


Meta-Analyzing Correlation Matrices in the Presence of Hierarchical Effect Size Multiplicity


Ronny Scherer^{1,2} and Diego G. Campos^{1,2}

¹ Centre for Educational Measurement at the University of Oslo (CEMO), Faculty of Educational Sciences, University of Oslo

² Centre for Research on Equality in Education (CREATE), Faculty of Educational Sciences, University of Oslo

Author Note

Ronny Scherer  <https://orcid.org/0000-0003-3630-0710>

Diego G. Campos  <https://orcid.org/0000-0002-8820-5881>

This study was pre-registered at <https://doi.org/10.17605/OSF.IO/ZNWR6>, and the respective data are openly available at <https://doi.org/10.17605/OSF.IO/XU8ES>. We have no known conflict of interest to disclose. This work is part of the Centres of Excellence scheme, funded by the Research Council of Norway, project number 331640, and the project “Mapping of Longitudinal data of Inequalities in Education (MapIE)”, funded by the European Commission under the Horizon Europe scheme, project number 101132474. We are grateful for the work by Schroeders, Kubina, and Gnambs (2022), who made their meta-analytic data set on the factor structure of the TAS-20 openly available.

Correspondence concerning this article should be addressed to Ronny Scherer, University of Oslo, Faculty of Educational Sciences, Centre for Educational Measurement at the University

of Oslo (CEMO) & Centre for Research on Equality in Education (CREATE), Blindernveien 31,
NO-0371 Oslo.

This manuscript is a preprint and has not yet undergone peer review. Status: 01 Oct 2024

Abstract

To synthesize evidence on the relations among multiple constructs, measures, or concepts, meta-analyzing correlation matrices across primary studies has become a crucial analytic approach. Common meta-analytic approaches employ univariate or multivariate models to estimate a pooled correlation matrix, which is subjected to further analyses, such as structural equation or network modeling. In practice, meta-analysts often extract multiple correlation matrices per study from multiple samples, study sites, labs, or countries, thus introducing hierarchical effect size multiplicity into the meta-analytic data. However, this feature has largely been ignored when pooling correlation matrices for meta-analysis. To contribute to the methodological development in this area, we describe a multilevel, multivariate, and random-effects modeling (MLMV-REM) approach, which pools correlation matrices meta-analytically and, at the same time, addresses hierarchical effect size multiplicity. Specifically, it allows meta-analysts to test various assumptions on the dependencies among random effects, aiding the selection of a meta-analytic baseline model. We describe this approach, present four working models within it, and illustrate them with an example and the corresponding R code.

Keywords: Correlation matrix, effect size multiplicity, multilevel meta-analysis, multivariate meta-analysis, meta-analytic structural equation modeling

Highlights

What is already known

- Meta-analyzing correlation matrices allows meta-analysts to test theories and hypotheses on the relations between multiple variables.
- Pooling correlation matrices meta-analytically across studies often relies on univariate or multivariate random-effects models.
- Hierarchical effect size multiplicity can arise when meta-analysts extract multiple correlation matrices from a study, often due to the inclusion of multiple samples, study sites, or countries.
- Existing pooling approaches either largely ignore hierarchical effect size multiplicity or only partially account for it.

What is new

- Multilevel extensions of multivariate random-effects models provide a framework for addressing hierarchical effect size multiplicity in the meta-analytic pooling of correlation matrices.
- Multilevel, multivariate, and random-effects models (MLMV-REM) enable meta-analysts to model the heterogeneity in correlation matrices within *and* between studies.
- These models can be estimated utilizing the multilevel and multivariate features in the R package “metafor”.

Potential impact for *Research Synthesis Methods* readers

- Meta-analysts who wish to synthesize correlation matrices in the presence of hierarchical effect size multiplicity may take a multilevel, multivariate, and random-effects approach.

- We aim to encourage the development and evaluation of models addressing effect size multiplicity in the meta-analytic pooling of correlation matrices.
- We offer an illustrative example along with the corresponding analytical R code that meta-analysts can adapt to their research.

Data Availability Statement

This study was pre-registered in the Open Science Framework (OSF) at <*Link blinded for peer review*>. The analytical R code, data, and supplementary materials are fully disclosed and contained in the respective OSF project at <*Link blinded for peer review*>. The data underlying the illustrative example were extracted from Schroeders, Kubera and Gnambs ¹ who have made them openly available at <https://osf.io/uxtkts/>.

Meta-Analyzing Correlation Matrices in the Presence of Hierarchical Effect Size Multiplicity

1. Introduction

Understanding how multiple constructs, concepts, or variables are related and testing theories that describe these relations requires the synthesis of *multiple* correlation coefficients.² Multiple correlation coefficients are often stored in correlation matrices, and synthesizing them across primary studies means to pool the correlation coefficients via meta-analytic methods into an overall, weighted average correlation matrix.³ Notably, correlation coefficients within a correlation matrix are not independent, as they are usually derived from the same sample. Given this correlational dependence (“correlational effect size multiplicity”), the pooling of correlation matrices often involves multivariate models that provide more accurate meta-analytic estimates and standard errors than multiple, separate univariate models.^{4,5}

With the growing availability of published studies and data, multiple independent samples, study sites, or countries may be included per primary study. This scenario is referred to as “hierarchical effect size multiplicity”.⁶ For instance, in their study of the relations among technology acceptance variables, Scherer et al.⁷ extracted multiple correlation matrices from primary studies, because some studies included more than one teacher sample (e.g., in-service and pre-service teachers). The resultant meta-analytic data were comprised of multiple correlation matrices that were hierarchically nested in studies.⁸

While dealing with hierarchical effect size multiplicity has been well-described for *univariate* outcomes (e.g., via multilevel random-effects models⁹ or cluster-robust variance estimation¹⁰ of one correlation coefficient), it has received less attention for *multivariate* outcomes, especially when correlation matrices are pooled. Similar to a univariate scenario,

ignoring hierarchical effect size multiplicity when meta-analyzing multivariate data may result in biased standard errors and heterogeneity estimates.^{11,12} In our view, existing approaches to pool correlation matrices meta-analytically largely ignored hierarchical effect size multiplicity or addressed it only to a limited extent.¹¹ For instance, pooling correlation matrices across studies—as a key stage in meta-analytic structural equation modeling—is often based on multivariate random-effects models assuming that correlation matrices are fully independent. Acknowledging this limitation, Wilson et al.¹³ developed a multilevel random-effects modeling approach—referred to as the “WPL approach”—that quantifies within- and between-study heterogeneity of the correlation coefficients within correlation matrices. However, this approach assumes that within- and between-heterogeneity are the same across correlations—an assumption that may or may not hold in practical applications.

To address these limitations, we present a multilevel, multivariate, and random-effects modeling (MLMV-REM) approach to pooling correlation matrices in the presence of hierarchical effect size multiplicity. Extending the existing approaches, the MLMV-REM approach allows meta-analysts to flexibly specify random-effects structures within and between studies, obtain information about correlation-specific heterogeneity, and examine moderator effects. In this tutorial paper, we describe the approach, highlight the key analytical decisions within, present four working models, and illustrate its application with an example, including the respective R code. Despite its flexibility, we hope to stimulate future evaluations of the MLMV-REM approach and to advance the development of approaches that deal with effect size multiplicity when pooling correlation matrices meta-analytically.

2. Existing Approaches to Pooling Correlation Matrices Meta-Analytically

In this section, we explain the different types of dependencies among correlation coefficients that may occur when meta-analyzing correlation matrices and review univariate and multivariate pooling approaches. While various methods have been developed to pool correlation matrices meta-analytically, each presents specific limitations. These shortcomings highlight the need for a more flexible and comprehensive method that can address both the correlational and hierarchical dependencies that may occur in meta-analytic data, ensuring more accurate and reliable meta-analytic conclusions.

2.1 Dependencies among Effect Sizes

While pooling correlation matrices across studies has gained popularity in meta-analysis, it still faces significant challenges. One of these challenges refers to effect size multiplicity that occurs due to the dependencies among effect sizes both within and across studies.^{11,13,14} Cheung⁴ summarized effect size dependencies as (a) dependencies among the sampling errors that occur due to the administration of multiple measures to the same sample in each study, (b) dependencies among true effect sizes at the population level across studies, and (c) dependencies due to hierarchies in the data, such as multiple correlation coefficients that are nested in studies.

In the context of pooling correlation matrices meta-analytically, such dependencies can occur not only because multiple correlation coefficients are stored in correlation matrices, but also because multiple correlation matrices are extracted from one study. For example, in a meta-analysis of psychological traits—such as anxiety, depression, and stress—studies might report multiple correlation matrices for samples defined by different clinical statuses (e.g., clinical vs. non-clinical samples). In another meta-analysis, authors may extract correlation matrices describing the relations among student achievement, motivation, and engagement in reading,

mathematics, and science, utilizing data from international large-scale assessments in education, such as the Programme for International Student Achievement (PISA) or the Trends in Mathematics and Science Study (TIMSS).^{8,15} Given the inclusion of several countries in these assessments, multiple correlation matrices are available per assessment cycle.

These examples illustrate that dependencies among multiple correlation coefficients may not only exist within a correlation matrix, because multiple correlations were derived from the same sample yet multiple constructs or measures (i.e., correlational effect size multiplicity).⁶ The inclusion of multiple correlation matrices per study that are derived from independent samples also creates hierarchical effect size multiplicity.¹⁶

Failing to account for such dependencies can bias meta-analytic results, lead to inaccurate pooled estimates, and compromise statistical inferences.^{6,16,17} Several statistical approaches have been proposed to address effect size multiplicity when pooling correlation matrices meta-analytically, which we describe briefly in the next sections. Notably, we exclude approaches that ignore effect size multiplicity—be it the fact that multiple correlation coefficients are stored in correlation matrices or multiple correlation matrices are available per study. For an overview of these approaches, we kindly refer readers to Stolwijk et al.¹¹

2.2 Univariate Approaches

Univariate approaches pool each correlation coefficient within a correlation matrix separately across studies via a series of standard, univariate, fixed- or random-effects models.^{4,18,19} For instance, taking a univariate approach, the correlation between anxiety and depression would be meta-analyzed independent from the correlations between anxiety and stress and between stress and depression. In situations where multiple correlations between the same constructs or measures are available due to the inclusion of multiple samples, univariate

approaches can account for this hierarchical effect size multiplicity via multilevel random-effects models. Such models estimate the variation of correlation coefficients within and between studies and allow meta-analysts to explore possible causes of this heterogeneity.^{20,21} However, this approach neglects the inherent multivariate structure of correlation matrices, where multiple correlations share common constructs and are derived from the same sample (i.e., correlational effect size multiplicity). Ignoring this dependence among correlations within matrices leads to an underestimation of standard errors and, consequently, inflated Type I error rates.^{22,23} Hence, univariate approaches are not recommended for synthesizing correlation matrices with or without hierarchical effect size multiplicity.⁵

2.3 Multivariate Approaches

Multivariate approaches can account for dependencies among correlation coefficients within studies, which arise from measuring multiple outcomes on the same sample.^{24,25} These models incorporate not only the sampling variances but also some information about the covariances of correlation coefficients, allowing for the estimation of a pooled correlation matrix that reflects the interrelationships between multiple outcomes.⁵ By simultaneously leveraging all available correlations, multivariate approaches “borrow strength”, for instance, from the within-study correlations, leading to more precise estimates of mean effect sizes and more reliable assessments of between-study heterogeneity.^{23,26,27} Moreover, multivariate approaches often result in joint confidence intervals and distributions and allow meta-analysts to predict an effect controlling for another, correlated effect.²⁷ However, the application of multivariate approaches in practice is often constrained by the limited availability of information about sampling covariances or correlations in primary studies, and estimation problems may occur due to the increased number of parameters and their joint distributions as compared to separate univariate

meta-analyses.^{27,28} In the following section, we briefly describe some multivariate approaches to pooling correlation matrices meta-analytically.

