European Journal of Psychology and Educational Research
Volume 7, Issue 2, 55 - 64.
ISSN: 2589-949X
https://www.ejper.com

Using R for Multivariate Meta-analysis on Educational Psychology Data: A Method Study

Gamon Savatsomboon* Mahasarakham University, THAILAND
Prasert Ruannakarn Mahasarakham University, THAILAND
Phamornpun Yurayat Mahasarakham University, THAILAND
Ong-art Chanprasitchai Mahasarakham University, THAILAND
Jibon Kumar Sharma Leihaothabam Manipur University, INDIA

Received: December 2, 2023 ▪ Revised: February 16, 2024 ▪ Accepted: April 24, 2024

Abstract: Using R to conduct univariate meta-analyses is becoming common for publication. However, R can also conduct multivariate meta-analysis (MMA). However, newcomers to both R and MMA may find using R to conduct MMA daunting. Given that, R may not be easy for those unfamiliar with coding. Likewise, MMA is a topic of advanced statistics. Thus, it may be very challenging for most newcomers to conduct MMA using R. If this holds, this can be viewed as a practice gap. In other words, the practice gap is that researchers are not capable of using R to conduct MMA in practice. This is problematic. This paper alleviates this practice gap by illustrating how to use R (the metaSEM package) to conduct MMA on educational psychology data. Here, the metaSEM package is used to obtain the required MMA text outputs. However, the metaSEM package is not capable of producing the other required graphical outputs. As a result, the metafor package is also used as a complimentary to generate the required graphical outputs. Ultimately, we hope that our audience will be able to apply what they learn from this method paper to conduct MMA using R in their teaching, research, and publication.

Keywords: *Educational psychology data, metaSEM package, multivariate meta-analysis, R, tutorial.*

To cite this article: Savatsomboon, G., Ruannakarn, P., Yurayat, P., Chanprasitchai, O., & Leihaothabam, J. (2024). Using R for multivariate meta-analysis on educational psychology data: A method study. *European Journal of Psychology and Educational Research, 7*(2), 55-64. https://doi.org/10.12973/ejper.7.2.55

Introduction

Using R to conduct univariate meta-analyses is becoming common for publication. However, R can also be used to conduct multivariate meta-analysis (MMA). R is popular because it offers several benefits. R is free. Several R packages can run MMA. These packages, in combination, can run complete MMA. Importantly, there are R communities that share resources on how to conduct the MMA. However, newcomers to both R and MMA may find it daunting to use R to conduct MMA. Given that, R may not be easy for those who are not familiar with coding. Likewise, MMA is a topic of advanced statistics. Thus, it may be very challenging for most newcomers to conduct MMA using R. If this holds, this can be viewed as a practice gap. In other words, the practice gap is that researchers are not capable of using R to conduct MMA in practice. This is problematic. Thus, this paper alleviates this practice gap by illustrating how to use R (the metaSEM package) to conduct MMA on educational psychology data. We use the metaSEM package authored by Cheung (2015) to conduct the MMA. In addition, we use the built-in dataset that comes with the metaSEM package. We also traced back to the paper of Berkey et al. (1998) that holds the dataset that the metaSEM package adapted from. We also traced back to the root source of the dataset (raw data) that Berkey et al. (1998) used to compute their effect sizes. In summary, we synthesize critical resources to put together a paper on how to use R to conduct the MMA, using a suitable (educational psychology) dataset along with the story behind the dataset. Here, the metaSEM package is used to obtain the required MMA text outputs. However, the metaSEM package is not capable of producing the required graphical outputs. As a result, the metaphor package is also used as a complimentary to generate the required graphical outputs. Ultimately, we hope that our audience will be able to apply what they learn from this paper to conduct MMA using R in their teaching, research,

* **Corresponding author:**
Gamon Savatsomboon, Mahasarakham University, Thailand. ✉ gamon.s@msu.ac.th

and publication. Finally, we hope that newcomers to R and MMA will apply what they learn from this paper in their teaching, research, and publication.

Literature Review

Common Categories of Meta-analysis

There are six common categories of meta-analysis (see Figure 1). The first category is univariate meta-analysis. The second category is multivariate meta-analysis. The third category is network meta-analysis. The fourth category is multilevel meta-analysis. The fifth category is structural equation meta-analysis. Finally, the sixth category is Bayesian meta-analysis. The categorization presented in Figure 1 is arranged by the authors of this paper. Thus, there might be other different arrangements by other authors. But, we have not come across one. However, the focus of this paper is on multivariate meta-analysis (see Figure 1).

