lines 7–9). Consequently, the mean of y_i in the ith study is modeled as $\beta_0 + \beta_1 x_i$. The variance structure of the model closely resembles that of a random effects model, with the exception that τ^2 is now the residual heterogeneity variance after accounting for the covariate x_i. Since x_i is treated as a design matrix or specified fixed values, the analysis has only one observed variable y_i, as indicated on line 12.

```
## Mixed-effects model with the covariate as a design matrix
m7 <- "yi ~ b0*1          ## Intercept of yi: b0
       yi ~~ data.vi*yi   ## Known sampling variance vi
       ui =~ 1*yi         ## ui: random effects
       ui ~~ tau2*ui      ## Residual heterogeneity variance tau2
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

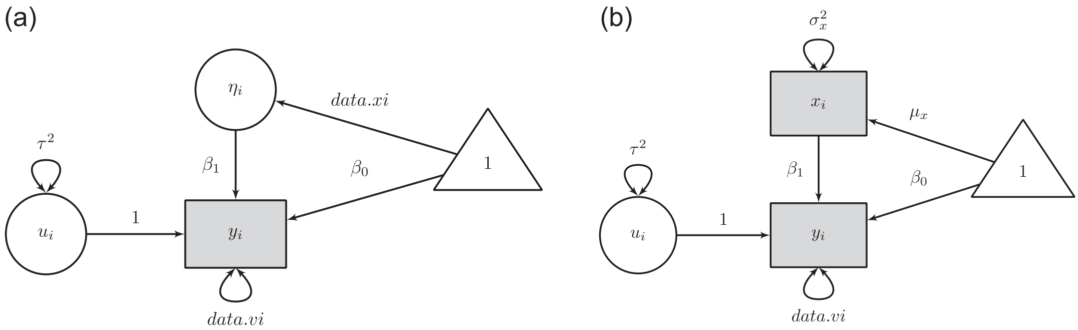


Figure 3. *Structural equation models of univariate meta-analysis with one study characteristic.*
Note: Figures (a) and (b) represent fixed-x and random-x mixed-effects meta-analyses, respectively. In these panels, y_i and $data.vi$ represent the effect size and its known sampling error variance in the ith study, respectively, while u_i and e_i denote the random effect and sampling error in the ith study, respectively. Additionally, μ represents the average effect, and τ^2 represents the random-effects (residual) heterogeneity variance.

```
      ## etai: a phantom latent variable to represent the covariate x
      etai =~ b1*yi     ## yi = b1*etai
      etai ~ data.xi*1  ## etai = xi
      etai ~~ 0*etai    ## Var(etai) = 0
"
## Only yi is an observed variable
ram7 <- lavaan2RAM(m7, obs.variables = "yi", std.lv = FALSE)
```

3.2.2. A random-X meta-regression

A standard practice in SEM is to treat x_i as another variable in the model. The random-X model of the ith study is shown below:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + u_i + e_i, \\ x_i &= \mu_x + e_{xi}, \\ \mathsf{Var}(u_i) &= \tau^2, \\ \mathsf{Var}(e_i) &= v_i, \\ \mathsf{Var}(e_{xi}) &= \sigma_x^2 \\ \mathsf{E}(x_i) &= \mu_x. \end{aligned} \tag{3.15}$$

The model-implied covariance matrix and the mean vector are

$$\begin{aligned} \boldsymbol{\mu}_i(\boldsymbol{\theta}) &= \begin{matrix} y_i \\ x_i \end{matrix} \begin{bmatrix} \beta_0 + \beta_1 \mu_x \\ \mu_x \end{bmatrix}, \\ \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) &= \begin{matrix} \\ y_i \\ x_i \end{matrix} \begin{matrix} \begin{matrix} y_i & \qquad\qquad x_i \end{matrix} \\ \begin{bmatrix} \beta_1^2 \sigma_x^2 + \tau^2 + v_i & \\ \beta_1 \sigma_x^2 & \sigma_x^2 \end{bmatrix} \end{matrix}. \end{aligned} \tag{3.16}$$

The main strength of treating x_i as another variable is its ability to handle missing effect sizes and covariates using FIML estimation without excluding any data,[51] for a discussion. However, a key limitation is that the data are assumed to follow a multivariate normal distribution. Another key distinction between the fixed-X and random-X approaches is how the explained variance R^2 is calculated. The fixed-X approach involves fitting two models: an unconditional model that excludes

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

covariates and a conditional model that incorporates the covariates. The R^2 is defined by[49]

$$R^2_{\text{fixed-x}} = \frac{\hat{\tau}_0^2 - \hat{\tau}_1^2}{\hat{\tau}_0^2}, \tag{3.17}$$

where $\hat{\tau}_0^2$ and $\hat{\tau}_1^2$ are the estimated heterogeneity variances without and with the covariates. In contrast, the random-X approach, commonly used in SEM, defines R^2 using only the conditional model. The formula is

$$R^2_{\text{random-x}} = \frac{\hat{\boldsymbol{\beta}}^{\mathrm{T}} \boldsymbol{S}_x \hat{\boldsymbol{\beta}}}{\hat{\boldsymbol{\beta}}^{\mathrm{T}} \boldsymbol{S}_x \hat{\boldsymbol{\beta}} + \hat{\tau}_1^2}, \tag{3.18}$$

where $\boldsymbol{S}_x$ is the sample covariance matrix of the covariates and $\hat{\boldsymbol{\beta}}$ is the vector of the regression coefficients.

The $R^2_{\text{random-x}}$ has several advantages over the $R^2_{\text{fixed-x}}$. First, the random-X approach retains all available data in the analysis in the presence of missing covariates, whereas the fixed-X approach must exclude studies with missing covariates. Lee and Beretvas[52] examined several methods for handling missing covariates in the meta-analysis, concluding that the random-X approach, using the FIML estimation, outperformed the listwise deletion approach under their simulation conditions. However, further commentary by White[53] and responses from Lee and Beretvas[54] highlight additional considerations and debates surrounding this issue. Second, the $R^2_{\text{fixed-x}}$ becomes negative when $\hat{\tau}_1^2 > \hat{\tau}_0^2$. Although $R^2_{\text{random-x}}$ can also be negative, it only occurs when $\hat{\tau}_1^2 < 0$, which is less likely than $\hat{\tau}_1^2 > \hat{\tau}_0^2$. Third, CIs for $R^2_{\text{random-x}}$ can be constructed using both Wald CI and LBCI methods, as it is defined as a function of other estimated parameters within the model. In contrast, constructing CIs for $R^2_{\text{fixed-x}}$ is less straightforward. One option is to apply a bootstrap method to the data. However, it may violate the assumption that the covariates are fixed values.

Figure 3(b) illustrates the SEM utilizing the random-X approach. In this model, x_i is treated as a variable characterized by its own mean and variance, as indicated in lines 6 and 7 of the following R code. The $R^2_{\text{random-x}}$ is defined in line 9. It is crucial to note that when converting the equations to the RAM specification, x_i must be included as a variable, as demonstrated in line 12. Upon fitting the models, it is observed that the results of both the fixed-X and random-X approaches are identical, as there is no missing data. The estimated coefficients, $\hat{\beta}_0$ and $\hat{\beta}_1$, are 0.1860 and −0.0013, respectively. The 95% CIs for these coefficients are [0.1479, 0.2239] for $\hat{\beta}_0$ and [−0.0033, 0.0006] for $\hat{\beta}_1$. Additionally, the $R^2_{\text{random-x}}$, along with its 95% LBCI, is reported as 0.0371 with a confidence interval of [0.0000, 0.1926].