Generalized Least Squares (GLS) Approach. Becker^{25,29} developed a multi-step GLS approach, in which multiple correlation coefficients can be synthesized while accounting for their dependencies. Specifically, in this approach, the correlation coefficients and sampling variances are extracted from the primary studies, and sampling covariances are estimated for each pair of correlation coefficient. Using the sampling covariance matrix and between-study heterogeneity estimates for each correlation coefficient to obtain weights under a random-effects assumption, a weighted average, that is, pooled correlation matrix is then obtained.²⁹ While the GLS approach can handle the dependencies among multiple correlations within correlation matrices, it does not account for hierarchical effect size multiplicity when multiple correlations are available for the same constructs across different samples within studies. It would likely need a multilevel extension of the step in which heterogeneity variances are estimated and incorporated in the weights.

Multivariate Random-Effects Modeling (MV-REM) Approach. Given that correlation matrices within studies contain multiple correlation coefficients from the same sample, dependencies among these effect sizes occur (i.e., correlational effect size multiplicity). A multivariate random-effects modeling approach can handle these dependencies and overcome issues associated with estimating multiple, separate, univariate random-effects models, such as the risk of bias and reduced precision.^{4,23} Specifically, multivariate random-effects models (MV-REM) decomposes multiple observed correlation coefficients into true effect sizes, random effects, and sampling error. Given the multivariate nature of the data, the sampling covariance matrix of a study does not only contain the sampling variances in the diagonal but also non-zero

sampling covariances in the off-diagonal.⁴ These covariance matrices are study-specific, so that two primary studies may show different degrees of dependencies. At the population level, the dependencies among effect sizes is modeled by the structure of the random effects, that is, the structure of the heterogeneity covariance matrix. In this matrix, the between-study heterogeneity variances are stored in the diagonal, and the off-diagonal represents the assumptions on the dependencies at the population level. Several structures have been proposed (e.g., unstructured, diagonal, or compound symmetric matrices)^{4,30}, and we will describe them in greater detail in section 3. Overall, while the MV-REM approach handles correlational effect size multiplicity within correlation matrices, it does not consider hierarchical effect size multiplicity. Hence, taking this approach, meta-analysts are limited to meta-analytic data in which each study contributes one correlation matrix.

Structural Equation Modeling (SEM) Approach. To synthesize correlation matrices across studies, Cheung³¹ translated multivariate meta-analysis into the structural equation modeling framework. In this framework, the vector of correlation coefficients and their variance components are modeled as mean and covariance structures in SEM. This approach follows two steps^{4,32}: In the first step, researchers obtain the sampling variance-covariance matrix for each study via confirmatory factor analyses. In the second step, the vector of correlation coefficients is pooled via a multivariate random-effects model that also incorporates the sampling covariance matrices. Besides a pooled correlation matrix, a between-study heterogeneity covariance matrix is estimated. The SEM approach overcomes some key issues multivariate meta-analysis has, such as the handling of missing correlation coefficients in some primary studies. Despite its utility in accounting for the dependencies among correlation coefficients within a correlation matrix (correlational effect size multiplicity), the SEM approach is less applicable when multiple

correlation matrices are reported for independent samples within a study (hierarchical effect size multiplicity), because it allows each study to contribute only one correlation to each cell of the synthesized correlation matrix.¹³ Hence, this approach is constrained by its assumption of independence between correlation matrices across studies.¹¹

Wilson-Polanin-Lipsey (WPL) Approach. The three-level random-effects modeling approach, proposed by Wilson and colleagues¹³, addresses both correlational and hierarchical effect size multiplicity when pooling correlation matrices. This approach includes information from all available (dependent) effect sizes per study and explicitly models these dependencies in the pooled correlation matrix. Furthermore, it estimates both within- and between-study variances, allowing for the examination of heterogeneity at different levels (Cheung, 2014b, 2019). In essence, the WPL approach is based on a multivariate, multilevel, and random-effects model, which yields a pooled correlation matrix, one within-study variance estimate, and one between-study variance estimate.^{11,13} However, the WPL approach assumes that these heterogeneity variance estimates are the same for all correlation coefficients. This assumption may not hold in practice, especially when the correlation coefficients represent associations among diverse constructs which may vary substantially across studies. In such a situation, meta-analysts may expect different amounts of heterogeneity across correlation coefficients. Overall, the three-level random-effects approach by Wilson and colleagues addresses both correlational and hierarchical dependencies, but its simplifying assumption about within- and between-study heterogeneity limits its utility.

3. A Multilevel, Multivariate, and Random-Effects Modeling (MLMV-REM) Approach

When multiple, independent samples provide correlation matrices, both correlational and hierarchical effect size multiplicity exist. One approach to account for these multiplicities is to

model the dependencies among multiple correlations within a correlation matrix via a *multivariate* component and to model the hierarchical nesting of multiple, independent correlation matrices in primary studies via a *multilevel* component. This approach can be based on multilevel, multivariate, and random-effects models (MLMV-REM)³³ that systematically extend the existing multivariate approaches (e.g., the WPL and MV-REM approaches).

3.1 Specification of the MLMV-REM

Suppose we have a meta-analytic data set that comprises correlation coefficients as effect sizes and their respective sampling variances. Let $\hat{\theta}_{ijk}$ denote the estimate of the i th correlation coefficient in the j th correlation matrix (or, respectively, the j th sample) extracted from the k th primary study. This estimate has a sampling variance v_{ijk} , which is often approximated as $v_{ijk} = (1 - \hat{\theta}_{ijk}^2)^2 / (n_{ijk} - 1)$, where n_{ijk} refers to the sample size underlying the correlation coefficient estimate $\hat{\theta}_{ijk}$.¹⁸ Given this terminology, the j th correlation matrix within the k th primary study that contains I correlation coefficients can then be written as $\hat{\boldsymbol{\theta}}_{jk} = [\hat{\theta}_{ijk}]_{i=1}^I$.

A generic version of a MLMV-REM can be specified by distinguishing between three levels of analysis. At level 1, the effect size estimates $\hat{\boldsymbol{\theta}}_{jk}$ are decomposed into true effect size $\boldsymbol{\lambda}_{jk}$ and sampling error.²¹

$$\text{Level 1 (sampling error):} \quad \hat{\boldsymbol{\theta}}_{jk} = \boldsymbol{\lambda}_{jk} + \mathbf{e}_{jk} \text{ with } \mathbf{e}_{jk} \sim MVN(\mathbf{0}, \mathbf{V}_{jk}).$$

Given the multivariate nature of the data (i.e., multiple correlations within a correlation matrix due to multiple outcomes or constructs of interest), the sampling variances v_{ijk} are stored in the diagonal of the sampling covariance matrices \mathbf{V}_{jk} , and possible covariances among the sampling correlation coefficients are stored off the diagonal. \mathbf{V}_{jk} take the following symmetric form:

$$\text{Sampling covariance matrix: } \mathbf{V}_{jk} = \begin{bmatrix} v_{1jk} & & & \\ cov_{(2,1)jk} & v_{2jk} & & \\ \vdots & \vdots & \ddots & \\ cov_{(I,1)jk} & cov_{(I,2)jk} & \cdots & v_{Ijk} \end{bmatrix}.$$

Oftentimes, meta-analysts only have access to the sampling variances v_{ijk} yet not to the sampling covariances $cov_{(i,i')jk}$ for the i th and i' th variables. Several solutions have been proposed to extract or approximate these covariances. For instance, Olkin and Finn³⁴ proposed an explicit expression for the covariance between two correlation coefficients that share one variable. Cheung⁴ suggested to translate meta-analytic models into the framework of structural equation modeling which makes available an asymptotic sampling covariance matrix. Wei and Higgins³⁵ described a strategy in which the available information about the correlations among effect sizes is meta-analyzed, and the resultant, pooled correlation coefficient is then used to impute the correlations for studies or matrices that do not provide this information. Another solution is to specify values of the correlations among effect sizes $r_{(i,i')jk}$, either informed by existing research or estimated correlations among constructs from the primary study data. The resultant sampling covariances are then derived as $cov_{(i,i')jk} = r_{(i,i')jk} \cdot \sqrt{v_{ijk}v_{i'jk}}$.

Alternatively, meta-analysts may then assume a non-zero sampling correlation r that is constant within each correlation matrix and equal across samples and studies.⁶ While this is a strong assumption, it still accounts for the dependencies among correlation coefficients to some degree, simplifies the sampling covariances to $cov_{(i,i')jk} = r \cdot \sqrt{v_{ijk}v_{i'jk}}$, and thus makes the estimation of \mathbf{V}_{jk} accessible, especially when information about the correlations among effect sizes is not available from the primary data. However, if multiple values of r may seem plausible from a substantive perspective, then sensitivity analyses could shed light on the impact of r on the meta-analytic results. Overall, the level-1 specification of the MLMV-REM is the first instance, in

which meta-analysts must decide on how to represent the effect size multiplicity within correlation matrices.

To further address the nesting of multiple correlation matrices (or samples) in primary studies, two additional levels of analysis are specified. Level 2 contains information about the within-study heterogeneity of correlation coefficients, and level 3 contains information about the between-study heterogeneity of correlation coefficients.^{9,20,21} The decomposition of effect sizes and the respective assumptions can be specified as follows:

$$\text{Level 2 (within studies): } \lambda_{jk} = \eta_k + \mathbf{u}_{jk} \text{ with } \mathbf{u}_{jk} \sim MVN(\mathbf{0}, \Sigma),$$

$$\text{Level 3 (between studies): } \eta_k = \mu + \mathbf{g}_k \text{ with } \mathbf{g}_k \sim MVN(\mathbf{0}, \Omega).$$

In this specification, the vector of correlation coefficients λ_{jk} is decomposed into an average study-level vector of correlations η_k and its matrix- or sample-specific deviations from it, \mathbf{u}_{jk} .