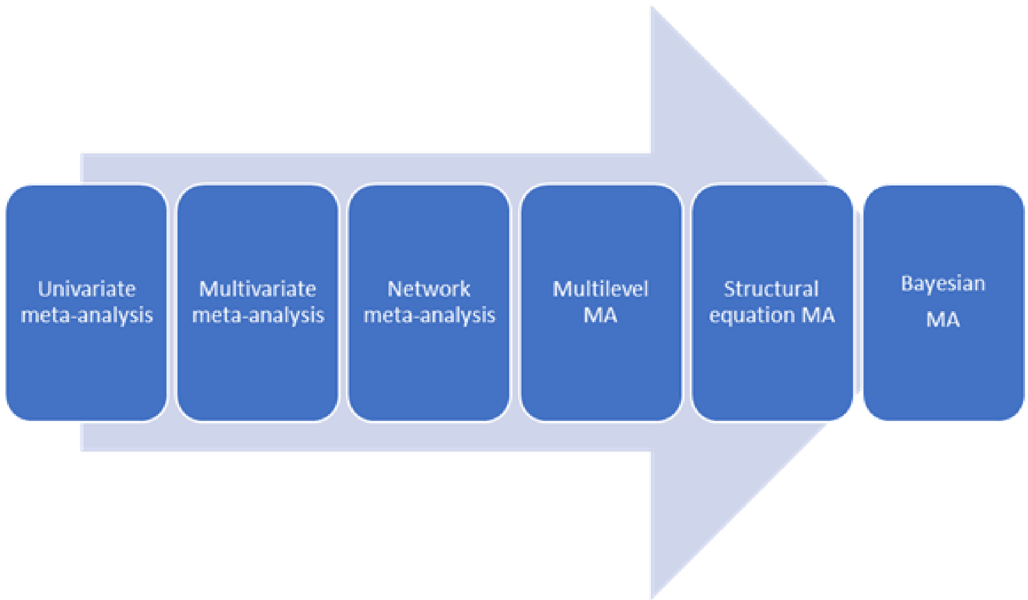

Figure 1. Common Categories of Meta-analysis (Savatsomboon et al., 2024)

Univariate Meta-analysis

For Figure 1, we largely draw upon the work of Savatsomboon et al. (2024). Their work has conceptual frameworks to accommodate each type of meta-analysis. According to Savatsomboon et al. (2024), a univariate meta-analysis is a statistical technique used to summarize the results of multiple studies investigating the same question. However, univariate meta-analysis (UMA) only examines one outcome variable. UMA takes the findings from each study and integrates them into a single overall estimate of the effect. Studies with larger sample sizes are weighted more heavily in the analysis. If our audience wants to learn more about UMA, the work of Savatsomboon et al. (2024) is a great resource.

Multivariate Meta-analysis

Gasparrini et al. (2012) state that MMA is an extension of UMA. MMA allows multiple outcomes of meta-analyses to be analyzed in a single study (see Figure 3). This is in line with how Hattle et al. (2022) describe MMA. Let's place our focus on a UMA before moving on to the MMA. It is critical to point out that UMA focuses on only one outcome (pooled effect size). Let's further elaborate on this. Based on Figure 3, if a meta-analysis study only focuses on the correlation between classroom management self-efficacy (CMSE) and emotional exhaustion (EE), it would only be considered a UMA. In short, the UMA focuses on one outcome (pooled effect size). For example, the one pooled effect size is the correlation between CMSE and EE. On the other hand, if a meta-analysis study focuses on multiple outcomes (pooled effect sizes), this would be considered an MMA. The multiple pooled effect sizes are the three relationships between CMSE and EE, CMSE and DP (depersonalization), and CMSE and PA (personal accomplishment) (see Figure 3).

PICO, Systematic Review, and PRISMA in Practice

PICO is required for MMA because it helps identify the research question. In other words, what effect sizes should be extracted (obtained), and why. Let's illustrate this claim. PICO stands for participants (P), intervention (I), control group (C), and outcome (O) (Schiavenato & Chu, 2021). Let's apply this to our MMA example case used in this paper. The example case presented in this paper draws upon the work of Aloe et al. (2014). Based on their work, the participants (P) are teachers. The intervention (I) is classroom management self-efficacy (CMSE). Here, the control group (C) is not applicable, because there is no comparison between experiment vs. control groups. The outcome (O) is a research question, what are the overall strengths of the relationships between classroom management self-efficacy (CMSE) and the three dimensions of burnout (emotional exhaustion (EE), depersonalization (DP), and (lowered) personal accomplishment (PA))? Again, MMA allows multiple effect sizes to be analyzed. In our case, we have three multiple pooled effect sizes (see Figure 3). Words of wisdom, it may be a good idea to draw a conceptual framework for a meta-study because it helps researchers see clearly what effect sizes need to be obtained for a study's systematic review.