```
1  ## Mixed-effects model with the covariate as a variable
2  m8 <- "yi ~ b0*1 + b1*xi ## Intercept (b0) and slope (b1)
3         yi ~~ data.vi*yi  ## Known sampling variance vi
4         ui =~ 1*yi        ## ui: random effects
5         ui ~~ tau2*ui     ## Residual heterogeneity variance tau2
6         xi ~ Meanx*1      ## Mean(xi) = Meanx
7         xi ~~ Varx*xi     ## Var(xi) = Varx
8         ## Define R2: explained variance/total variance
9         R2 := b1^2*Varx/(b1^2*Varx + tau2)
10 "
11 ## Both yi and xi are observed variables
12 ram8 <- lavaan2RAM(m8, obs.variables = c("yi", "xi"), std.lv = FALSE)
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

3.2.3. Regressing a covariate on a true effect size

In traditional meta-analysis, the effect size is typically treated as the outcome variable, while the observed variables are predictors. However, there are instances where these roles may be reversed, with the effect size being used to predict the observed variables. Several researchers, for example,[55,56], have been interested in exploring the relationship between effect size and the future impact of studies, measured in terms of citations. Since effect sizes are reported before the accumulation of citations, it is logical to use the effect size as a predictor in such analyses. The common practice involves conducting a regression analysis by regressing the number of citations on the effect sizes.[55] A significant issue arises in this context: each y_i is distributed with a known sampling variance v_i, which is usually ignored in the applications. It is well-documented in the literature that ignoring the precision in the predictors can lead to biased estimates of the regression coefficients and *SE*s.[57] Moreover, the calculated R^2 is also biased because it ignores the known sampling variance v_i.

This problem can be effectively addressed using SEM. In SEM, the external variable z_i can be regressed on the *true* effect size f_{yi} rather than the observed effect size y_i. The following equations depict the model where the number of citations z_i is regressed on the true effect size f_{yi}:

$$\begin{aligned} y_i &= f_{yi} + e_i, \\ z_i &= \beta_0 + \beta_1 f_{yi} + e_{zi}, \\ \mathsf{Var}\,(e_i) &= v_i, \\ \mathsf{E}\,(f_{yi}) &= \mu_y, \\ \mathsf{Var}\,(f_{yi}) &= \tau^2, \\ \mathsf{Var}\,(e_z) &= \sigma_z^2. \end{aligned} \tag{3.19}$$

The model-implied variance and mean of the *i*th study are

$$\begin{aligned} \boldsymbol{\mu}_i(\boldsymbol{\theta}) &= \begin{matrix} y_i \\ z_i \end{matrix}\begin{bmatrix} \mu_y \\ \beta_0 + \beta_1\mu_y \end{bmatrix}, \\ \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) &= \begin{matrix} y_i \\ z_i \end{matrix}\begin{bmatrix} \tau^2 + v_i & \\ \beta_1\tau^2 & \beta_1^2\tau^2 + \sigma_z^2 \end{bmatrix}. \end{aligned} \tag{3.20}$$

Figure 4 illustrates the graphical model where the true effect size is conceptualized as a latent variable f_{yi}. This latent variable is then used to predict the number of citations z_i. The following R code is employed to implement the model. To reduce the need for extensive typing, only the true effect size f_{yi} is explicitly defined as a latent variable in line 2. In contrast, the sampling error e_i is not explicitly specified. Instead, the known sampling variance v_i is assigned to the observed effect size y_i using the syntax `yi~~data.vi*yi` in line 6. The R^2, which does not involve v_i, is defined on line 10. The results from the model indicate that the parameter regression coefficient $\hat{\beta}_1$ is estimated to be -1.9755, with the $SE = 1.0733$. The R^2 value is estimated to be 0.0672, with a 95% Wald CI of [0.0000, 0.2043] after truncating the negative estimated lower bound to zero. In contrast, the result of the wrong approach by regressing the outcome on the sample effect size is $\hat{\beta}_1 = -1.4114$ with the $SE = 0.8051$ and an R^2 value of 0.0495.

```
1  ## Regressing a covariate on the true effect size
2  m9 <- "fyi =~ 1*yi          ## fyi: true effect size
3         fyi ~~ tau2*fyi      ## Var(fyi) = tau2
4         fyi ~ Meany*1        ## Mean(fyi) = Meany
5         yi ~ 0*1
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

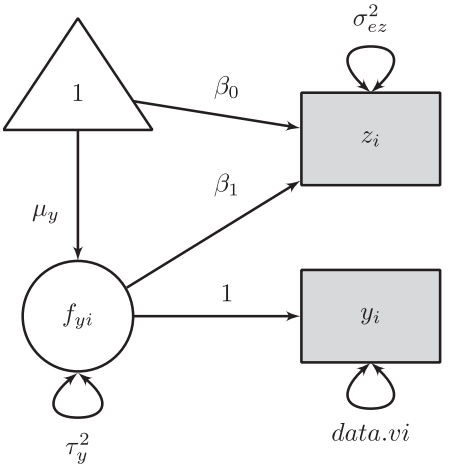


Figure 4. *A structural equation model of regressing one study characteristic z_i on the true effect size f_{yi}.*
Note: y_i and $data.vi$ represent the effect size and its known sampling error variance in the *i*th study, respectively. f_{yi} and z_i represent the true effect size and the outcome variable in the *i*th study, respectively. Furthermore, μ_y signifies the average effect, while τ^2_y indicates the variance of random-effects heterogeneity.