At level 2, these deviations may exhibit within-study heterogeneity, which is stored in the respective heterogeneity matrix Σ . Accordingly, η_k is then further decomposed into an overall effect size estimate μ and the study-specific deviations from it, \mathbf{g}_k . At level 3, these deviations may exhibit between-study heterogeneity, which is stored in the respective heterogeneity matrix Ω . Combining all levels of analysis results in the following, overall model specification:

$$\hat{\theta}_{jk} = \mu + \mathbf{g}_k + \mathbf{u}_{jk} + \mathbf{e}_{jk} \text{ with } \mathbf{e}_{jk} \sim MVN(\mathbf{0}, \mathbf{V}_{jk}), \mathbf{u}_{jk} \sim MVN(\mathbf{0}, \Sigma), \mathbf{g}_k \sim MVN(\mathbf{0}, \Omega).$$

This MLMV-REM allows meta-analysts to quantify correlation-specific heterogeneity within and between primary studies (see Table 1, Model 1). In this context, “correlation-specific” means that the parameters are estimated for each of the correlations within a correlation matrix. For instance, if meta-analysts wish to synthesize the correlations among four variables, then each correlation matrix contains six correlations off the diagonal, which represent the associations among the six pairs of variables. In our specification of the MLMV-REM, each of these six types

of correlations have their own estimates of the within- and between-study heterogeneity variances. In the next section, we describe a range of structures the heterogeneity matrices Σ and Ω may have.

Finally, the MLMV-REM can be further extended to a mixed-effects meta-regression model by adding explanatory variables, such as sample, study, publication, or measurement characteristics. In such an extension, \mathbf{u}_{jk} and \mathbf{g}_k may become residuals in a regression sense, depending at which level the explanatory variables are defined.³⁶ Moreover, heterogeneity could be explained specifically for each correlation coefficient at two levels of analysis (i.e., within and between studies). This entails that multiple measures of variance explanations are accessible for each correlation coefficient, once explanatory variables are added to the MLMV-REM (i.e., $R^2_{(2)i}$ and $R^2_{(3)i}$ for each of the $i = 1, \dots, I$ correlations).^{4,37}

3.2 Specification of the Level-2 and Level-3 Random Effects

As noted earlier, heterogeneity can be modeled with the MLMV-REM both within and between studies for each type of correlation. Specifically, the symmetric heterogeneity matrices Σ and Ω store the correlation-specific heterogeneity variances in the diagonal and the heterogeneity covariances in the off-diagonal. For I correlation coefficients, these matrices can take the following, often referred to as “unstructured” forms:

$$\begin{aligned} \text{Level-2 heterogeneity covariance matrix: } \Sigma &= \begin{bmatrix} \tau_1^2 & & & \\ \rho_{(2,1)}\tau_2\tau_1 & \tau_2^2 & & \\ \vdots & \vdots & \ddots & \\ \rho_{(I,1)}\tau_I\tau_1 & \rho_{(I,2)}\tau_I\tau_2 & \cdots & \tau_I^2 \end{bmatrix} \\ \text{Level-3 heterogeneity covariance matrix: } \Omega &= \begin{bmatrix} \gamma_1^2 & & & \\ \phi_{(2,1)}\gamma_2\gamma_1 & \gamma_2^2 & & \\ \vdots & \vdots & \ddots & \\ \phi_{(I,1)}\gamma_I\gamma_1 & \phi_{(I,2)}\gamma_I\gamma_2 & \cdots & \gamma_I^2 \end{bmatrix} \end{aligned}$$

These random-effects covariance matrices contain the within-study variances τ_i^2 , the between-study variances γ_i^2 for each correlation, and the covariances that are composed of the within- or between-study standard deviations, τ_i and γ_i , and the correlations among the within- or between-level random effects of the correlation coefficients, $\rho_{(i,i')}$ and $\phi_{(i,i')}$ for $i, i' = 1, \dots, I$.

This form of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ is “unstructured” and allows meta-analysts to include the covariances among random effects that are specific to each pair (i, i') of correlation coefficients. However, although the specification of this form has a large degree of flexibility, meta-analysts may not be able to estimate all parameters within the heterogeneity covariance matrices. For instance, the likely small number of available correlations and primary studies could lead to unstable variances and small statistical power of the respective estimates. Moreover, information about the correlations among the random effects within and between studies may be lacking. Hence, meta-analysts often must simplify the structures of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$. We would like to highlight two approaches to simplifying the heterogeneity matrices. The first approach concerns the correlations among the random effects, $\rho_{(i,i')}$ and $\phi_{(i,i')}$. Like the specification of the sampling covariance matrices \mathbf{V}_{jk} , meta-analysts could fix $\rho_{(i,i')}$ and $\phi_{(i,i')}$ to some non-zero constant values, ρ and ϕ , that are the same for each random effect or approximate them in some way. The structure with constant random-effects correlations is referred to as the “heteroscedastic compound symmetry (HCS)” structure. If ρ or ϕ are fixed to zero, then the matrices $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ have a diagonal structure—a special case of the HCS structure in which meta-analysts can model only the within- and between-study heterogeneity variances and assume that random effects at level 2 and 3 are independent. This diagonal structure is often used in the first stage of MASEM, in which multiple correlation matrices are pooled via a multivariate random-effects model.⁴ The heterogeneity matrices with an HCS or diagonal structure are as follows:

$$\begin{aligned} \text{HCS structures: } \mathbf{\Sigma} &= \begin{bmatrix} \tau_1^2 & & & \\ \rho\tau_2\tau_1 & \tau_2^2 & & \\ \vdots & \vdots & \ddots & \\ \rho\tau_I\tau_1 & \rho\tau_I\tau_2 & \dots & \tau_I^2 \end{bmatrix}, \mathbf{\Omega} = \begin{bmatrix} \gamma_1^2 & & & \\ \phi\gamma_2\gamma_1 & \gamma_2^2 & & \\ \vdots & \vdots & \ddots & \\ \phi\gamma_I\gamma_1 & \phi\gamma_I\gamma_2 & \dots & \gamma_I^2 \end{bmatrix}. \\ \\ \text{Diagonal structures: } \mathbf{\Sigma} &= \begin{bmatrix} \tau_1^2 & & & \\ 0 & \tau_2^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \tau_I^2 \end{bmatrix}, \mathbf{\Omega} = \begin{bmatrix} \gamma_1^2 & & & \\ 0 & \gamma_2^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \gamma_I^2 \end{bmatrix}. \end{aligned}$$

Notably, these structures contain heterogeneity variances that are specific to the random effects at levels 2 and 3 and thus to the correlation coefficients. In other words, for each correlation coefficient (i.e., effect size), within- and between-study variance components are estimated. A second approach to simplifying $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ is to constrain these variance components to be equal across random effects. Imposing this constraint to an HCS structure results in a so-called “compound symmetric (CS)” structure. In the special case where ρ or ϕ are fixed to zero, the structure is that of a “scaled identity” structure.³⁰ Notably, the term “scaled identity structure” is based on the fact that the two heterogeneity matrices have the forms $\mathbf{\Sigma} = \tau^2 \mathbf{I}_I$ and $\mathbf{\Omega} = \gamma^2 \mathbf{I}_I$, where \mathbf{I}_I represents an $I \times I$ identity matrix.

$$\begin{aligned} \text{CS structures: } \mathbf{\Sigma} &= \begin{bmatrix} \tau^2 & & & \\ \rho\tau^2 & \tau^2 & & \\ \vdots & \vdots & \ddots & \\ \rho\tau^2 & \rho\tau^2 & \dots & \tau^2 \end{bmatrix}, \mathbf{\Omega} = \begin{bmatrix} \gamma^2 & & & \\ \phi\gamma^2 & \gamma^2 & & \\ \vdots & \vdots & \ddots & \\ \phi\gamma^2 & \phi\gamma^2 & \dots & \gamma^2 \end{bmatrix}. \\ \\ \text{Scaled identity structures: } \mathbf{\Sigma} &= \begin{bmatrix} \tau^2 & & & \\ 0 & \tau^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \tau^2 \end{bmatrix}, \mathbf{\Omega} = \begin{bmatrix} \gamma^2 & & & \\ 0 & \gamma^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \gamma^2 \end{bmatrix}. \end{aligned}$$

If meta-analysts decide for one of the structures that assume fixed correlations among random effects, ρ or ϕ , in our view, they should examine the extent to which the meta-analytic findings and inferences are sensitive to the choice of these fixed values.^{10,21}

3.3 Working Models and Decision Scheme

Table 1 provides an overview of four working models within the MLMV-REM approach, which estimate weighted average effect sizes and the within- and between-study variance components and assume constant correlations among the level-2 and level-3 random effects. Model 1 allows the heterogeneity estimates at levels 2 and 3 to be specific to each correlation coefficient. Hence, Σ and Ω may simplify to an HCS or a diagonal structure. Model 2 estimates correlation-specific within-study heterogeneity variances but constrains the between-study heterogeneity variances to being equal across correlation coefficients. Hence, Σ follows an HCS or a diagonal structure, while Ω follows a CS or scaled identity structure. This model assumes that the correlation coefficients have the same amount of heterogeneity between (yet not within) studies. Conversely, Model 3 estimates correlation-specific between-study heterogeneity and constrains the within-study heterogeneity to being equal across correlation coefficients. Finally, Model 4 imposes the equality constraints to both the within- and between-study heterogeneity variances. In this model, Σ and Ω both have a CS or scaled identity structure. Notably, this model is similar to the meta-analytic model underlying the WPL approach.¹³

This selection of working models shows that meta-analysts have several options to specify and simplify similar or different level-2 and level-3 variance structures (see Table 1). We chose Models 1-4 as working models because, in our view, they are relevant for meta-analytic modeling practice for several reasons. First, these models allow meta-analysts to estimate correlation-specific heterogeneity variances. The information about the extent to which specific correlations within correlation matrices vary within and between studies identifies the levels at which heterogeneity is located and could be explained by potential moderators.

Second, as noted earlier, the oftentimes lacking or limited information about the correlations among the level-2 and level-3 random effects in a meta-analysis may not allow meta-analysts to estimate these correlations for each correlation coefficient separately. Hence, fixing these random-effects correlations to some constant value represents a pragmatic solution to still accounting for random-effects dependencies despite the lack of information about their extent.⁶

Third, the constant random-effects correlations ρ and ϕ can also be fixed to zero, assuming the independence of random effects at levels 2 and 3. This assumption is sometimes made when meta-analysts are only interested in estimating the heterogeneity variances.^{4,5} However, in this situation, the dependencies among the level-2 and level-3 random effects are not accounted for and may introduce some bias to the meta-analytic estimates and their standard errors.