Now, let's move on to a systematic review. According to Ahn and Kang (2018), a systematic review comprises a plan and search strategy to obtain the required primary studies to be included in a multivariate meta-analysis study. A systematic review is needed for other types of meta-analysis studies. For this paper, only primary studies that measure the relationships between CMSE and the three variables EE, DP, and PA are included. Sixteen papers are included in our MMA study. The dataset is borrowed from the work of Aloe et al. (2014). The concrete result of a systematic review is PRISMA.

PRISMA, in principle, reports the inclusion and exclusion of the primary studies under systematic review. PRISMA is a summary of steps for the inclusion and exclusion of primary studies to be included/excluded in a single MMA study (Page et al., 2021). However, the work including the dataset of Aloe et al. (2014) that we use as a main example in our study only discusses the inclusion criteria in their work. In addition, they have not provided PRISMA in their study. Thus, we will use the generic PRISMA below to explain how it applies to MMA. There are six lines of rectangular boxes in Figure 2. First, records (MMA studies) can come from two sources: databases and other sources. Second, records have to be removed if there are any duplicates. In other words, we do not include the same studies in an MMA study. Third, records are screened in terms of their relevancy. The removal of the records needs to be explained on the right side. Fourth, full articles (studies) need to be included. Explanations are needed as to why these studies are included. Those excluded need to be explained as to why they are excluded. Fifth, qualitative synthesis studies do not apply to our work, because MMA is a quantitative synthesis. Finally, we end up with a total number of studies that will be included in our MMA study. In our example, we include 16 records/studies/articles in our MMA study (see Figure 4). Again, we use the work of Aloe et al. (2014) as our main example in this paper.

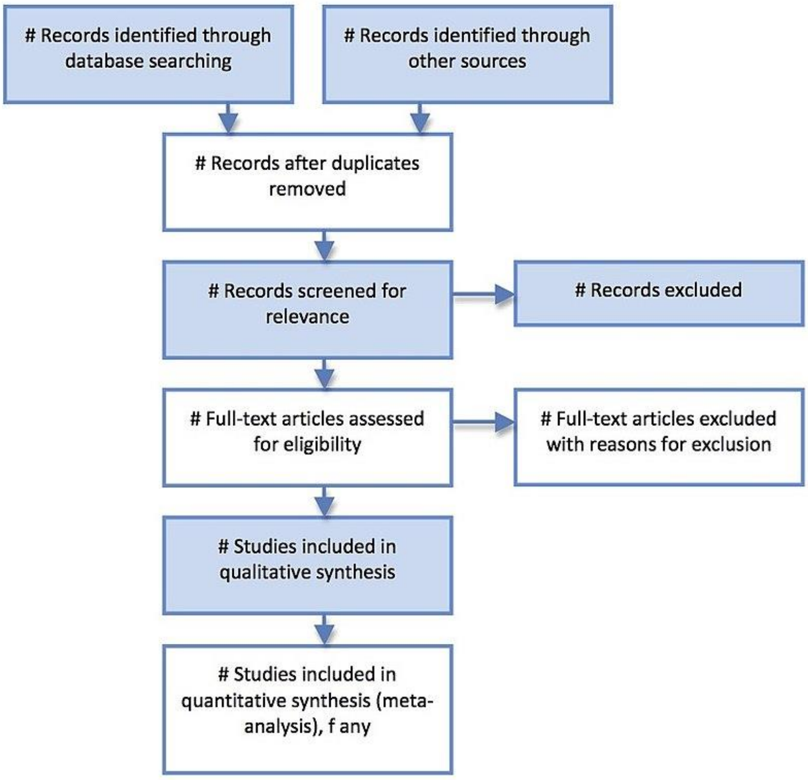

Figure 2. Generic PRISMA (Atizinha, 2012).

Conceptual Framework

Figure 3 presents the building blocks of a conceptual framework of our MMA example study. It is not common to include a conceptual framework in an MMA study. However, we feel that a conceptual framework needs to be included in this paper for clarity purposes. Based on Figure 3, the conceptual framework of the paper comprises two major columns. The first column is intervention. This can be called an independent variable (IV). CMSE (Classroom management self-efficacy) is an intervention used in the sixteen primary studies included in our example study. The second column comprises outcomes (pooled effect sizes). This can be called dependent variables. Burnout is the main variable (outcome measure) in the primary studies. However, burnout has three dimensions: emotional exhaustion (EE), depersonalization (DP), and (lowered) personal accomplishment (PA). For elaborated definitions of classroom self-efficacy, burnout, EE, DP, and PA, please consult the work of Aloe et al. (2014). In summary, the (pooled) effect sizes are the three Pearson Product Moment correlations: r1 (CMSE-EE), r2 (CMSE-DP), and r3 (CMSE-PA). It is important to point out that the correlations of the three pairs of variables are negative (Brown, 2012). Please be reminded that PA is a (lowered) personal accomplishment.