```
6         yi ~~ data.vi*yi    ## Var(ei) = vi
7         zi ~ b0*1 + b1*fyi  ## Intercept (b0) and slope (b1)
8         zi ~~ ErrVarz*zi    ## Error variance of zi
9         ## Define an R2
10        R2 := b1^2*tau2/(b1^2*tau2 + ErrVarz)
11  "
```

3.2.4. Meta-regression with partially standardized covariates

When the predictors' scales differ, interpreting regression coefficients using raw scores (unstandardized solutions) can be challenging.[58] For example, our previous meta-regression analysis noted that when the individualism score (x_i) of the countries in the studies increases by one unit, the expected effect size y_i decreases by 0.0013. Although this change is not statistically significant, we will continue to use it in our discussion. The standard deviation of x_i is 19.1356, making it difficult to contextualize the value.

Standardized coefficients are typically preferred, especially when predictors are on different scales. Standardization facilitates comparison and interpretation by ensuring that all variables have the same variance, thereby allowing for more accurate comparison and interpretation. There are two types of standardization: complete and partial. Complete standardization involves standardizing the predictors and the outcome variable, while partial standardization only standardizes the predictors. In meta-analysis, partial standardization is preferred because it maintains the scale of the effect size, which is essential for indicating residual heterogeneity.

A naive approach might involve manually standardizing the variables before conducting a meta-regression. However, this method is known to produce biased *SE*s because it does not consider the uncertainty of the standardization process.[59,60] A more accurate approach is to treat the standardized variables as latent variables within the SEM framework.[61] This method properly incorporates the uncertainty and provides more accurate *SE*s.

Figure 5(a) and the subsequent equations depict a model for a partially standardized solution. This model has two latent variables, f_{yi} and f_{xi}, representing the unstandardized true effect size and partially standardized predictor, respectively. The standard deviation of the predictor is represented by its factor loading σ_{xi} (`sd_xi` in line 8), whereas the variance of f_{xi} is fixed at 1 (line 9). The residual variance of

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

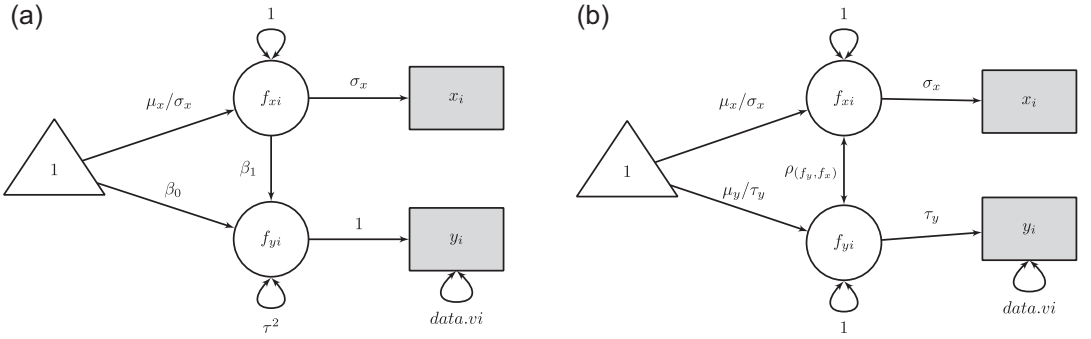


Figure 5. *Structural equation models of partially standardized meta-regression (a) and correlation between a true effect size and a covariate (b).*

Note: Figures (a) and (b) illustrate models with partially standardized effects and correlation between an effect size and a covariate. In these panels, y_i and $data.vi$ represent the effect size and its known sampling error variance in the ith study, respectively, while u_i and e_i denote the random effect and sampling error in the ith study, respectively. Additionally, μ represents the average effect, and τ^2 represents the random-effects (residual) heterogeneity variance.

the true effect size is represented by τ^2 (`tau2_res` in line 3), which is used to define the R^2 in line 14. After fitting the partially standardized solution, the estimated R^2 is 0.0371. The lower bound of the 95% Wald CI is truncated to zero, whereas the upper bound is 0.1425:

$$\begin{aligned} y_i &= f_{yi} + e_i, \\ x_i &= \sigma_x f_{xi}, \\ f_{yi} &= \beta_0 + \beta_1 f_{xi} + u_i, \\ \mathrm{Var}\,(e_i) &= v_i, \\ \mathrm{Var}\,(u_i) &= \tau^2_{f_y}, \\ \mathrm{Var}\,(f_{xi}) &= 1, \\ \mathrm{E}(f_{yi}) &= \beta_0 + \beta_1 \mu_x/\sigma_x, \\ \mathrm{E}\,(x_i) &= \mu_x/\sigma_x. \end{aligned} \tag{3.21}$$

The model-implied variance and mean of the ith study are

$$\begin{aligned} \boldsymbol{\mu}_i(\boldsymbol{\theta}) &= \begin{matrix} y_i \\ x_i \end{matrix}\begin{bmatrix} \beta_0 + \beta_1\mu_x/\sigma_x \\ \mu_x \end{bmatrix}, \\ \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) &= \begin{matrix} \\ y_i \\ x_i \end{matrix}\begin{matrix} \begin{matrix} y_i & x_i \end{matrix} \\ \begin{bmatrix} \beta_1^2 + \tau^2_{f_y} + v_i & \\ \beta_1\sigma_x & \sigma_x^2 \end{bmatrix} \end{matrix}. \end{aligned} \tag{3.22}$$

```
1  ## Partial standardized xi but not yi
2  m10 <- "fyi =~ 1*yi          ## fyi: Latent effect size of yi
3          fyi ~~ tau2_res*fyi  ## tau2_res: residual heterogeneity variance
4          fyi ~ b0*1 + b1*fxi  ## b0: intercept, b1: slope
5          yi ~ 0*1
6          yi ~~ data.vi*yi     ## Var(ei) = vi
7          ## fxi: Standardized score of xi
8          fxi =~ sd_xi*xi      ## sd_xi: sd of xi
9          fxi ~~ 1*fxi         ## Var(fxi) =1 for standardization
10         fxi ~ mu_x_sd*1
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

```
        xi ~ 0*1
        xi ~~ 0*xi
        ## Define R2
        R2 := b1^2/(b1^2 + tau2_res)
"
```

3.2.5. Correlating a covariate with a true effect size

The correlation coefficient is crucial in statistics. It is the cornerstone of regression analysis, SEM, and many multivariate statistics. It provides a quick understanding of how two variables are linearly related to one another, and it is a simple tool for detecting potential multicollinearity among predictors, which can be problematic in regression models. As a result, it is a common practice to present a correlation table of the variables in academic papers, offering a snapshot of the linear relationships between them.

In the context of meta-analyses, correlations between effect sizes and covariates are sometimes reported. However, researchers often calculate these correlations by correlating the sample effect sizes y_i and the covariates.[55,56,62] Similarly, researchers sometimes calculate the correlation between two effect sizes by ignoring their sampling error.[63] As discussed in the previous section, this approach can introduce bias into the estimated correlations and their *SE*s.[64] The presence of sampling variances v_i can distort the true relationships between variables, leading to inaccurate and potentially misleading conclusions. It is, therefore, essential to account for these variances to obtain more reliable and valid correlation estimates.

Figure 5(b) and the following equations define the model to measure the correlation between the true effect size f_{yi} and a covariate x_i. The variances of the standardized f_{yi} and *latent* covariate f_{xi} are fixed at one, where τ_{f_y} and σ_x represent the standard deviations of true effect size and x_i. Then the covariance between f_{yi} and f_{xi} is the required correlation $\rho_{(x,f_y)}$. It should be noted that the means of f_{yi} and f_{xi} are also scaled by $1/\tau_{f_y}$ and $1/\sigma_x$. The model is

$$\begin{aligned} y_i &= \tau_{f_y} f_{yi} + e_i, \\ x_i &= \sigma_x f_{xi}, \\ \mathsf{Var}\,(e_i) &= v_i, \\ \mathsf{Var}\,(f_{yi}) &= \mathsf{Var}\,(f_{xi}) = 1, \\ \mathsf{Cor}\,(f_{yi}, f_{xi}) &= \rho_{(x,f_y)} \\ \mathsf{E}\,(f_{yi}) &= \mu_y/\tau_{f_y}, \\ \mathsf{E}\,(f_{x_i}) &= \mu_x/\sigma_x. \end{aligned} \tag{3.23}$$