Fourth, in working models 2-4, the within- and/or between-study heterogeneity variances are constrained to being equal across correlation coefficients. While these models make a strong and oftentimes unrealistic assumption on the heterogeneity in the meta-analytic data, they are still useful for meta-analytic practices. Specifically, comparing a model with such equality constraints to working model 1 allows meta-analysts to test hypotheses on the similarity or differences in heterogeneity variance estimates across the correlation coefficients.⁴ The constraints in Models 2-4 offer meta-analysts ways to simplify the meta-analytic models and thus strive for model parsimony.

Fifth, Models 1-4 represent several substantive assumptions that may represent meta-analysts' theories and hypotheses on the heterogeneity in the meta-analytic data. Model 1 represents the most flexible model among the working models and is useful in situations where

the correlation-specific effect sizes and heterogeneity variances are in meta-analysts' key interest. Moreover, Model 4 may also be useful to model meta-analytic data are comprised of highly heterogeneous samples within primary studies of different characteristics. Model 2 assumes that correlation coefficients can vary within studies to a different extent and between studies to the same extent. This model could be useful for meta-analytic data that are comprised of highly heterogeneous samples within primary studies that follow the same study protocol (e.g., multi-lab studies). Conversely, Model 3 might be useful to model meta-analytic data with highly homogeneous samples within studies but studies that differ substantially in their characteristics (e.g., different sampling protocols). Model 4 may be useful in situations where (a) meta-analysts wish to account for heterogeneity both within and between studies yet do not have a large-enough meta-analytic data base to model correlation-specific heterogeneity variances; (b) within- and between-study heterogeneity variances are homogeneous (e.g., consistently close to zero across all correlation coefficients); or (c) meta-analysts wish to test hypotheses on the equality of the heterogeneity variances.

To select one of the working models, meta-analysts must make several analytic decisions. Figure 1 provides an overview of these decisions and their possible outcomes. First, when meta-analysts have extracted the correlation matrices and the respective correlation coefficients from the primary study data, they must decide on the structure of the sampling covariance matrices \mathbf{V}_{jk} . Specifically, assuming that correlations within correlation matrices/samples are independent results in diagonal sampling covariance matrices which essentially only contain the sampling variances of the correlation coefficients. Alternatively, meta-analysts could assume some dependence among the sampling errors and then specify \mathbf{V}_{jk} with sampling covariances off the diagonals. In scenarios where multiple measures of the same sample have been taken over time

or multiple, correlated constructs were assessed for each sample, a dependency structure with some correlation among sampling errors is conceptually reasonable and preferable over an independence structure.^{4,6,21} The extracted correlation coefficients $\hat{\Theta}_{jk}$ and sampling covariance matrices can then be submitted to meta-analytic modeling.

Second, if effect size multiplicity is present in the meta-analytic data, meta-analysts must decide if they wish to account for it. In case they wish to ignore effect size multiplicity—for instance, due to close-to-zero within-study heterogeneity in the correlation coefficients—they may choose a multivariate random-effects model to pool the correlation matrices.^{2,11,14} To model effect size multiplicity, meta-analysts may choose a working model within the MLMV-REM approach.

Third, meta-analysts must decide on the structures of the within- and between-study heterogeneity matrices. Specifically, do they allow for correlation-specific within- and between study variances in correlation coefficients and which assumption do they make on the random-effects correlations within and between studies (i.e., independence assumption with $\rho = \phi = 0$ or dependence assumptions with some non-zero values of ρ and ϕ)? While the former decision determines which of the four working models is selected, the latter decision must be made for any model that has been selected.

In addition to these substantive decisions, meta-analysts may also rely on their fit information and the results of model comparisons. For instance, working models 2 and 3 are both nested in Model 1, so that, in addition to inspecting information criteria, likelihood-ratio testing could provide information about which model may be preferred. In this sense, the selection of a working model is both a substantive and statistical decision.

4. Illustrative Example

To support meta-analysts with translating their decisions on the dependencies among sampling errors, the modeling of within- and between-study heterogeneity variances, and the choices for the level-2 and level-3 correlations among random effects, we illustrate the implementation of Models 1-4 in the R software with an example data set. Specifically, we showcase the model estimation using the R package “metafor”³⁰ and the generation of cluster-robust standard errors using the R package “clubSandwich”³⁸. Please find the full data in Supplementary Material S1 and the detailed R code with the respective output in Supplementary Material S2 and S3.

4.1 Description of the Meta-Analytic Data

We chose a subset of an openly available data set that was published by Schroeders et al.¹ This meta-analytic data set contains the correlations among the items of the Toronto Alexithymia Scale (TAS-20) and/or the statistical parameters to derive them (e.g., standardized factor loadings and factor correlations). For illustrative purposes, we chose the data focusing on participants’ self-reported difficulties with describing feelings (DDF). This dimension was measured by the following five items of the TAS-20 (see Schroeders et al.):

- Item 2: *It is difficult for me to find the right words for my feelings.*
- Item 4: *I am able to describe my feelings easily. (reverse-coded)*
- Item 11: *I find it hard to describe how I feel about people.*
- Item 12: *People tell me to describe my feelings more.*
- Item 17: *It is difficult for me to reveal my innermost feeling, even to close friends.*

Given the focus on these five items, the correlation matrices we intended to pool meta-analytically, were 5×5 matrices with 10 correlation coefficients each. The data were accessed

from a csv-file, stored in an R object named `tas20`, and then supplemented by an effect size identifier `ESID`.

```
tas20 <- read.csv2("TAS20-Data.csv", header = TRUE)

tas20$ESID <- seq_len(nrow(tas20))

head(tas20, 3) ## First three correlations in study 1
```

Output:

```
##      Reference SampleID      Cell Correlation StudyID PubYear Country PubType    N
## 1  aluja2020         83    I2-I4      0.1402      57    2020   Spain        1 602
## 2  aluja2020         83    I2-I11    0.3714      57    2020   Spain        1 602
## 3  aluja2020         83    I4-I11    0.2137      57    2020   Spain        1 602
##      PropFemale Age Clinical ESID
## 1      25.25 52.8         0     1
## 2      25.25 52.8         0     2
## 3      25.25 52.8         0     3
```

To facilitate the meta-analytic modeling and to indicate the different analytical levels, we created identifiers of the effect sizes (`ESID`), the samples (`SampleID`), the primary studies (`StudyID`), and the type of correlation (`Cell`; e.g., “I2-I4” indicates the correlation between items 2 and 4). The correlation coefficients were stored in the variable `Correlation`, and the respective sampling variances in the variable `SamplingVar`.

The dataset contained 880 correlation coefficients that were extracted from 88 samples in 62 primary studies, totaling 69722 participants. On average, the samples were comprised of 792 participants ($M = 792.3$, $SD = 1508.8$, $Mdn = 327$, $Min = 99$, $Max = 12706$), and each study included 1.4 correlation matrices ($SD = 0.6$, $Mdn = 1$, $Min = 1$, $Max = 4$).

4.2 Sampling Covariance Matrices

If meta-analysts assume independent sampling errors, then they may directly use the sampling variances of the correlation coefficients as input for the meta-analytic working model.

If meta-analysts decided to account for the dependencies among sampling errors within correlation matrices for a specific sample, they may create sampling covariance matrices using the information about the known sampling variances and the assumed correlation among the sampling errors of the correlation coefficients (r). There are several ways to generate the sampling matrices. For instance, the R package “clubSandwich” contains the `impute_covariance_matrix()` function, which creates a list of sampling covariance matrices (V) for each correlation matrix (`SampleID`).

```
V <- impute_covariance_matrix(vi = tas20$SamplingVar, cluster =
                             tas20$SampleID, r = 0.6, check_PD = TRUE)
```

This function further allows meta-analysts to specify the type of dependencies (e.g., correlations among multiple measures, autoregressive structures, or both) and to check the sampling covariance matrices for positive definiteness (`check_PD = TRUE`). Notably, we used the identifier of independent samples (`SampleID`) to indicate the correlation matrix, because each independent sample provided one correlation matrix. The object V contains the full sampling covariance matrix with a block design to indicate the correlation matrices.

```
## Sampling covariance matrix of the first sample
V[1:10, 1:10]
```

Output:

```
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 0.0015991253 0.0008437111 0.0009340172 0.0008852038 0.0009237858
[2,] 0.0008437111 0.0012365241 0.0008213248 0.0007784009 0.0008123279
```

```

[3,] 0.0009340172 0.0008213248 0.0015153913 0.0008617165 0.0008992749
[4,] 0.0008852038 0.0007784009 0.0008617165 0.0013611361 0.0008522771
[5,] 0.0009237858 0.0008123279 0.0008992749 0.0008522771 0.0014823734
[6,] 0.0008062322 0.0007089575 0.0007848403 0.0007438231 0.0007762430
[7,] 0.0008708351 0.0007657658 0.0008477291 0.0008034253 0.0008384429
[8,] 0.0009159673 0.0008054526 0.0008916637 0.0008450638 0.0008818963
[9,] 0.0007739516 0.0006805717 0.0007534162 0.0007140413 0.0007451632
[10,] 0.0008104787 0.0007126917 0.0007889742 0.0007477410 0.0007803316
      [,6]      [,7]      [,8]      [,9]     [,10]
[1,] 0.0008062322 0.0008708351 0.0009159673 0.0007739516 0.0008104787
[2,] 0.0007089575 0.0007657658 0.0008054526 0.0006805717 0.0007126917
[3,] 0.0007848403 0.0008477291 0.0008916637 0.0007534162 0.0007889742
[4,] 0.0007438231 0.0008034253 0.0008450638 0.0007140413 0.0007477410
[5,] 0.0007762430 0.0008384429 0.0008818963 0.0007451632 0.0007803316
[6,] 0.0011291076 0.0007317494 0.0007696732 0.0006503396 0.0006810328
[7,] 0.0007317494 0.0013173067 0.0008313467 0.0007024510 0.0007356036
[8,] 0.0007696732 0.0008313467 0.0014573871 0.0007388564 0.0007737272
[9,] 0.0006503396 0.0007024510 0.0007388564 0.0010405013 0.0006537651
[10,] 0.0006810328 0.0007356036 0.0007737272 0.0006537651 0.0011410332

```

Another option to create sampling covariance matrices is to use the `vcalc()` function in the R package `metafor`. This function has similar utilities as the above-described function.

Nonetheless, it contains some additional features, such as the possibility to substitute matrices that are not positive definite by their nearest positive definite matrix (`nearpd = TRUE`).

```

V <- vcalc(vi = SamplingVar, cluster = SampleID, obs = ESID,
           data = tas20, rho = 0.6, checkpd = TRUE, nearpd = TRUE)

```

The sampling covariance matrices generated with these two options were identical for our illustrative example (average difference = $-3.4 \cdot 10^{-21}$, maximal absolute difference = $3.5 \cdot 10^{-18}$). In the following analyses, we used the V matrix generated from the `vcalc()` function.