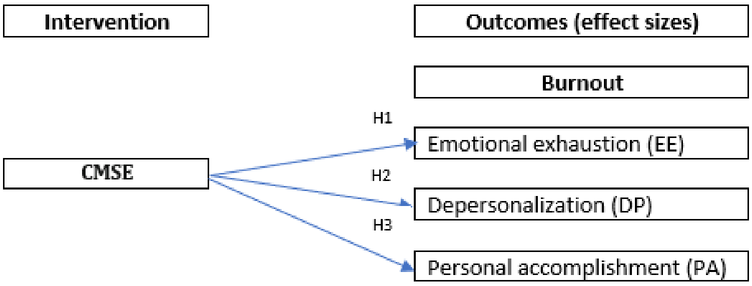

Figure 3. MMA Conceptual Framework

Hypothesis Development

Based on Figure 3, three MMA hypotheses are proposed.

H1: The pooled effect size (correlation between CMSE and EE) is significant.

H2: The pooled effect size (correlation between CMSE and DP) is significant.

H3: The pooled effect size (correlation between CMSE and PA) is significant.

Methodology

The Dataset

The dataset comprises 16 studies (see Figure 4). The first column is Study which includes the names of the primary studies. The second column is Year which includes the year of each study. Columns (EE, DP, PA) include the effect sizes (correlations). These correlations are the correlations between CMSE and EE, CMSE and DP, and CMSE and PA. Columns (V_EE, V_DP, V_PA) include sample variances of EE, DP, and PA. Columns (C_EE_DP, C_EE_PA, C_PD_PA) include covariances of the 3 pairs of variables. All these can be found in Figure 4. The dataset will be obtained identically to Figure 4 after running the R codes. The R file is located at https://osf.io/puzdm/?view_only=c53b34a6ba4c4770b6a69cc330f25f22. The cloud host is Open Science Framework (OSF). The R file includes all the required codes to run all the analyses presented in this method paper. Again, we used the dataset of Aloe et al. (2014) which is publicly available.

Aloe14 ×

Filter

	Study	Year	EE	DP	PA	V_EE	V_DP	V_PA	C_EE_DP	C_EE_PA	C_DP_PA
1	Betoret	2009	-0.38	-0.32	0.62	0.0016	0.0018	0.0011	0.0005	-0.0002	-0.0003
2	Brouwers & Tomic	2000	-0.40	-0.39	0.56	0.0013	0.0009	0.0008	0.0006	-0.0004	-0.0004
3	Bumen	2010	-0.31	-0.34	0.48	0.0014	0.0014	0.0012	0.0007	-0.0003	-0.0004
4	Chang	2009	-0.32	-0.41	0.41	0.0021	0.0019	0.0019	0.0009	-0.0010	-0.0011
5	Durr	2008	-0.47	-0.54	0.71	0.0061	0.0063	0.0041	0.0032	-0.0010	-0.0012
6	Evers et al.	2002	-0.26	-0.31	0.39	0.0093	0.0067	0.0066	0.0028	-0.0015	-0.0045
7	Evers et al.	2004	-0.30	-0.33	0.56	0.0020	0.0013	0.0010	0.0007	-0.0004	-0.0006
8	Friedman	2003	-0.15	-0.33	0.23	0.0045	0.0045	0.0048	0.0011	-0.0018	-0.0014
9	Gold	1985	-0.08	-0.10	0.12	0.0019	0.0015	0.0016	0.0010	-0.0005	-0.0007
10	Huk	2011	-0.08	-0.11	0.06	0.0136	0.0118	0.0107	0.0062	-0.0041	-0.0051
11	Kress	2007	0.00	-0.20	0.23	0.0143	0.0102	0.0087	0.0081	-0.0020	-0.0026
12	Kumarakulasingam	2002	-0.22	-0.30	0.56	0.0058	0.0066	0.0049	0.0029	-0.0007	-0.0010
13	Martin et al.	2012	-0.24	-0.40	0.49	0.0016	0.0010	0.0007	0.0009	-0.0003	-0.0004
14	Ozdemir	2007	-0.42	-0.46	0.62	0.0020	0.0020	0.0015	0.0005	-0.0002	-0.0003
15	Skaalvik and Skaalvik	2007	-0.35	-0.34	0.37	0.0042	0.0049	0.0047	0.0014	-0.0011	-0.0013
16	Williams	2012	-0.26	-0.26	0.40	0.0048	0.0040	0.0033	0.0021	-0.0007	-0.0009

Figure 4. The Dataset Displaying in R

Data Analyses and Results

Required R Packages for Conducting MMA

First, the R base (R Project, n.d.) needs to be installed. Second, RStudio (Wickham et al., 2019), is optional, but strongly recommended. Third, the metaSEM package needs to be installed. The metaSEM is required for conducting the MMA.