The model-implied covariance matrix and mean vector of the *i*th study are

$$\begin{aligned} \boldsymbol{\mu}_i(\boldsymbol{\theta}) &= \begin{matrix} y_i \\ x_i \end{matrix}\begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix}, \\ \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) &= \begin{matrix} \\ y_i \\ x_i \end{matrix}\begin{matrix} \begin{matrix} y_i & \quad x_i \end{matrix} \\ \begin{bmatrix} \tau_{f_y}^2 + v_i & \\ \rho_{(x,f_y)}\tau_{f_y}\sigma_x & \sigma_x^2 \end{bmatrix} \end{matrix}. \end{aligned} \tag{3.24}$$

The following R code implements this model. The key idea is to define two standardized latent variables (`fyi` and `fxi`) to represent the standardized effects of the true effect and covariate. The corresponding factor loadings (`tauy` and `sd_xi`) now represent the standard deviations of the variables (lines 3 and 9). After fitting the model, the estimated correlation and its 95% Wald CI are $\hat{\rho} = -0.1925, [-0.4667, 0.0816]$, which is not statistically significant at 0.05. If we disregard the known

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

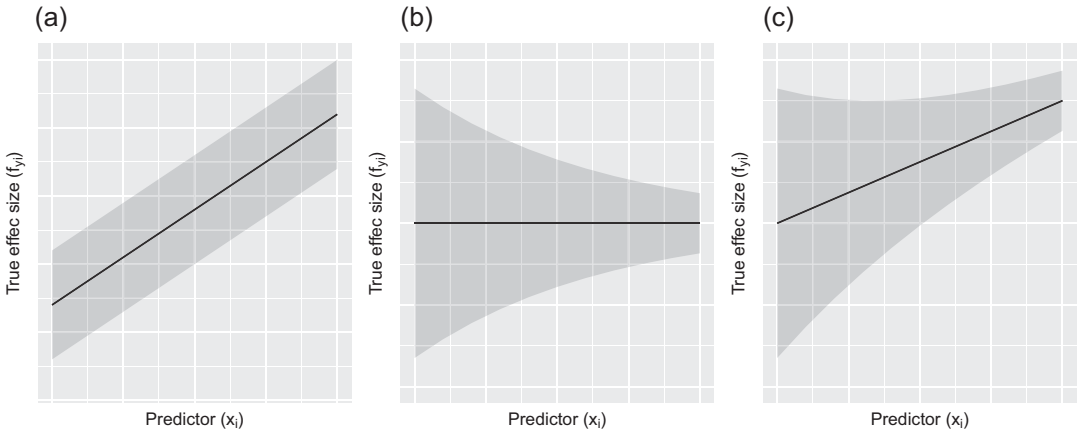


Figure 6. *Some meta-regression models.*
Note: Figure (a) shows a standard mixed-effects model on the true effect size (f_{y_i}) with $f_{yi} = \beta_0 + \beta_1 x_i + u_i$ and $\mathsf{Var}(u_i) = \tau^2$. Figure (b) shows a location-scale model on the heterogeneity variance with $f_{yi} = \mu + u_i$ and $\mathsf{Var}(u_i) = \exp(\alpha_0 + \alpha_1 x_i)$. Panel (c) shows a location-scale model on both the mean and the heterogeneity variance with $f_{yi} = \beta_0 + \beta_1 x_i + u_i$ and $\mathsf{Var}(u_i) = \exp(\alpha_0 + \alpha_1 x_i)$. The shaded areas represent the 95% predictive intervals of the true effect size.

sample variance of the effect size and correlate the sample effect size with the covariate, the estimated correlation is −0.1734, with a 95% Wald CI of [−0.4074, 0.0820].

```
## Correlation between fyi and fxi
m11 <- "## fyi: Standardized true effect size of yi
        fyi =~ tauy*yi     ## tauy: sd of fyi
        fyi ~~ 1*fyi       ## Var(fyi) = 1
        fyi ~ y_mean_div_tauy*1   ## rescaled mean
        yi ~ 0*1
        yi ~~ data.vi*yi
        ## fxi: Standardized score of xi
        fxi =~ sd_xi*xi   ## sd_xi: sd of xi
        fxi ~~ 1*fxi      ## Var(fxi) = 1 for standardization
        fxi ~~ Cor*fyi    ## Correlation between fyi and fxi
        fxi ~ x_mean_div_sigmax*1 ## rescaled mean
        xi ~ 0*1
        xi ~~ 0*xi
"
```

3.2.6. Location-scale and nonlinear models

Most meta-regression models in the literature are linear models. Figure 6(a) illustrates a linear relationship between the predictor x_i and the true effect size f_{yi}. The shaded area represents the 95% predictive intervals (PIs), which approximately include 95% of the true effect sizes. While nonlinear relationships, such as quadratic or cubic terms, can still be accommodated within linear models by including polynomial terms, there are cases where genuinely nonlinear models may be required.[65–69] This section addresses these advanced models.

The location-scale models have been utilized in multilevel models.[70] and SEM[71] Recently, Viechtbauer and López-López[23] extended these models for meta-analysis. The core idea is to model the mean and variance structures separately. For example, Figure 6(b) depicts a model with an average effect where the heterogeneity decreases as x_i increases. Conversely, Figure 6(c) illustrates that x_i has a positive relationship with the expected mean but a negative relationship with the residual heterogeneity variances. These models are similar to the regression models with heteroskedasticity.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

In the location-scale model with additive error, the model-implied variance and mean for the *i*th study are as follows:

$$\begin{aligned} \mu_i(\theta) &= \beta_0 + \beta_1 x_i, \\ \Sigma_i(\theta) &= \exp(\alpha_0 + \alpha_1 x_i) + v_i, \end{aligned} \tag{3.25}$$

where $\tau_0^2 = \exp(\alpha_0)$ is the residual heterogeneity variance when $x_i = 0$. An exponential function is typically employed to ensure that the predicted residual heterogeneity variance is nonnegative. The mean structure is identical to that in a standard meta-regression. The intriguing aspect is the variance structure. Although we use x_i in this example's mean and variance structures, the predictors can be different for these two structures.

The following R code implements the location-scale model in Equation (3.25). Lines 2–6 define a generic random-effects model with the average effect `mu` and residual heterogeneity variance `tau2`. Lines 8 and 9 define the linear relationship for `mu` and the nonlinear relationship for `tau2`. `b0` (β_0) and `b1` (β_1) are the regression coefficients in predicting the mean structure `mu`, whereas `a0` (α_0) and `a1` (α_1) are the regression coefficients in predicting the variance structure `tau2`. Since we aim to replace the parameters on the left-hand side with the expressions on the right-hand side, we need to add the `replace.constraints = TRUE` argument in the analysis (see the discussion in Section 3.1.4). After fitting the model, neither regression coefficient is statistically significant in predicting the mean and variance structures. The parameter estimates and their 95% Wald CIs are ($\hat{\beta}_1 = -0.0013, [-0.0031, 0.0005]$) and ($\hat{\alpha}_1 = 0.0054, [-0.0204, 0.0312]$). The results are identical to those in the `metafor` package.

```
1  ## Location-scale model with an additive heterogeneity variance
2  m12 <- "yi ~ mu*1          ## Mean(yi) = mu
3          yi ~~ data.vi*yi   ## Var(ei) = vi
4          ## ui is the latent variable of the random effect
5          ui =~ 1*yi         ## yi = 1*ui
6          ui ~~ tau2*ui      ## Var(ui) = tau2 of the residual heterogenenity variance
7          ## Specify a linear relationship on the mean structure
8          mu == b0 + b1*data.xi
9          ## Specify a nonlinear relationship on the variance structure
10         tau2 == exp(a0 + a1*data.xi)
11 "
```