4.3 Approaches to Pooling Correlation Matrices

Once the correlation matrices and sampling (co-)variance matrices have been extracted or generated, meta-analysts can employ the working models we proposed to estimate a pooled correlation matrix and heterogeneity (co-)variances. In this section, we describe in detail the implementation of working model 1 and explain briefly the implementation of Models 2-4 in R.

Model 1. Given the multivariate nature of the data (i.e., multiple correlations per sample that are based on multiple, measured variables in the primary studies) and the hierarchical effect size multiplicity (i.e., multiple correlation matrices nested in primary studies), we used the `rma.mv()` function in “metafor” to estimate Model 1. Figure 2 shows an example code with explanations of its elements. Specifically, this function contains the specification of the correlation coefficients (`yi`), the sampling covariance matrices (`V`), and the data set in which these elements are stored (`data`). Next, the random effects within (`|ESID`) and between studies (`|StudyID`) are specified as a list. Given that Model 1 assumes correlation-specific estimates of these random effects, the option `~factor(Cell)` is added to this list. Suppose that the meta-analyst does have some evidence that the level-2 and level-3 random effects can be considered independent. Specifying an HCS structure within and between studies (`struc=c("HCS", "HCS")`), she may assume a correlation of zero within and zero between studies (`rho=0` and, respectively, `phi=0`). Next, the estimation procedure is set; in our example, we chose restricted maximum likelihood estimation (`method="REML"`). Finally, we specified the variable indicating the type of correlation as a moderator and turned off the

intercept of this meta-regression part (`mods=~factor(Cell)-1`). This way, we get estimates of each pooled correlation coefficient directly. Labelling the `rma.mv()` object

“`tas20.mlrmvrem1`” and using the `summary()` function, we obtained the following output of this model estimation:

```
## Multivariate Meta-Analysis Model (k = 880; method: REML)
##      logLik      Deviance      AIC      BIC      AICc
##      823.5375   -1647.0750  -1587.0750  -1444.0202  -1584.8581
##
## Variance Components:
## outer factor: ESID      (nlvls = 880)
## inner factor: factor(Cell) (nlvls = 10)
##
##      estim      sqrt k.lvl  fixed  level
## tau^2.1    0.0005  0.0232   88    no  I11-I12
## tau^2.2    0.0003  0.0184   88    no  I11-I17
## tau^2.3    0.0011  0.0332   88    no  I12-I17
## tau^2.4    0.0002  0.0136   88    no  I2-I11
## tau^2.5    0.0014  0.0373   88    no  I2-I12
## tau^2.6    0.0020  0.0450   88    no  I2-I17
## tau^2.7    0.0108  0.1041   88    no  I2-I4
## tau^2.8    0.0043  0.0657   88    no  I4-I11
## tau^2.9    0.0021  0.0463   88    no  I4-I12
## tau^2.10   0.0047  0.0686   88    no  I4-I17
## rho        0.0000                      yes
##
## outer factor: StudyID      (nlvls = 62)
## inner factor: factor(Cell) (nlvls = 10)
##
##      estim      sqrt k.lvl  fixed  level
```

```

## gamma^2.1      0.0038  0.0614      88      no  I11-I12
## gamma^2.2      0.0039  0.0626      88      no  I11-I17
## gamma^2.3      0.0043  0.0653      88      no  I12-I17
## gamma^2.4      0.0029  0.0541      88      no  I2-I11
## gamma^2.5      0.0034  0.0581      88      no  I2-I12
## gamma^2.6      0.0036  0.0601      88      no  I2-I17
## gamma^2.7      0.0051  0.0714      88      no  I2-I4
## gamma^2.8      0.0039  0.0627      88      no  I4-I11
## gamma^2.9      0.0082  0.0904      88      no  I4-I12
## gamma^2.10     0.0053  0.0726      88      no  I4-I17
## phi            0.0000                                yes
##
## Test for Residual Heterogeneity:
## QE(df = 870) = 12485.1269, p-val < .0001
## Test of Moderators (coefficients 1:10):
## QM(df = 10) = 6731.5291, p-val < .0001
##
## Model Results:
##
##              estimate      se      zval      pval      ci.lb      ci.ub
## factor(Cell)I11-I12    0.3288  0.0096  34.2474 <.0001  0.3100  0.3477 ***
## factor(Cell)I11-I17    0.3605  0.0095  37.8147 <.0001  0.3418  0.3792 ***
## factor(Cell)I12-I17    0.2947  0.0104  28.3118 <.0001  0.2743  0.3151 ***
## factor(Cell)I2-I11     0.4631  0.0083  55.8157 <.0001  0.4468  0.4793 ***
## factor(Cell)I2-I12     0.3656  0.0098  37.4558 <.0001  0.3465  0.3847 ***
## factor(Cell)I2-I17     0.4005  0.0103  38.9950 <.0001  0.3803  0.4206 ***
## factor(Cell)I2-I4      0.4327  0.0153  28.3126 <.0001  0.4028  0.4627 ***
## factor(Cell)I4-I11     0.3602  0.0118  30.4748 <.0001  0.3370  0.3834 ***
## factor(Cell)I4-I12     0.2919  0.0136  21.4181 <.0001  0.2652  0.3186 ***
## factor(Cell)I4-I17     0.3255  0.0130  25.1005 <.0001  0.3001  0.3509 ***
## ---

```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

First, the output contains information about the log-likelihood, deviance, and information criteria which can be used for further model comparisons. Second, the output displays the within-study variance estimates ($\tau^2_{.1} - \tau^2_{.10}$) and the between-study variance estimates ($\gamma^2_{.1} - \gamma^2_{.10}$) for each of the ten correlation coefficients. These estimates exhibited a considerable range, with some variances being close to zero and others as high as 0.0108 ($\tau^2_{.7}$). Third, the Q statistics are shown, indicating, for instance, that statistically significant (residual) heterogeneity existed in the data ($Q_E[870] = 12485.13, p < .01$). Fourth, the estimates of the weighted average correlation coefficients, their standard errors, z -values, p -values, and the respective confidence intervals are shown (section `Model results`). Notably, all pooled coefficients were positive and were statistically significantly different from zero. If meta-analysts wish to further obtain cluster-robust standard errors of the pooled correlation coefficient estimates (with primary studies as clusters, indicated by `StudyID`), they may use the convenience function `robust()` in “metafor” that calls the “clubSandwich” package:

```
## Cluster-robust standard errors

tas20.mlmvrem1.robust <- robust(tas20.mlmvrem1, cluster =
                                StudyID, clubSandwich = TRUE)

summary(tas20.mlmvrem1.robust)
```

Output:

```
## [...]

## Model Results:
```

##	estimate	se ¹	tval ¹	df ¹	pval ¹	ci.lb ¹
## factor(Cell) I11-I12	0.3288	0.0134	24.5620	57.95	<.0001	0.3020
## factor(Cell) I11-I17	0.3605	0.0129	27.8425	58.39	<.0001	0.3346


```

## factor(Cell) I12-I17      0.2947  0.0150  19.6470  58.39  <.0001  0.2647
## factor(Cell) I2-I11      0.4631  0.0111  41.6699  57.86  <.0001  0.4408
## factor(Cell) I2-I12      0.3656  0.0141  25.9749  57.71  <.0001  0.3374
## factor(Cell) I2-I17      0.4005  0.0138  28.9693  58.11  <.0001  0.3728
## factor(Cell) I2-I4       0.4327  0.0181  23.8985  57.12  <.0001  0.3965
## factor(Cell) I4-I11      0.3602  0.0152  23.6903  57.7   <.0001  0.3298
## factor(Cell) I4-I12      0.2919  0.0174  16.7327  59.75  <.0001  0.2570
## factor(Cell) I4-I17      0.3255  0.0168  19.3269  58.34  <.0001  0.2918
##
##                               ci.ub1
## factor(Cell) I11-I12    0.3556  ***
## factor(Cell) I11-I17    0.3864  ***
## factor(Cell) I12-I17    0.3247  ***
## factor(Cell) I2-I11     0.4853  ***
## factor(Cell) I2-I12     0.3938  ***
## factor(Cell) I2-I17     0.4281  ***
## factor(Cell) I2-I4      0.4690  ***
## factor(Cell) I4-I11     0.3906  ***
## factor(Cell) I4-I12     0.3268  ***
## factor(Cell) I4-I17     0.3592  ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 1) results based on cluster-robust inference (var-cov estimator: CR2,
##    approx t/F-tests and confidence intervals, df: Satterthwaite approx)

```

Table 2 summarizes the key estimates of Model 1. However, the previous output does not contain any information about the preference of Model 1 over alternative models. Hence, in the next section, we describe the estimation of Models 2-4 and perform model comparisons to select an appropriate working model for the meta-analytic pooling of the correlation matrices.

Models 2-4. The specification of the working models 2-4 is like that of Model 1. While Model 1 allowed for correlation-specific estimates of the heterogeneity variances, Models 2-4 restricted these estimates to be the same across correlations. This is reflected in the specification of the structures of level-2 and level-3 random effects. Table 1 shows sample codes to estimate Models 2-4 in “metafor”. For instance, in Model 2, the random-effects structures are specified as `struc=c("HCS", "CS")`, and the respective output contains the following variance

estimates:

```
## [...]
## Variance Components:
## outer factor: ESID          (nlvls = 880)
## inner factor: factor(Cell) (nlvls = 10)
##
##
```

	estim	sqrt	k.lvl	fixed	level
## tau^2.1	0.0005	0.0230	88	no	I11-I12
## tau^2.2	0.0004	0.0189	88	no	I11-I17
## tau^2.3	0.0012	0.0351	88	no	I12-I17
## tau^2.4	0.0002	0.0126	88	no	I2-I11
## tau^2.5	0.0013	0.0367	88	no	I2-I12
## tau^2.6	0.0020	0.0451	88	no	I2-I17
## tau^2.7	0.0118	0.1085	88	no	I2-I4
## tau^2.8	0.0044	0.0664	88	no	I4-I11
## tau^2.9	0.0044	0.0661	88	no	I4-I12
## tau^2.10	0.0055	0.0740	88	no	I4-I17
## rho	0.0000			yes	

```
##
## outer factor: StudyID      (nlvls = 62)
## inner factor: factor(Cell) (nlvls = 10)
##
```

```
##          estim    sqrt  fixed
## gamma^2    0.0038  0.0616    no
## phi        0.0000          yes
```

Notably, the within-study variance estimates are correlation-specific ($\tau^2_{.1} - \tau^2_{.10}$), and only one heterogeneity variance is estimated at the between level (γ^2). Similarly, Model 3 can be specified using the option `struc=c("CS", "HCS")` to implement the assumption of correlation-specific heterogeneity variances between studies but one uniform variance estimate within studies. Finally, the structure option in Model 4 specifies two compound structures `struc=c("CS", "CS")`. Please find the output of the estimation of Models 3 and 4 in the Supplementary Material S2. Table 2 shows the parameters of Models 2-4.