Bringing the Data into R

There are several ways to bring the data into R. First, if the data is already built-in in the metaSEM package, for example, users have to just call that dataset (e.g. Aloe14) from the metaSEM package. The example built-in dataset used in this paper is called Aloe14 available in the metaSEM package. The second way is to type the data directly into R, but it must be done properly. Third, you can store your data on an Excel file and instruct R to read the data from that Excel file from where the file is located, for example, on Drive D. There are more ways to bring the data into R. Readers are encouraged to explore them.

R Codes for Conducting the MMA

Line 1 calls the metaSEM package to be in use. Line 2 displays the dataset. Lines 3-5 conduct the MMA. Line 6 is intentionally left blank (no execution). Line 7 is a comment line (no execution). Line 8 reruns the results of Line 3 to get rid of failed errors. Line 9 prints the text output of MMA. Line 10 is blank. Line 11 is a comment line. Line 12 extracts the variance component of the random effects. Line 13 is intentionally left blank. Line 14 is a comment line. Line 15 converts the results of Line 12 into a matrix. Line 16 is intentionally left blank. Line 17 is a comment line. Line 18 adds the names of dimensions. Line 19 is blank. Line 20 prints the covariance. Line 21 is blank. Line 22 is a comment line. Line 23 converts into a correlation matrix. Line 24 is blank. Line 25 is a comment line. Line 26 plots the multivariate effect sizes. The R codes in Figure 5 are available at https://osf.io/puzdm/?view_only=c53b34a6ba4c4770b6a69cc330f25f22 .

```
1 library(metaSEM)
2 Aloe14
3 meta1 <- meta(y=cbind(EE,DP,PA),
4         v=cbind(V_EE, C_EE_DP, C_EE_PA, V_DP, C_DP_PA, V_PA),
5         data=Aloe14)
6
7 ## Rerun it to remove the error code
8 meta1 <- rerun(meta1)
9 summary(meta1)
10
11 ## Extract the variance component of the random effects
12 ( coef1 <- coef(meta1, select="random") )
13
14 ## Convert it into a symmetrix matrix by row major
15 my.cov <- vec2symMat(coef1, byrow=TRUE)
16
17 ## Add the dimensions for ease of interpretation
18 dimnames(my.cov) <- list( c("EE", "DP", "PA"),
19                    c("EE", "DP", "PA") )
20 my.cov
21
22 ## Convert it into a correlation matrix
23 ( cov2cor(my.cov) )
24
25 ## Plot the multivariate effect sizes
26 plot(meta1, main="", axis.labels=c("EE", "DP", "PA"))
```

Figure 5. R Codes for Conducting the MMA

Random-effect Model and Heterogeneity

First of all, we chose the random-effect model over a fixed model for several reasons. First, we assume that the studies' pooled effect sizes are inherently not the same (not leaning towards fixed). Second, the number of (sixteen) studies can be viewed as limited. Third, when heterogeneity is high, a random-effects model is generally preferred over a fixed-effects model in meta-analysis. This acknowledges the between-study variability and provides a more conservative estimate of the overall effect. Based on the random-effect model, the three effect sizes (correlations between CMSE and EE, CMSE and DP, and CMSE and PA) are all significant, $p < 0.05$ (see Figure 6). Fixed- and mixed models are also available in the metaSEM package to pool the effect sizes of the individual studies. However, explanations are needed to explain why users choose the fixed-, random-, or mixed- effects model. Here, we adopted the random-effect model because we anticipated that there was a high degree of heterogeneity across the sixteen studies included in this MMA study. The inconsistency of interventions among the sixteen studies may cause the degree of heterogeneity to be quite high.

```
Call:
meta(y = cbind(EE, DP, PA), v = cbind(V_EE, C_EE_DP, C_EE_PA,
   V_DP, C_DP_PA, V_PA), data = Aloe14)

95% confidence intervals: z statistic approximation (robust=FALSE)
Coefficients:
            Estimate   Std.Error     lbound      ubound
Intercept1 -0.27787801  0.02933890 -0.33538120 -0.22037482
Intercept2 -0.32888552  0.02765563 -0.38308956 -0.27468147
Intercept3  0.43364914  0.04355500  0.34828291  0.51901537
Tau2_1_1    0.01041632  0.00500650  0.00060375  0.02022888
Tau2_2_1    0.00856245  0.00423539  0.00026124  0.01686365
Tau2_2_2    0.00907552  0.00434070  0.00056789  0.01758314
Tau2_3_1   -0.01635020  0.00695076 -0.02997344 -0.00272697
Tau2_3_2   -0.01379872  0.00629766 -0.02614190 -0.00145554
Tau2_3_3    0.02702532  0.01101256  0.00544110  0.04860953
           z value Pr(>|z|)
Intercept1  -9.4713  < 2e-16 ***
Intercept2 -11.8922  < 2e-16 ***
Intercept3   9.9564  < 2e-16 ***
Tau2_1_1     2.0806  0.03747 *
Tau2_2_1     2.0216  0.04321 *
Tau2_2_2     2.0908  0.03655 *
Tau2_3_1    -2.3523  0.01866 *
Tau2_3_2    -2.1911  0.02845 *
Tau2_3_3     2.4540  0.01413 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Q statistic on the homogeneity of effect sizes: 256.7292
Degrees of freedom of the Q statistic: 45
P value of the Q statistic: 0

Heterogeneity indices (based on the estimated Tau2):
                              Estimate
Intercept1: I2 (Q statistic)   0.7917
Intercept2: I2 (Q statistic)   0.7969
Intercept3: I2 (Q statistic)   0.9344

Number of studies (or clusters): 16
Number of observed statistics: 48
Number of estimated parameters: 9
Degrees of freedom: 39
-2 log likelihood: -98.57498
OpenMx status1: 0 ("0" or "1": The optimization is considered fine.
Other values may indicate problems.)
```