Since the location-scale model has been implemented in the `metafor` package, some readers may wonder about the necessity of implementing their own model given this existing functionality. The present SEM approach offers the flexibility to specify a wide range of arbitrary models tailored to specific research needs. For example, the location-scale model with two or more effect sizes can be easily accommodated within the SEM-based meta-analysis framework. Additionally, we can also explore models that have not yet been implemented in current software. For instance, to my knowledge, the location-scale model has not been extended to incorporate multiplicative error or hybrid models (cf. the study[34] for other ideas). As an example, the following model-implied variance and mean in the *i*th study shows a location-scale model with a multiplicative error:

$$\begin{aligned} \mu_i(\theta) &= \beta_0 + \beta_1 x_i, \\ \Sigma_i(\theta) &= \exp(\alpha_0 + \alpha_1 x_i) v_i, \end{aligned} \tag{3.26}$$

where $\phi_0 = \exp(\alpha_0)$ is the heterogeneity scaling factor when $x_i = 0$. The implementation of the location-scale model with a multiplicative error is shown below. After fitting the model, the predictor is statistically significant in predicting the mean structure ($\hat{\beta}_1 = -0.0026$, 95%

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

Wald CI = [−0.0047, −0.0005]) but not for the variance structure ($\hat{\alpha}_1 = -0.0110$, 95% Wald CI = [−0.0313, 0.0093]). In theory, we can also create a location-scale model for the hybrid model that incorporates both multiplicative and additive heterogeneity. Still, it is yet to be determined what the interpretations of the parameters are due to the large number of parameters involved.

```
## Location-scale model with a multiplicative heterogeneity variance
m13 <- "yi ~ mu*1          ## Mean(yi) = mu
        yi ~~ sigma2*yi    ## Var(yi) = sigma2
        ## Specify nonlinear on mean and variance structures
        mu == b0 + b1*data.xi
        sigma2 == exp(a0 + a1*data.xi)*data.vi
"
```

4. Multivariate meta-analyses

In this section, we expand the univariate meta-analysis to a multivariate meta-analysis. Multivariate meta-analysis has been developed and applied for decades.[72–74] It is commonly used when a study contributes more than one conceptually distinct effect.[75] For instance, there may be multiple treatment groups but only one control group in intervention or experimental studies. The effect sizes calculated from these multiple treatment groups are correlated because they all share the same control group.[76] This sampling correlation needs to be considered when analyzing these effect sizes. In observational studies, correlation matrices are a natural form of multivariate meta-analysis.[77] These matrices allow researchers to analyze the relationships between multiple variables simultaneously.

We use a sample dataset from Chan et al.[78] to demonstrate models with a bivariate meta-analysis. These authors investigated the misinformation and the debunking effects. The primary studies included in the meta-analysis involved three groups of participants: (1) those exposed to a misinformation message, (2) those exposed to both the misinformation message and a debunking message, and (3) control participants who received neither message. To illustrate, we selected the standardized mean difference (Hedges' g) for two effects: the misinformation effect (the standardized difference between groups 1 and 3) and the debunking effect (the standardized difference between groups 2 and 1). As group 1 was involved in calculating both effect sizes, their sampling errors were correlated.[76] For convenience, we refer to the effects of misinformation and debunking as y_{1i} and y_{2i}, respectively. Additionally, we use the year of publication x_i as a covariate in our illustrations.

The original study comprised 52 independent samples, each nested within one of eight research reports. We only selected two effect sizes for this illustration and excluded studies comparing groups 2 and 3. As a result, the final sample for our illustration comprises 34 independent samples, each nested within one of six research reports. It is worth noting that the data and analyses differ between Chan et al.'s original study and our current tutorial. Our focus is on demonstrating the techniques rather than providing substantive interpretations. Readers should refer to Chan et al.[78] for more detailed analyses and interpretations.

4.1. Bivariate random-effects model by regressing a true effect on another true effect

Researchers sometimes aim to use one effect size to predict another effect size. In medical research, researchers investigate the effect of baseline risk (an effect size) on treatment benefit (another effect size).[79–83] Similar examples are also found in psychology. For instance,[84] examined the follow-up effect size of intervention effects on social-emotional and cognitive skills in children and adolescents. These authors used the immediate treatment effect to predict the follow-up treatment effect. Another example is Chan and Albarracín,[85] a follow-up study of Chan et al.,[78] which aimed to study the debunking effect while controlling for the misinformation effect. Researchers typically test these hypotheses using

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

meta-regression, treating one effect size as a predictor while ignoring the known sampling variance of the predictor. As discussed in this article, this approach can lead to incorrect parameter estimates and *SE*s.

Figure 7 presents the structural equation model in Equation (4.1) that can correctly analyze such data. The central idea is to formulate the true effect sizes as two latent variables ($f_i^{(1)}$ and $f_i^{(2)}$). This allows for the straightforward inclusion of predictive relationships in the model:

$$
\begin{aligned}
\underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_i}_{\boldsymbol{y}_i} &= \underbrace{\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}_i}_{\boldsymbol{f}_i} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_i}_{\boldsymbol{e}_i}, \\
\underbrace{\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}_i}_{\boldsymbol{f}_i} &= \begin{bmatrix} 0 & 0 \\ \beta_1^{(2)} & 0 \end{bmatrix} \underbrace{\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}_i}_{\boldsymbol{f}_i} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_i}_{\boldsymbol{u}_i} \\
\mathsf{Cov}\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_i\right) &= \underbrace{\begin{bmatrix} \tau_{y1,y1} & \\ 0 & \tau_{y2,y2} \end{bmatrix}}_{\boldsymbol{T}}, \\
\mathsf{Cov}\left(\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_i\right) &= \underbrace{\begin{bmatrix} v_{y1,y1} & \\ v_{y1,y2} & v_{y2,y2} \end{bmatrix}}_{\boldsymbol{V}_i}, \\
\mathsf{E}\left(\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}_i\right) &= \underbrace{\begin{bmatrix} \mu^{(1)} \\ \beta_0^{(2)} + \mu^{(1)}\beta_1^{(2)} \end{bmatrix}_i}_{\boldsymbol{\mu}_i}.
\end{aligned}
\tag{4.1}
$$

The model-implied mean vector and covariance matrix in the ith study are

$$
\begin{aligned}
\boldsymbol{\mu}_i(\boldsymbol{\theta}) &= \boldsymbol{\mu}_i, \\
\boldsymbol{\Sigma}_i(\boldsymbol{\theta}) &= \begin{array}{c} \\ y_1 \\ y_2 \end{array} \begin{array}{c} \begin{array}{cc} y_1 & \quad\quad y_2 \end{array} \\ \begin{bmatrix} \tau_{y1,y1} & \\ \beta_1^{(2)}\tau_{y1,y1} & (\beta_1^{(2)})^2\tau_{y1,y1} + \tau_{y2,y2} \end{bmatrix} \end{array} + \boldsymbol{V}_i.
\end{aligned}
\tag{4.2}
$$