Model selection. To select a working model, we compared the information criteria and performed likelihood-ratio tests across Models 1-4. Table 2 contains the respective values of the AIC, BIC, and a corrected AIC. Consistently across these information criteria, Model 2 showed the lowest values and thus the best fit. Moreover, a likelihood-ratio test between Models 1 and 2 suggested that both models did not differ statistically significantly in their fit, $\chi^2(9) = 6.4, p = .70$. We performed this test using the `anova()` function:

```
## Model 1 vs. Model 2
anova(tas20.mlmvrem1, tas20.mlmvrem2)
```

Output:

##	df	AIC	BIC	AICc	logLik	LRT	pval	QE
## Full	30	-1587.0750	-1444.0202	-1584.8581	823.5375			12485.1269
## Reduced	21	-1598.6390	-1498.5007	-1597.5494	820.3195	6.4360	0.6956	12485.1269

Further likelihood-ratio tests suggested the preference of Model 2 over Models 3 and 4. This preference of Model 2 indicated that the between-level variance estimates did not differ

4.4 Moderator Analyses

Categorical moderator. There are several ways of exploring the possible moderating effect of a categorical variable using the MLMVREM approach. One way is to extend the model specification of Model 2 in “metafor” by specifying a crossed effect between the moderator and the variable that indicates the type of correlation (i.e., `factor(Cell)`). This specification still contains the direct effect of `factor(Cell)`, which serve as the intercepts, that is, the expected weighted average correlations for each type of correlations when the sample is non-clinical. Hence, the model specification does no longer contain an overall intercept (-1).

[illegible]

```

method = "REML",

mods = ~ factor(Cell) - 1 +

factor(Cell):factor(ClinicalBinary))

## Model summary

summary(tas20.clinical)

```

Output:

```

## [...]

## Model Results:

##
##
## estimate      se      zval      pval
## factor(Cell) I11-I12      0.3347  0.0105  31.9536 <.0001
## factor(Cell) I11-I17      0.3645  0.0102  35.7235 <.0001
## factor(Cell) I12-I17      0.2988  0.0112  26.7920 <.0001
## factor(Cell) I2-I11      0.4683  0.0097  48.1327 <.0001
## factor(Cell) I2-I12      0.3684  0.0111  33.1291 <.0001
## factor(Cell) I2-I17      0.4046  0.0115  35.0509 <.0001
## factor(Cell) I2-I4      0.4536  0.0165  27.4535 <.0001
## factor(Cell) I4-I11      0.3793  0.0128  29.5723 <.0001
## factor(Cell) I4-I12      0.3024  0.0133  22.6891 <.0001
## factor(Cell) I4-I17      0.3430  0.0137  24.9700 <.0001
## factor(Cell) I11-I12:factor(ClinicalBinary) 1 -0.0255  0.0162  -1.5815  0.1138
## factor(Cell) I11-I17:factor(ClinicalBinary) 1 -0.0182  0.0152  -1.1948  0.2322
## factor(Cell) I12-I17:factor(ClinicalBinary) 1 -0.0186  0.0181  -1.0278  0.3041
## factor(Cell) I2-I11:factor(ClinicalBinary) 1 -0.0237  0.0136  -1.7370  0.0824
## factor(Cell) I2-I12:factor(ClinicalBinary) 1 -0.0139  0.0180  -0.7703  0.4411
## factor(Cell) I2-I17:factor(ClinicalBinary) 1 -0.0183  0.0190  -0.9636  0.3352
## factor(Cell) I2-I4:factor(ClinicalBinary) 1 -0.0839  0.0308  -2.7213  0.0065
## factor(Cell) I4-I11:factor(ClinicalBinary) 1 -0.0766  0.0226  -3.3906  0.0007
## factor(Cell) I4-I12:factor(ClinicalBinary) 1 -0.0394  0.0236  -1.6701  0.0949
## factor(Cell) I4-I17:factor(ClinicalBinary) 1 -0.0681  0.0246  -2.7751  0.0055

```

```
## [...]
```

The estimates of the model parameters indicate statistically significant and negative moderator effects of the type of the sample on three correlation coefficients: I2-I4 ($B = -0.084$, $SE = 0.031$, $p = .007$), I4-I11 ($B = -0.077$, $SE = 0.023$, $p = .001$), and I4-I17 ($B = -0.068$, $SE = 0.025$, $p = .006$). Hence, these correlations tended to be smaller for clinical samples than for non-clinical samples. The cluster-robust standard errors supported this result. Next, we computed the variance explanations for each variance component.

```
## Level: Primary studies
100*max((1-tas20.clinical$gamma2/tas20.mlmvrem2$gamma2), 0)

## Level: Effect sizes
100*max((1-tas20.clinical$tau2[8]/tas20.mlmvrem2$tau2[8]), 0)
```

Output:

```
## [1] 1.56858
## [1] 11.57651
```

About 1.6% of the between-study variation in effect sizes were explained by the type of the sample, while about 11.6% of the within-study variation of the correlation I4-I11 were explained.

Another way of exploring the effect sizes across the moderator's categories is to estimate the subgroup-specific effect sizes directly. To achieve this, the moderator option in the above-described specification is changed into:

```
mods =~ factor(Cell):factor(ClinicalBinary)-1
```

Consequently, meta-analysts obtain the effect size estimates of each correlation for studies with non-clinical (code 0) and clinical samples (code 1).

Output:

```
## [...]
## Model Results:
##
##               estimate      se      zval      pval
## factor(Cell)I11-I12:factor(ClinicalBinary)0    0.3347  0.0105  31.9536  <.0001
## factor(Cell)I11-I17:factor(ClinicalBinary)0    0.3645  0.0102  35.7235  <.0001
## [...]
## factor(Cell)I11-I12:factor(ClinicalBinary)1    0.3092  0.0154  20.1282  <.0001
## factor(Cell)I11-I17:factor(ClinicalBinary)1    0.3463  0.0147  23.6034  <.0001
## [...]
```

However, this way, the similarities and differences in effects are not tested statistically.

Continuous moderator. Introducing continuous moderators works in the same way as adding categorical moderators. However, given that continuous moderators do not have discrete categories, subgroup-specific estimates of the effect sizes are not accessible. In this example, we chose the proportion of female study participants as a possible moderator without any variance-stabilizing transformation for better interpretability and to avoid issues related to back-transformations.^{39,40} The model specification code is similar to that of the categorical example and contains a modified mods option:

```
mods =~ factor(Cell) - 1 + factor(Cell):PropFemale
```

The results indicated that two correlations were moderated by the proportion of female study participants: I11-I17 ($B = -0.001$, $SE = 0.001$, $p = .001$) and I2-I4 ($B = 0.014$, $SE = 0.001$, $p = .018$). Only the latter effect was supported by the cluster-robust standard errors. Please find the detailed output in the Supplementary Material S2.

4.5 Sensitivity Analyses

As noted earlier, some decisions on model parameters in the MLMV-REM may be informed by prior research or meta-analysts' substantive expertise. To examine the sensitivity of

the meta-analytic findings, we explored how different, a-priori fixed values of r (0.0, 0.2, 0.4, 0.6, 0.8)¹, ρ (0.0, 0.5, 1.0), and ϕ (0.0, 0.5, 1.0) affected (a) model selection, and (b) key parameter estimates of Model 1, such as the weighted average correlation coefficients or heterogeneity variances.

Overall, Models 1 and 2 were suitable meta-analytic working models to pool the correlation matrices across the conditions. When ρ and ϕ were small, Model 2 was likely to be preferred over Model 1. In Model 1, for a given correlation coefficient r , the weighted average correlation coefficients were not sensitive to the choices of ρ and ϕ . Nonetheless, the relation between ρ (or, respectively, ϕ) and the within-study variance estimates was mostly U-shaped, and larger values of ρ and ϕ were associated with higher between-study variance estimates. For given values of ρ and ϕ , the weighted average correlation coefficients were stable across the values of r . Again, the relation between the within-study variance estimates and r was U-shaped, while the relation was negative linear for the between-study variance estimates. The detailed results of these sensitivity analyses are shown in Supplementary Material S2 and S3. Hence, the heterogeneity estimates in Model 1 were sensitive to the choices of r , ρ , and ϕ , and simulation studies are needed to examine them systematically across more conditions.

5. Discussion and Some Recommendations

In this tutorial paper, we have described a multilevel, multivariate, and random-effects approach to addressing hierarchical effect size multiplicity when pooling correlation matrices across primary studies. This approach can be specified via several working models under various assumptions on the correlation-specific within- and between-study variances. We illustrated the

¹ Notably, the sampling covariance matrices that were based on the assumption $r = 1.0$ were non-positive definite. Hence, we did not specify this value in our sensitivity analyses.

implementation of four working models using the R package “metafor” and the respective cluster-robust standard errors using the R package “clubSandwich”. This illustration contained a description of the models’ assumptions, the R code to specify and estimate them, and the resultant R output.

As noted earlier, the MLMV-REM approach allows meta-analysts to model within- and between-study heterogeneity, either specific for each correlation coefficient or under the assumption of equal heterogeneity across correlations at the respective levels of analysis. To our best knowledge, this approach is the first to make explicit these heterogeneity variances—existing approaches either provide one heterogeneity variance estimate for all correlation coefficients at the within and between level¹³ or correlation-specific heterogeneity variance estimates at a single level of analysis.⁴ We therefore consider the MLMV-REM approach an extension and integration of the existing approaches. This approach does not only provide meta-analysts with information about associations among multiple constructs, concepts, or variables and their heterogeneity but also has the potential to advance meta-analytic structural equation and network modeling, in which meta-analysts can test their theories and hypotheses about these associations.¹¹

The working models within the MLMV-REM approach represented several assumptions meta-analysts may have on the degree of sampling error dependence and the structure of the within- and between-level random effects. This does not only allow meta-analysts to test hypotheses on these structures (e.g., equal vs. correlation-specific heterogeneity variances) but enables them to select a model for the pooling of correlation matrices that represents the structure and nature of their meta-analytic data. We recommend selecting a meta-analytic working model by estimating and comparing models with several assumptions, evaluating model

fit, and considering substantive aspects, such as theories or existing evidence from other studies. Moreover, we encourage meta-analysts to also consider the feasibility of working models with many correlation-specific parameter estimates. Sometimes, too few studies and correlation matrices are available to reliably estimate such models with sufficient statistical power.⁴

The MLMV-REM approach has several limitations and leaves some open questions that point into future research directions. First, while this approach offers a flexible specification of model assumptions on the structure of random effects, model selection becomes not only an opportunity but a necessity. Oftentimes, meta-analytic data contain too few studies to estimate all possible variance components of correlations within a correlation matrices⁴. In this situation, meta-analysts would need to strive for model parsimony and select a model that may not capture all correlation-specific heterogeneity variances within and between studies but some. However, modern methods such as Bayesian meta-analysis may address the challenges of few available studies, and simulations studies could shed more light on the meta-analytic sample size requirements and the factors associated with them.