Figure 6. Text Outputs

In terms of heterogeneity, the p-value for the Q statistic is 0. This is not desirable. A non-significant p-value is desirable. High heterogeneity in a meta-analysis indicates that the studies included show a significant amount of variability in their findings. In other words, the effect sizes of the studies are spread out and do not all cluster closely together. This can happen for several reasons. It might be easier to examine heterogeneity through I^2. The heterogeneity of the three pooled effect sizes are as follows: 79.17%, 79.69%, and 93.44%. These are also not desirable. The acceptable levels are 50% or lower. Given the high heterogeneity, researchers must be cautious in interpreting the results. Finally, by carefully considering heterogeneity and taking appropriate steps to address it, researchers can ensure their meta-analysis provides a more robust and reliable picture of the available evidence. However, dealing with the heterogeneity issue is beyond the scope of this paper.

Forest plots

Forest plots can be generated using the univariate meta-analysis method for each pooled effect size, given that the data are provided. If sufficient data are provided, users can use other R packages, the meta package, for example, to perform univariate meta-analyses to obtain the forest plots of the three effect sizes. Here, forest plots are intentionally omitted because it is of great interest in UMA, but not in MMA.

Results of Hypothesis Testing

Three hypotheses were tested. They were all supported.

H1: The pooled effect size (correlation between CMSE and EE) is statistically significant, $p < 0.05$.
H2: The pooled effect size (correlation between CMSE and DP) is statistically significant, $p < 0.05$.
H3: The pooled effect size (correlation between CMSE and PA) is statistically significant, $p < 0.05$.

Plots of Correlations Among the Pooled Effect Sizes

Based on Figure 7, the correlations among the pooled effect sizes (DP and EE, PA and EE, and PA and DP) are highly correlated. This is in line with the assumption that the pooled effect sizes of an MMA are assumed to be correlated because these pooled effect sizes are the results of the same intervention (CMSE). The results below are desirable.