The following R code implements this model. There are two important points worth discussing. First, we are estimating the means of the true effect sizes ($f_i^{(1)}$ and $f_i^{(2)}$) (lines 9 and 10) while the means of the observed effect sizes ($y_i^{(1)}$ and $y_i^{(2)}$) are fixed at 0 in line 3. Since we are regressing the true effect sizes, estimating the means of these true effect sizes is essential. Second, there is no `T_21` in the (residual) heterogeneity covariance matrix (lines 19 and 20) because we have included the regression coefficient $\beta_1^{(2)}$ in the model (line 22). After fitting the model, the estimated coefficients and their 95% Wald CIs are $\hat{\beta}_0^{(2)} = -0.3621, [-1.0554, 0.3312]$ and $\hat{\beta}_1^{(2)} = 0.6836, [0.4361, 0.9310]$. The estimated R^2 is 0.7989, [0.5850, 1.0000], where the upper bound is truncated to 1.00.

```
1  ## Prepare a dataset
2  dat2 <- data.frame(y1i=Chan17$g_misinfo,
3                     y2i=Chan17$g_debunk,
4                     V11i=Chan17$v_misinfo,
5                     V21i=Chan17$c_mis_deb,
6                     V22i=Chan17$v_debunk)
7
```

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

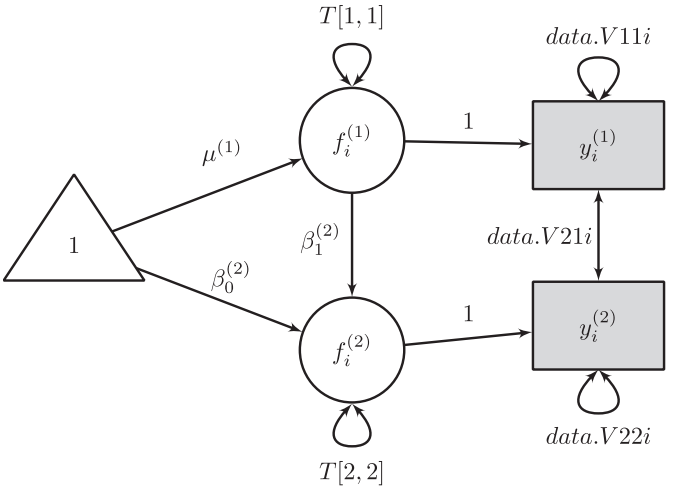


Figure 7. *A structural equation model of a bivariate random-effects meta-analysis by regressing a true effect on another true effect.*

Note: $y_i^{(1)}$ and $y_i^{(2)}$ are two effect sizes in the ith study. The terms $u_i^{(1)}$ and $u_i^{(2)}$, as well as $e_i^{(1)}$ and $e_i^{(2)}$, represent the two random effects and sampling errors in the ith study, respectively. The matrices $\boldsymbol{T}$ and $\boldsymbol{V}_i$ (specifically $data.v11i$, $data.v21i$, and $data.v22i$) are the 2×2 heterogeneity covariance matrix of the random effects and the known sampling error covariance matrix, respectively. $\mu^{(1)}$ is the average effect of $f_i^{(1)}$. $\beta_0^{(2)}$ and $\beta_1^{(2)}$ are the intercept and slope in regressing $f_i^{(2)}$ on $f_i^{(1)}$.

```
## Regressing true effect sizes f2i on f1i
m14 <- "## Means of y1 and y2 are fixed at 0
        y1i + y2i ~ 0*1
        ## Known sampling variances and covariance
        y1i ~~ data.V11i*y1i
        y2i ~~ data.V22i*y2i
        y1i ~~ data.V21i*y2i
        ## f1i and f2i: True effect sizes
        f1i =~ 1*y1i
        f2i =~ 1*y2i
        ## Heterogeneity variances
        f1i ~~ T_11*f1i
        f2i ~~ T_22*f2i
        ## Regressing f2i on f1i
        f2i ~ b0*1+ b1*f1i
        f1i ~ mu_1*1
        ## Define R2 on f2i
        R2 := b1^2*T_11/(b1^2*T_11 + T_22)
"
```

5. Discussion

This tutorial showed how various standard and advanced meta-analytic models can be expressed as structural equation models. The SEM-based meta-analysis offers a straightforward approach to defining models using equations, making it accessible to non-experts in meta-analysis. Once specified, these models can be converted into graphical models, which help understand the relationships between variables. Additionally, users can calculate the model-implied mean vector and covariance matrix, further enhancing their understanding of the fitted models. These components (equations, graphical models, and model-implied statistics) are valuable tools for users to comprehend the structure and

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

assumptions of their meta-analytic models. Most importantly, users can fit the proposed models to their data using the FIML estimation method, which is efficient when missing data are MAR. This tutorial also highlighted the flexibility of SEM by showing that users can define new models among effect sizes that may not currently be available in standard meta-analysis packages. This ability allows researchers to explore innovative modeling approaches tailored to their research questions.

Despite the many appealing features of the SEM approach, it has some limitations that should be considered. Firstly, only the FIML estimation method is readily available, although it is possible to specify REML for some models. However, creating these models is not straightforward.[29] Additionally, various estimators are used in meta-analysis (e.g., Hunter-Schmidt, Hedges, and DerSimonian-Laird),[27] and it is unlikely that they can be implemented within the SEM framework. Secondly, SEM uses a Wald or *z*-statistic to test parameter estimates, while most meta-analysis packages prefer a *t*-statistic, which generally performs better, especially in small sample sizes. Nevertheless, researchers can manually address this limitation by replacing the *z*-statistic with the *t*-statistic.[52]

Another limitation concerns the handling of non-independent effect sizes. While extending the SEM-based meta-analysis to accommodate non-independent effect sizes is relatively straightforward using multivariate meta-analysis,[9] other types of non-independence, such as nested effect sizes, pose additional challenges. A SEM-based three-level meta-analysis has been proposed and implemented in the `metaSEM` package,[10] but it is currently limited to only one effect size. Specifying three-level meta-analyses with multiple effect sizes using equations has yet to be implemented. Users who wish to conduct such analyses can still implement them using the `OpenMx` code.[86] Additionally, the popular robust variance estimation (RVE) method[87,88] in meta-analysis has not been implemented for SEM-based meta-analysis. Future research should explore these directions to expand the capabilities and applicability of SEM in meta-analysis.

In conclusion, this tutorial demonstrates that SEM-based meta-analysis is a flexible and powerful tool for learning about and conducting meta-analyses. It can supplement existing meta-analysis tools to address complex research questions. Despite its limitations, SEM offers unique advantages in specifying, visualizing, and fitting sophisticated meta-analytic models. The SEM-based meta-analysis holds great promise as a favorable framework for conducting future meta-analyses.

Acknowledgements. Part of this research was presented at the Annual Meeting of the Society of Research Synthesis Methodology in June 2024 in Amsterdam, the Netherlands, and at the Australasian Mathematical Psychology Conference in February 2026 in Singapore.

Author contributions. Conceptualization: M.W.-L.C.; Data curation: M.W.-L.C.; Data visualization: M.W.-L.C.; Methodology: M.W.-L.C.; Writing—original draft: M.W.-L.C.