Second, the current implementation of the MLMV-REM approach synthesizes correlation matrices meta-analytically and estimates the respective heterogeneity (co-)variances. Hence, meta-analysts who are interested in testing hypotheses and theories about the relations among variables would then submit these elements to structural equation or network modeling. This two-stage approach addresses hierarchical effect size multiplicity at stage 1. We encourage meta-analysts to develop and explore alternative approaches that may integrate these two analytic stages into one stage. Such an integration could improve estimation efficiency and reduce possible bias in the variance estimates.⁴¹

Third, the MLMV-REM approach requires substantive decisions on the model specification, such as the structure of random effects or the sampling covariance matrix \mathbf{V} and the choices of the correlations between sampling errors in the matrix \mathbf{V} and the correlations between random effects in the matrices $\mathbf{\Sigma}$ and $\mathbf{\Omega}$. These decisions also include the choices of the correlations among sampling errors and the within- and between-level random effects. Oftentimes, researchers do not have specific hypotheses or prior evidence on these correlations, which leaves them to guessing these values. In future model specifications, however, it may be possible to estimate such correlations.^{6,30,33} Overall, we recommend backing the analytical decisions meta-analysts have to make by conducting sensitivity analyses to examine the impact of these decisions on meta-analytic estimates.

Fourth, effect size multiplicity may be hierarchical, correlational, or both.⁶ In our view, the MLMV-REM we have proposed is best suited for meta-analytic data sets with hierarchical multiplicity, in which multiple independent correlation matrices are nested in primary studies. In such a scenario, the variance components have a clear interpretation as within- and between-study heterogeneity variances of the correlation coefficients. In scenarios in which multiple correlation matrices are available for the same sample within a study, our proposed approach may not be best-suited and may only approximate the true weighted average correlation coefficients and their heterogeneity. Specifically, the MLMV-REM approach we proposed assumes zero correlation between correlation matrices within studies. Current methodological developments in the meta-analytic pooling of univariate effect sizes, such as the working models proposed by Pustejovsky and Tipton⁶, and meta-analytic models describing longitudinal data³⁰ could serve as starting points to develop an approach that handles meta-analytic data with correlational multiplicity.

In our view, the MLMV-REM approach represents *one* possible approach to address hierarchical effect size multiplicity when pooling correlation matrices meta-analytically, and alternative approaches may be developed in the future. This approach will have to stand trial against other approaches, and, in our view, investigating its accuracy, efficiency, and statistical power is a natural next step towards evaluating its performance.

References

1. Schroeders U, Kubera F, Gnambs T. The Structure of the Toronto Alexithymia Scale (TAS-20): A Meta-Analytic Confirmatory Factor Analysis. *Assessment*. 2022;29(8):1806-1823. doi:10.1177/10731911211033894
2. Becker BJ, Aloe AM, Cheung MW-L. Meta-analysis of correlations, correlation matrices, and their functions. In: Schmid CH, Stijnen T, White IR, eds. *Handbook of meta-analysis*. Routledge/Taylor & Francis Group; 2021:347–369.
3. Cheung MW-L, Cheung SF. Random-effects models for meta-analytic structural equation modeling: review, issues, and illustrations. *Research Synthesis Methods*. 2016/06/01 2016;7(2):140-155. doi:10.1002/jrsm.1166
4. Cheung MW-L. *Meta-Analysis: A Structural Equation Modeling Approach*. John Wiley & Sons Ltd.; 2015.
5. Jak S. *Meta-Analytic Structural Equation Modelling*. Springer International Publishing; 2015.
6. Pustejovsky JE, Tipton E. Meta-Analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science*. 2022;23:425-438. doi:10.1007/s11121-021-01246-3
7. Scherer R, Siddiq F, Tondeur J. The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*. 2019/01/01/ 2019;128:13-35. doi:10.1016/j.compedu.2018.09.009

8. Campos DG, Cheung MW-L, Scherer R. A primer on synthesizing individual participant data obtained from complex sampling surveys: A two-stage IPD meta-analysis approach. *Psychological Methods*. 2023;doi:10.1037/met0000539
9. Fernández-Castilla B, Jamshidi L, Declercq L, Beretvas SN, Onghena P, Van den Noortgate W. The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*. 2020/10/01 2020;52(5):2031-2052. doi:10.3758/s13428-020-01373-9
10. Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*. 2010/01/01 2010;1(1):39-65. doi:10.1002/jrsm.5
11. Stolwijk I, Jak S, Eichelsheim V, Hoes M. Dealing With Dependent Effect Sizes in MASEM. *Zeitschrift für Psychologie*. 2022/01/01 2022;230(1):16-32. doi:10.1027/2151-2604/a000485
12. Ke Z, Zhang Q, Tong X. Bayesian Meta-Analytic SEM: A One-Stage Approach to Modeling Between-Studies Heterogeneity in Structural Parameters. *Structural Equation Modeling: A Multidisciplinary Journal*. 2019/05/04 2019;26(3):348-370. doi:10.1080/10705511.2018.1530059
13. Wilson SJ, Polanin JR, Lipsey MW. Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods*. 2016/06/01 2016;7(2):121-139. doi:10.1002/jrsm.1199
14. Sheng Z, Kong W, Cortina JM, Hou S. Analyzing matrices of meta - analytic correlations: Current practices and recommendations. *Research Synthesis Methods*. 2016;7(2):187-208. doi:10.1002/jrsm.1206

15. Scherer R, Siddiq F, Nilsen T. The potential of international large-scale assessments for meta-analyses in education. *Large-scale Assessments in Education*. 2024/02/14 2024;12(1):4. doi:10.1186/s40536-024-00191-1
16. López-López JA, Page MJ, Lipsey MW, Higgins JPT. Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods*. 2018/09/01 2018;9(3):336-351. doi:10.1002/jrsm.1310
17. Cheung MW-L. A Guide to Conducting a Meta-Analysis with Non-Independent Effect Sizes. *Neuropsychology Review*. 2019/12/01 2019;29(4):387-396. doi:10.1007/s11065-019-09415-6
18. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd; 2009.
19. Becker BJ, Aloe AM, Cheung MW-L. Meta-analysis of correlations, correlation matrices, and their functions. In: Schmid CH, Stijnen T, White IR, eds. *Handbook of Meta-Analysis*. 1st ed. CRC Press Routledge/Taylor & Francis Group; 2021:347-369:chap 16.
20. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*. 2013/06/01 2013;45(2):576-594. doi:10.3758/s13428-012-0261-6
21. Harrer M, Cuijpers P, Furukawa TA, Ebert DD. *Doing Meta-Analysis in R: A Hands-on Guide*. Chapman & Hall/CRC Press; 2022.
22. Moeyaert M, Ugille M, Beretvas NS, Ferron J, Bunuan R, Van Den Noortgate W. Methods for dealing with multiple outcomes in meta-analysis: A comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology*. 2017;20(6):559–572. doi:10.1080/13645579.2016.1252189

23. Riley RD. Multivariate meta-analysis: The effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2009;172(4):789–811. doi:10.1111/j.1467-985X.2008.00593.x
24. Becker BJ, Aloe AM. Model-based meta-analysis and related approaches. In: Cooper H, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis and Meta-Analysis*. 3rd ed. Russell Sage Foundation; 2019:339-366:chap 16.
25. Becker BJ. Using Results From Replicated Studies to Estimate Linear Models. *Journal of Educational Statistics*. 1992/12/01 1992;17(4):341-362. doi:10.3102/10769986017004341
26. Jackson D, White IR, Price M, Copas J, Riley RD. Borrowing of strength and study weights in multivariate and network meta-analysis. *Statistical Methods in Medical Research*. 2017/12/01 2015;26(6):2853-2868. doi:10.1177/0962280215611702
27. Jackson D, White IR, Riley RD. Multivariate meta-analysis. In: Schmid CH, Stijnen T, White IR, eds. *Handbook of Meta-Analysis*. CRC Press; 2021:163-186:chap 9.
28. Ahn S, Ames AJ, Myers ND. A Review of Meta-Analyses in Education: Methodological Strengths and Weaknesses. *Review of Educational Research*. 2012/12/01 2012;82(4):436-476. doi:10.3102/0034654312458162
29. Card NA. *Applied meta-analysis for social science research*. Guilford Press; 2012.
30. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. 2010;36(3)
31. Cheung MW-L. Multivariate Meta-Analysis as Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*. 2013/07/01 2013;20(3):429-454. doi:10.1080/10705511.2013.797827

32. Cheung MW-L, Chan W. Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*. 2005;10(1):40-64. doi:10.1037/1082-989X.10.1.40
33. McShane BB, Böckenholt U. Multilevel multivariate meta-analysis made easy: An introduction to MLMVmeta. *Behavior Research Methods*. 2023/08/01 2023;55(5):2367-2386. doi:10.3758/s13428-022-01892-7
34. Olkin I, Finn JD. Testing correlated correlations. *Psychological Bulletin*. 1990;108(2):330–333. doi:10.1037/0033-2909.108.2.330
35. Wei Y, Higgins JPT. Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine*. 2013/03/30 2013;32(7):1191-1205. doi:10.1002/sim.5679
36. Snijders TAB, Bosker RJ. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 2nd ed. Sage Publications; 2012.
37. Rights JD, Sterba SK. Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods*. 2019;24(3):309–338. doi:10.1037/met0000184
38. *clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections*. R package version 0.5.3. 2021. <https://CRAN.R-project.org/package=clubSandwich>
39. Lin L, Xu C. Arcsine-based transformations for meta-analysis of proportions: Pros, cons, and alternatives. *Health Science Reports*. 2020/09/01 2020;3(3):e178. doi:<https://doi.org/10.1002/hsr2.178>
40. Schwarzer G, Chemaitelly H, Abu-Raddad LJ, Rücker G. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single

proportions. *Research Synthesis Methods*. 2019/09/01 2019;10(3):476-483.

doi:<https://doi.org/10.1002/jrsm.1348>

41. Jak S, Cheung MW-L. Meta-analytic structural equation modeling with moderating effects on SEM parameters. *Psychological Methods*. 2020;25(4):430-455.

doi:10.1037/met0000245

Tables

Table 1*Specification of Meta-Analytic Working Models with the MLMV-REM Approach*