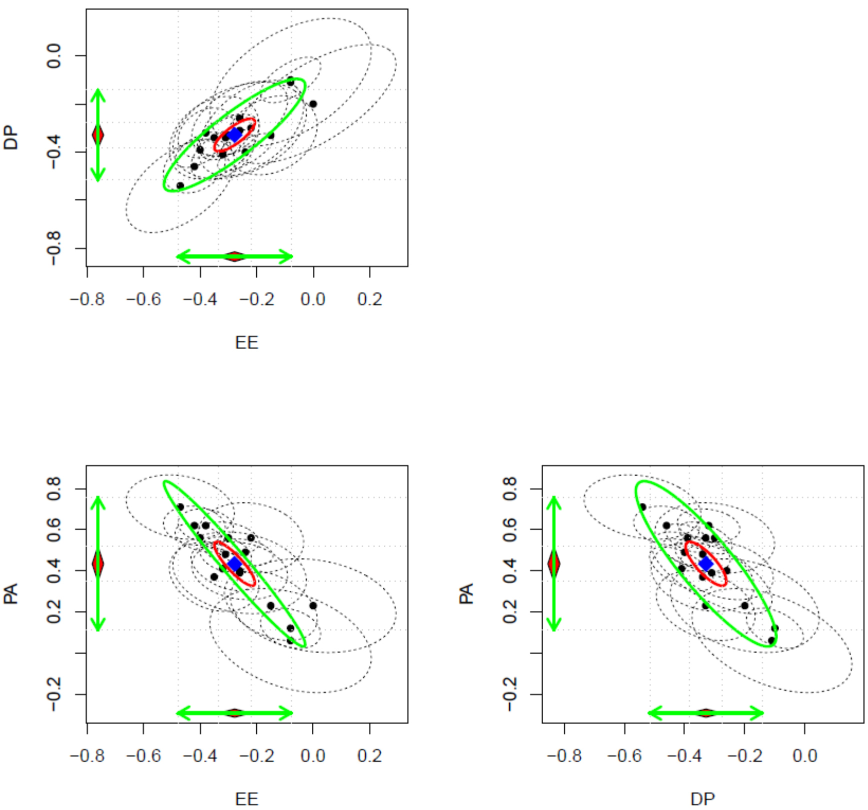

Figure 7. Correlations Among Dependent Variables

Funnel Plots and Publication Bias

Funnel plots may be produced for MMA. To do so, the UMA approach may be adopted. Again, the funnel plots are of great interest in UMA, but not MMA. Thus, the funnel plots were intentionally omitted. For publication bias, the original work of Aloe et al. (2014) conducts the publication bias test using Egger's regression test for funnel plot asymmetry for each set of effect sizes. The results indicated that there were no statistically significant asymmetries for classroom management self-efficacy with emotional exhaustion ($p = 0.06$) and classroom management self-efficacy with depersonalization ($p = 0.29$). This is desirable. However, it was statistically significant for the relationships between classroom management self-efficacy with (lowered) personal accomplishment ($p = 0.02$). This is not desirable. This raises some concern about the possibility of publication bias for this set of effect sizes. Thus, the results for this set of effects should be interpreted more cautiously. However, the publication bias test carried out by Aloe et al. (2014) is for univariate meta-analysis (not multivariate meta-analysis). Unfortunately, the metaSEM package does not offer an option to test publication bias for multivariate meta-analysis yet. However, other R packages can perform Egger's regression test, for example, the dmetar package (Harrer et al., n.d.). Thus, users, if so wish, can conduct Egger's regression test to obtain publication bias results, given there is enough data in the dataset to perform the analysis. Again, publication bias can be computed at the UMA level. At the MMA level, publication bias computation is not available as an option yet.

In addition to what we have suggested above in terms of testing for publication biases, other methods can be taken into consideration. First, visual inspection of funnel plots, multivariate extensions of funnel plot asymmetry tests can include statistical tests like Egger's test and Begg's test have been adapted for multivariate scenarios. These tests assess whether the observed distribution of effect sizes deviates from what would be expected under no publication bias. Another option is multivariate rank accumulation tests. These tests compare the ranks of effect sizes across multiple outcomes with what would be expected under no publication bias. They can be a powerful tool for detecting bias, especially when combined with other methods. Finally, meta-regression analyses are another option, including study characteristics (e.g., sample size, funding source) as predictors in a regression model along with the effect sizes can help to identify potential sources of heterogeneity and publication bias.

Discussion

The objective of this study is to illustrate how to use R to conduct MMA on psychological data. We use the dataset from the work of Aloe et al. (2014). To do this, we developed a conceptual framework to help our audience see clearly the relationships of variables being tested. They hypothesized that classroom management self-efficacy has a negative effect on burnout (see Figure 3). Self-efficacy is a protective factor against burnout (Aloe et al., 2013). The term self-efficacy is defined as a person's belief in their own ability to succeed at a particular task or goal (Dictionary.com, n.d.). Thus, if the self-efficacy of teachers increases, their feelings of burnout will decrease. Burnout was measured by three dimensions: EE, DP, and PA. This translated into three MMA hypotheses (see Figure 3). This is consistent with the work of Savas et al. (2014). According to Savas et al., they also found that self-efficacy predicted burnout negatively. In other words, the relationship between self-efficacy and burnout is negatively significant. Savas et al. measured burnout through the same three dimensions: EE, DP, and PA. We tested the three hypotheses. The three hypotheses were significant (see Figure 6). According to Aloe et al., results from sixteen studies indicate that there is a significant relationship between classroom management self-efficacy and the three dimensions of burnout, suggesting that teachers with higher levels of CMSE are less likely to experience feelings of burnout. In terms of theoretical implication, Aloe et al. contribute significantly to the body of literature on classroom management self-efficacy and burnout. In terms of practical implications, the findings of Aloe et al. have major policy implications. For example, the findings have implications for school teaching policy on countering teachers' burnout.

Conclusion

We successfully used R to conduct MMA on psychology data. The dataset belongs to Aloe et al. (2014). We developed a conceptual framework based on the main variables included in the work of Aloe et al. (2014). As demonstrated throughout the paper, R is capable of performing the MMA at the basic level. The metaSEM is the package used in this method paper. The metaSEM package has all the capabilities to conduct the MMA at the basic level. We chose the metaSEM package because it offers many learning resources on conducting MMA. However, the graphical capability of the metaSEM package is still limited. Thus, we also used the metafor package to generate the required graphical outputs (see Figure 6). In addition, the package does not offer Egger's regression test to probe for publication bias. But, Egger's test can be performed by other R packages, for example, the dmetar package. Again, publication bias analysis is common at UMA, but not at MMA.