Competing interests. The author declares that no competing interests exist.

Data availability statement. Complete examples and R code are available on Zenodo at https://doi.org/10.5281/zenodo.20962641.

Funding statement. This research was supported by a grant from the Singapore Ministry of Education Academic Research Fund Tier 1 (FY2021-FRC1-008).

Ethical standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Disclosure of use of AI tools. ChatGPT and Grammarly were used to check grammar and improve the writing.

References

[1] Rosenthal R, DiMatteo MR. Meta-analysis: recent developments in quantitative methods for literature reviews. *Annu Rev Psychol*. 2001;52: 59–82.
[2] Morris SB. Meta-analysis in organizational research: a guide to methodological options. *Annu Rev Organ Psych Organ Behav*. 2023;10: 225–259.
[3] Tanner-Smith EE, Grant S. Meta-analysis of complex interventions. *Annu Rev Public Health*. 2018;39: 135–151.
[4] Sutton AJ. *Methods for Meta-Analysis in Medical Research*. John Wiley; 2000.
[5] Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36: 1–48.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

[6] R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria, 2024. http://www.R-project.org/.
[7] Cheung MWL. Fixed-effects meta-analyses as multiple-group structural equation models. *Struct Equ Model Multidiscip J*. 2010;17: 481–509.
[8] Cheung MWL. A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychol Methods*. 2008;13: 182–202.
[9] Cheung MWL. Multivariate meta-analysis as structural equation models. *Struct Equ Model Multidiscip J*. 2013;20: 429–454.
[10] Cheung MWL. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol Methods*. 2014;19: 211–229.
[11] Cheung MWL. *Meta-Analysis: A Structural Equation Modeling Approach*. John Wiley & Sons, Inc.; 2015. https://doi.org/10.1002/9781118957813.
[12] Cheung MWL. Structural equation modeling (SEM)-based meta-analysis. In: Hoyle RH, ed. *Handbook of Structural Equation Modeling*. 2nd ed. Guilford Press, 2023: 664–680. https://www.guilford.com/books/Handbook-of-Structural-Equation-Modeling/Rick-Hoyle/9781462544646/contents.
[13] Boker S, Neale M, Maes H, et al. OpenMx: an open source extended structural equation modeling framework. *Psychometrika*. 2011;76: 306–317.
[14] Neale MC, Hunter MD, Pritikin JN, et al. OpenMx 2.0: extended structural equation and statistical modeling. *Psychometrika*. 2016;81: 535–549.
[15] Muthén BO, Muthén LK. *Mplus Users Guide*. 8th ed. Muthén & Muthén; 2017.
[16] Cheung MWL. metaSEM: an R package for meta-analysis using structural equation modeling. *Front Psychol*. 2015;5.
[17] Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48: 1–36.
[18] Epskamp S. semPlot: unified visualizations of structural equation models. *Struct Equ Model Multidiscip J*. 2015;22: 474–483.
[19] Cheung MWL. symSEM: symbolic computation for structural equation models. 2024. https://cran.r-project.org/package=symSEM.
[20] McArdle JJ, McDonald RP. Some algebraic properties of the reticular action model for moment structures. *Br J Math Stat Psychol*. 1984;37: 234–251.
[21] Neale MC, Miller MB. The use of likelihood-based confidence intervals in genetic models. *Behav Genet*. 1997;27: 113–120.
[22] Jaramillo F, Mulki JP, Marshall GW. A meta-analysis of the relationship between organizational commitment and salesperson job performance: 25 years of research. *J Bus Res*. 2005;58: 705–714.
[23] Viechtbauer W, López-López JA. Location-scale models for meta-analysis. *Res Synth Methods* 2022;13: 697–715. https://doi.org/10.1002/jrsm.1562.
[24] Bonett DG. Meta-analytic interval estimation for bivariate correlations. *Psychol Methods*. 2008;13: 173–181.
[25] Bonett DG, Price RM Jr. Varying coefficient meta-analysis methods for odds ratios and risk ratios. *Psychol Methods*. 2015;20: 394–406.
[26] Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Methods*. 1998;3: 486–504.
[27] Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J Educ Behav Stat*. 2005;30: 261–293.
[28] Veroniki AA, Jackson D, Bender R, et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. *Res Synth Methods*. 2019;10: 23–43.
[29] Cheung MWL. Implementing restricted maximum likelihood estimation in structural equation models. *Struct Equ Model Multidiscip J*. 2013;20: 157–167.
[30] Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med*. 1999;18: 2693–2708.
[31] Stanley TD, Doucouliagos H. Neither fixed nor random: weighted least squares meta-regression. *Res Synth Methods*. 2017;8: 19–42.
[32] Stanley TD, Doucouliagos H. Neither fixed nor random: weighted least squares meta-analysis. *Stat Med*. 2015;34: 2116–2127.
[33] Stanley TD, Ioannidis JPA, Maier M, Doucouliagos H, Otte WM, Bartoš F. Unrestricted weighted least squares represent medical research better than random effects in 67,308 Cochrane meta-analyses. *J Clin Epidemiol*. 2023;157: 53–58.
[34] Kulinskaya E, Olkin I. An overdispersion model in meta-analysis. *Stat Model*. 2014;14: 49–76.
[35] van Aert RCM, Jackson D. A new justification of the Hartung-Knapp method for random-effects meta-analysis based on weighted least squares regression. *Res Synth Methods*. 2019;10: 515–527.
[36] Stanley TD, Doucouliagos H, Havranek T. Reducing the biases of the conventional meta-analysis of correlations. *Res Synth Methods*. 2025;16: 42–59.
[37] Nam IS, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Stat Med*. 2003;22: 2309–2333.
[38] DuMouchel W. Bayesian metaanalysis. In: Berry D, ed. *Statistical Methodology in the Pharmaceutical Sciences*. CRC Press; 1989: 509–529.
[39] Baker RD, Jackson D. Meta-analysis inside and outside particle physics: two traditions that should converge? *Res Synth Methods*. 2013;4: 109–124.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