Working model	Effect sizes	Within-study heterogeneity	Between-study heterogeneity	Random-effects correlations	Example specification in metafor
Model 1	Correlation-specific	Correlation-specific, HCS structure	Correlation-specific, HCS structure	Fixed ρ and ϕ	<code>rma.mv(yi = Correlation, V = V, data = tas20, random = list(~ factor(Cell) ESID, ~ factor(Cell) StudyID), method = "REML", struc = c("HCS", "HCS"), rho = 0, phi = 0, mods = ~ factor(Cell) - 1)</code>
Model 2	Correlation-specific	Correlation-specific, HCS structure	Constrained, CS structure	Fixed ρ and ϕ	<code>rma.mv(yi = Correlation, V = V, data = tas20, random = list(~ factor(Cell) ESID, ~ factor(Cell) StudyID), method = "REML", struc = c("HCS", "CS"), rho = 0, phi = 0, mods = ~ factor(Cell) - 1)</code>
Model 3	Correlation-specific	Constrained, CS structure	Correlation-specific, HCS structure	Fixed ρ and ϕ	<code>rma.mv(yi = Correlation, V = V, data = tas20, random = list(~ factor(Cell) ESID, ~ factor(Cell) StudyID), method = "REML", struc = c("CS", "HCS"), rho = 0, phi = 0, mods = ~ factor(Cell) - 1)</code>
Model 4	Correlation-specific	Constrained, CS structure	Constrained, CS structure	Fixed ρ and ϕ	<code>rma.mv(yi = Correlation, V = V, data = tas20, random = list(~ factor(Cell) ESID, ~ factor(Cell) StudyID),</code>

```
method = "REML", struc = c("CS",  
"CS"), rho = 0, phi = 0, mods = ~  
factor(Cell) - 1)
```

Note. The variable “Cell” is defined as a factor indicating the different types of correlations in the correlation matrix (e.g., “I2-I4”, “I11-I12”). “Correlation-specific” means that the parameters are estimated for each of the ten correlations. “Constrained” means that the parameter is constrained to being equal across the types of correlations. “Correlation” contains the extracted Pearson correlation coefficients. “V” represents the sampling covariance matrix. “Cell” represents a nominal variable of the names of the different types of correlations. The random-effects correlations ρ and ϕ refer to the correlations among random effects within and, respectively, between studies. These correlations were fixed to zero here. “HCS” specifies a heteroscedastic compound symmetric structure of the random effects, while “CS” specifies a compound symmetric structure.³⁰

Table 2

Weighted Average Effect Sizes and Variance Components in Working Models 1-4 with Independent Random Effects (with $\rho = \phi = 0$)

Item-item correlation	Weighted average effect sizes			Heterogeneity		Information criteria		
	\bar{r}	SE	95 % CI	Within-study heterogeneity τ^2	Between-study heterogeneity γ^2	AIC	BIC	AICc
Model 1 (Correlation-specific effect sizes, within- and between-study heterogeneity)						-1587.1	-1444.0	-1584.9
I11-I12	0.329	0.013	[0.302, 0.356]	0.001	0.004			
I11-I17	0.360	0.013	[0.335, 0.386]	0.000	0.004			
I12-I17	0.295	0.015	[0.265, 0.325]	0.001	0.004			
I2-I11	0.463	0.011	[0.441, 0.485]	0.000	0.003			
I2-I12	0.366	0.014	[0.337, 0.394]	0.001	0.003			
I2-I17	0.400	0.014	[0.373, 0.428]	0.002	0.004			
I2-I4	0.433	0.018	[0.396, 0.469]	0.011	0.005			
I4-I11	0.360	0.015	[0.330, 0.391]	0.004	0.004			
I4-I12	0.292	0.017	[0.257, 0.327]	0.002	0.008			
I4-I17	0.325	0.017	[0.292, 0.359]	0.005	0.005			
Model 2 (Correlation-specific effect sizes and within-study heterogeneity, constrained between-study heterogeneity)						-1598.6	-1498.5	1597.5
I11-I12	0.329	0.013	[0.302, 0.355]	0.001	0.004			
I11-I17	0.360	0.013	[0.335, 0.386]	0.000	0.004			
I12-I17	0.295	0.015	[0.265, 0.325]	0.001	0.004			
I2-I11	0.463	0.011	[0.441, 0.485]	0.000	0.004			
I2-I12	0.365	0.014	[0.337, 0.394]	0.001	0.004			
I2-I17	0.400	0.014	[0.373, 0.428]	0.002	0.004			
I2-I4	0.433	0.018	[0.397, 0.469]	0.012	0.004			
I4-I11	0.360	0.015	[0.330, 0.390]	0.004	0.004			
I4-I12	0.293	0.017	[0.258, 0.327]	0.004	0.004			
I4-I17	0.326	0.017	[0.292, 0.360]	0.005	0.004			

Model 3 (Correlation-specific effect sizes and between-study heterogeneity, constrained within-study heterogeneity)						-1553.9	-1453.8	-1552.8
I11-I12	0.328	0.014	[0.301, 0.355]	0.002	0.003			
I11-I17	0.360	0.013	[0.334, 0.387]	0.002	0.002			
I12-I17	0.295	0.015	[0.265, 0.325]	0.002	0.003			
I2-I11	0.461	0.012	[0.438, 0.485]	0.002	0.002			
I2-I12	0.366	0.014	[0.337, 0.394]	0.002	0.003			
I2-I17	0.401	0.014	[0.373, 0.428]	0.002	0.003			
I2-I4	0.434	0.019	[0.396, 0.472]	0.002	0.013			
I4-I11	0.360	0.015	[0.329, 0.391]	0.002	0.006			
I4-I12	0.292	0.017	[0.257, 0.327]	0.002	0.008			
I4-I17	0.325	0.017	[0.291, 0.359]	0.002	0.007			
Model 4 (Correlation-specific effect sizes, constrained within- and between-study heterogeneity)						-1530.6	-1473.4	-1530.2
I11-I12	0.328	0.013	[0.302, 0.355]	0.003	0.005			
I11-I17	0.361	0.013	[0.335, 0.387]	0.003	0.005			
I12-I17	0.295	0.015	[0.265, 0.325]	0.003	0.005			
I2-I11	0.461	0.011	[0.439, 0.484]	0.003	0.005			
I2-I12	0.365	0.014	[0.338, 0.393]	0.003	0.005			
I2-I17	0.401	0.014	[0.373, 0.428]	0.003	0.005			
I2-I4	0.436	0.019	[0.398, 0.473]	0.003	0.005			
I4-I11	0.361	0.015	[0.330, 0.391]	0.003	0.005			
I4-I12	0.293	0.017	[0.258, 0.328]	0.003	0.005			
I4-I17	0.326	0.017	[0.292, 0.360]	0.003	0.005			

Note. \bar{r} = Weighted average Pearson correlation coefficient with cluster-robust standard error (*SE*) and 95% confidence interval (*CI*). “I” refers to the item in the TAS20. Correlations are labelled using the item numbers (e.g., “I11-I12” refers to the correlation between items 11 and 12). AIC = Akaike’s Information Criterion, BIC = Bayesian Information Criterion, AICc = Corrected AIC³⁰. All models assume independent random effects (i.e., $\rho = \phi = 0$) and block-diagonal sampling covariance matrices.

Figures

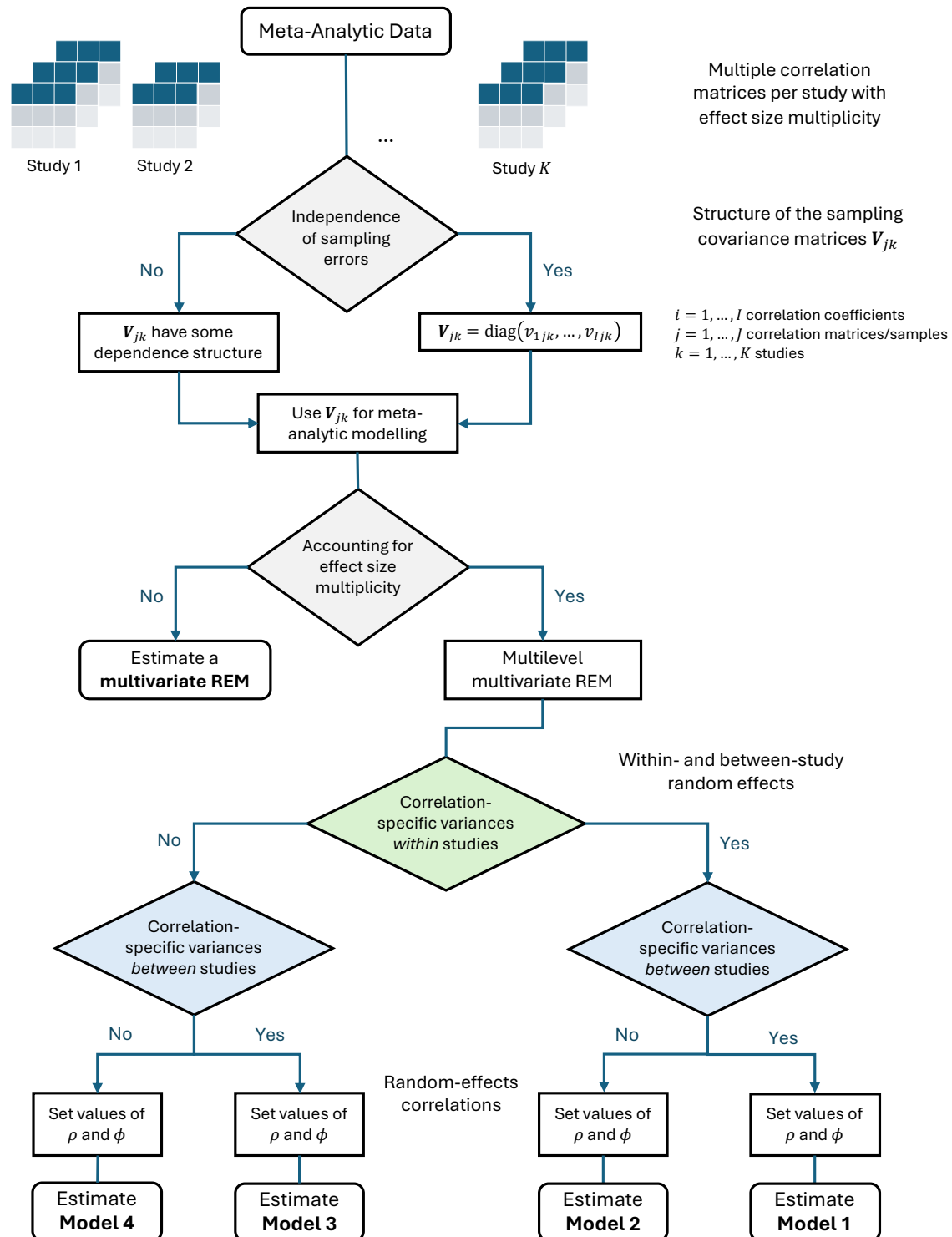
Figure 1*Decision Scheme for Selecting a Working Model*

Figure 2

Elements of the Analytic Code to Specify Working Model 1 in the R package “metafor”