Recommendations

In terms of practice recommendations, based on our findings, MMA can help us see clearly whether the social treatments (i.e. classroom management) significantly affect the outcome variables (EE, DP, and PA) in significant ways. For practical implication, practitioners (e.g. school administrators and teachers) can use these findings to frame their school teaching policy concerning classroom management. Other fields, such as management where different treatments (e.g. managerial practices) are implemented and affecting multiple (organization) outcomes, can also benefit from MMA research. Based on our findings, MMA is suitable for meta-analysis research in the field of psychology. However, researchers in other fields can also use MMA to test a treatment (e.g. innovative management practices) that affects multiple (organizational) outcomes. In other words, MMA does not have to be applied in psychological research. Given R capabilities, we encourage researchers to explore the capabilities of the metaSEM package and other R packages that can perform MMA. So, they can have all the capabilities of R at their disposal. Ultimately, we encourage academic researchers to use R for MMA for their teaching, research, and publication. Finally, future method papers can focus on the remaining categories of meta-analysis presented in Figure 1. Future research can use educational psychology data or data from other wide variety of fields (e.g. business, economics, and health sciences).

Limitations

The dataset is appropriate to the defined scope of the paper (e.g. testing the three pooled effect sizes of the example study). These three pooled effect sizes were the correlations between CMSE and EE, CMSE and DP, and CMSE and PA. However, we did not produce both forest and funnel plots. Again, generating the forest and funnel plots has to take the UMA approach, which is not within the scope of this paper.

Ethics Statement

This research project received approval from the ethics committee of Mahasarakham University (approval number 537-592/2023).

Conflict of Interest

The authors declare no conflicts of interest.

Acknowledgment

This research project was financially supported by Mahasarakham University.

Authorship Contribution Statement

Savatsomboon: Conceptualization, design, analysis, writing. Ruannakarn: Editing/reviewing, supervision, critical revision of manuscript. Yurayat: Conceptualization, design, supervision. Chanprasitchai: Conceptualization, design, supervision. Leihaothabam: Conceptualization, design, final approval.

References

Ahn, E., & Kang, H. (2018). Introduction to systematic review and meta-analysis. *Korean Journal of Anesthesiology*, *71*(2), 103-112. https://doi.org/10.4097/kjae.2018.71.2.103

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*, *26*, 101-126. https://doi.org/10.1007/s10648-013-9244-0

Atizinha. (2012, March 25). *File:PRISMA flow diagram.jpg*. Wikipedia. https://en.m.wikipedia.org/wiki/File:PRISMA_flow_diagram.jpg

Berkey, C. S., Hoaglin, D. C., Antczak-Bouckoms, A., Mosteller, F., & Colditz, G. A. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine*, *17*(22), 2537-2550. https://bit.ly/4ax0tRC

Brown, C. G. (2012). A systematic review of the relationship between self-efficacy and burnout in teachers. *Educational and Child Psychology*, *29*(4), 47-63. https://doi.org/10.53841/bpsecp.2012.29.4.47

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

Dictionary.com. (n.d.). Self-efficacy In *Dictionary.com dictionary*. https://bit.ly/3yxOBl8

Gasparrini, A., Armstrong, B., & Kenward, M. G. (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*, *31*(29), 3821-3839. https://doi.org/10.1002/sim.5471

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. D. (n.d.). *dmetar: Doing meta-analysis in R.* dmetar. http://dmetar.protectlab.org/

Hattle, M., Burke, D. L., Trikalinos, T., Schmid, C. H., Chen, Y., Jackson, D., & Riley, R. D. (2022). Multivariate meta-analysis of multiple outcomes: Characteristics and predictors of borrowing of strength from Cochrane reviews. *Systematic Reviews, 11*, Article 149. https://doi.org/10.1186/s13643-022-01999-0

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*. Article 89. https://doi.org/10.1186/s13643-021-01626-4

R Project. (n.d.). The R project for statistical computing. https://www.R-project.org/

Savas, A. C., Bozgeyik, Y., & Eser, I. (2014). A study on the relationship between teacher self efficacy and burnout. *European Journal of Educational Research*, *3*(4), 159-166. https://doi.org/10.12973/eu-jer.3.4.159

Savatsomboon, G., Yurayat, P., Chanprasitchai, O., Narkbunnum, W., Sharma, J. K., & Svetsomboon, S. (2024). A proposed categorization of meta-analysis, their respective example conceptual frameworks, and applicable R packages for education research: A review. *Journal of Practical Studies in Education*, *5*(3), 1-7. https://doi.org/10.46809/jpse.v5i3.83

Schiavenato, M., & Chu, F. (2021). PICO: What it is and what it is not. *Nurse Education in Practice*, 56, 103194. https://doi.org/10.1016/j.nepr.2021.103194

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686