[40] Schmid CH. Heterogeneity: multiplicative, additive or both? *Res Synth Methods*. 2017;8: 119–120.
[41] Hartung J. An alternative method for meta-analysis. *Biom J*. 1999;41: 901–916.
[42] Hartung J, Knapp G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat Med*. 2001;20: 1771–1782.
[43] Sidik K, Jonkman JN. A simple confidence interval for meta-analysis. *Stat Med* 2002;21. https://doi.org/10.1002/sim.1262.
[44] Röver C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. *BMC Med Res Methodol*. 2015;15: 99.
[45] Lin L, Aloe AM. Evaluation of various estimators for standardized mean difference in meta-analysis. *Stat Med* 2021;40. https://doi.org/10.1002/sim.8781.
[46] IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol*. 2014;14: 25.
[47] Sera F, Armstrong B, Blangiardo M, Gasparrini A. An extended mixed-effects framework for meta-analysis. *Stat Med* 2019;38. https://doi.org/10.1002/sim.8362.
[48] Tipton E, Pustejovsky JE, Ahmadi H. A history of meta-regression: technical, conceptual, and practical developments between 1974 and 2018. *Res Synth Methods*. 2019;10: 161–179.
[49] Raudenbush SW. Analyzing effect sizes: random effects models. In: Cooper HM, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis and Meta-Analysis*. 2nd ed. Russell Sage Foundation; 2009: 295–315.
[50] Rindskopf D. Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika*. 1984;49: 37–47.
[51] Enders CK. *Applied Missing Data Analysis*. Guilford Press; 2010.
[52] Lee J, Beretvas SN. Comparing methods for handling missing covariates in meta-regression. *Res Synth Methods* 2023;14. https://doi.org/10.1002/jrsm.1585.
[53] White IR. The importance of plausible data-generating mechanisms in simulation studies: a response to 'comparing methods for handling missing covariates in meta-regression' by Lee and Beretvas (doi:10.1002/jrsm.1585). *Res Synth Methods* 2023;14. https://doi.org/10.1002/jrsm.1605.
[54] Lee J, Beretvas SN. Missingness mechanism messiness in meta-analysis: a response to "the importance of plausible data-generating mechanisms in simulation studies". *Res Synth Methods* 2023;14. https://doi.org/10.1002/jrsm.1603.
[55] Hanel PHP, Haase J. Predictors of citation rate in psychology: inconclusive influence of effect and sample size. *Front Psychol*. 2017;8.
[56] Lortie CJ, Aarssen LW, Budden AE, Leimu R. Do citations and impact factors relate to the real numbers in publications? A case study of citation rates, impact, and effect sizes in ecology and evolutionary biology. *Scientometrics*. 2013;94: 675–682.
[57] Fuller WA. *Measurement Error Models*, Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.; 1987. Visited May 15, 2018. https://doi.org/10.1002/9780470316665.
[58] Schmidt FL. Statistical and measurement pitfalls in the use of meta-regression in meta-analysis. *Career Dev Int*. 2017;22: 469–476.
[59] Yuan KH, Chan W. Biases and standard errors of standardized regression coefficients. *Psychometrika*. 2011;76: 670–690.
[60] Jones JA, Waller NG. Computing confidence intervals for standardized regression coefficients. *Psychol Methods*. 2013;18: 435–453.
[61] Cheung MWL. Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behav Res Methods*. 2009;41: 425–438.
[62] Barto EK, Rillig MC. Dissemination biases in ecology: effect sizes matter more than quality. *Oikos* 2012;121. https://doi.org/10.1111/j.1600-0706.2011.19401.x.
[63] Open Science Collaboration. Estimating the reproducibility of psychological. *Science*. 2015;349: aac4716.
[64] Schauer JM, Fitzgerald KG, Peko-Spicer S, Whalen MCR, Zejnullahi R, Hedges LV. An evaluation of statistical methods for aggregate patterns of replication failure. *Ann Appl Stat*. 2021;15: 208–229.
[65] Boucher M, Bennetts M. The many flavors of model-based meta-analysis: part I—introduction and landmark data. *CPT Pharmacometrics Syst Pharmacol*. 2016;5: 54–64.
[66] Boucher M, Bennetts M. Many flavors of model-based meta-analysis: part II – modeling summary level longitudinal responses. *CPT Pharmacometrics Syst Pharmacol*. 2018;7: 288–297.
[67] Rota M, Bellocco R, Scotti L, et al. Random-effects meta-regression models for studying nonlinear dose–response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Stat Med*. 2010;29: 2679–2687.
[68] Bleidorn W, Schwaba T, Zheng A, et al. Personality stability and change: a meta-analysis of longitudinal studies. *Psychol Bull*. 2022;148: 588–619.
[69] Breit M, Scherrer V, Tucker-Drob EM, Preckel F. The stability of cognitive abilities: a meta-analytic review of longitudinal studies. *Psychol Bull*. 2024;150: 399–439.
[70] Hedeker D, Mermelstein RJ, Demirtas H. An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*. 2008;64: 627–634.
[71] McNeish D. Specifying location-scale models for heterogeneous variances as multilevel SEMs. *Organ Res Methods*. 2021;24: 630–653.
[72] Berkey CS, Anderson JJ, Hoaglin DC. Multiple-outcome meta-analysis of clinical trials. *Stat Med*. 1996;15: 537–557.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.

[73] Raudenbush SW, Becker BJ, Kalaian H. Modeling multivariate effect sizes. *Psychol Bull*. 1988;103: 111–120.
[74] Jackson D, Riley R, White IR. Multivariate meta-analysis: potential and promise. *Stat Med*. 2011;30: 2481–2498.
[75] Cheung MWL. A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychol Rev*. 2019;29: 387–396.
[76] Gleser LJ, Olkin I. Stochastically dependent effect sizes. In: Cooper H, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis and Meta-Analysis*. 2nd ed. Russell Sage Foundation; 2009: 357–376.
[77] Becker BJ, Aloe AM, Cheung MWL. Meta-analysis of correlation matrices and structural equation models. In: Schmid CH, Stijnen T, White IR, eds. *Handbook of Meta-Analysis*. Chapman and Hall/CRC; 2020: 347–369. https://doi.org/10.1201/9781315119403.
[78] Chan MS, Jones CR, Hall Jamieson K, Albarracín D. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. *Psychol Sci*. 2017;28: 1531–1546.
[79] Arends LR, Hoes AW, Lubsen J, Grobbee DE, Stijnen T. Baseline risk as predictor of treatment benefit: three clinical meta-re-analyses. *Stat Med*. 2000;19: 3497–3518.
[80] Ghidey W, Stijnen T, HCv H. Modelling the effect of baseline risk in meta-analysis: a review from the perspective of errors-in-variables regression. *Stat Methods Med Res*. 2013;22: 307–323.
[81] Guolo A. Flexibly modeling the baseline risk in meta-analysis. *Stat Med*. 2013;32: 40–50.
[82] Guolo A. Improving likelihood-based inference in control rate regression. *Stat Med* 2018;37. https://doi.org/10.1002/sim.7511.
[83] Guolo A. Measurement errors in control risk regression: a comparison of correction techniques. *Stat Med* 2022;41. https://doi.org/10.1002/sim.9228.
[84] Hart ER, Bailey DH, Luo S, Sengupta P, Watts TW. Fadeout and persistence of intervention impacts on social-emotional and cognitive skills in children and adolescents: a meta-analytic review of randomized controlled trials. *Psychol Bull*. 2024;150(10): 1207–1236. https://doi.org/10.1037/bul0000450.
[85] Chan MS, Albarracín D. A meta-analysis of correction effects in science-relevant misinformation. *Nat Hum Behav*. 2023;7: 1514–1525.
[86] Pritikin JN, Hunter MD, Tv O, Brick TR, Boker SM. Many-level multilevel structural equation modeling: an efficient evaluation strategy. *Struct Equ Model Multidiscip J*. 2017;24: 684–698.
[87] Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Res Synth Methods*. 2010;1: 39–65.
[88] Pustejovsky JE, Tipton E. Meta-analysis with robust variance estimation: expanding the range of working models. *Prev Sci*. 2022;23: 425–438.

Cite this article: Cheung, M.W.-L. A tutorial on fitting flexible meta-analytic models with structural equation modeling in R. *Research Synthesis Methods*. 2026: 1–30. https://doi.org/10.1017/rsm.2026.10109

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 00:00:54, subject to the Cambridge Core terms of use.