Correcting for publication bias in multivariate and multilevel meta-analysis: A multivariate step function selection model approach

Robbie C.M. van Aert

June 11, 2026

Tilburg University

Publication bias and dependent effect sizes

- Publication bias is the selective publication of studies and usually favoring statistically significant outcomes
- Publication bias methods have primarily been developed for meta-analyses including independent effect sizes

Publication bias and dependent effect sizes

- Publication bias is the selective publication of studies and usually favoring statistically significant outcomes

- Publication bias methods have primarily been developed for meta-analyses including independent effect sizes

- Dependent effect sizes are common in psychology, education, and medicine [1–4] $\rightarrow$ 4.5 effect sizes per study

- Two types of dependencies:
  - Hierarchical effects $\rightarrow$ multilevel meta-analysis model
  - Correlated effects $\rightarrow$ multivariate meta-analysis model

Publication bias and dependent effect sizes

- Publication bias is the selective publication of studies and usually favoring statistically significant outcomes

- Publication bias methods have primarily been developed for meta-analyses including independent effect sizes

- Dependent effect sizes are common in psychology, education, and medicine [1–4] $\rightarrow$ 4.5 effect sizes per study

- Two types of dependencies:
  - Hierarchical effects $\rightarrow$ multilevel meta-analysis model
  - Correlated effects $\rightarrow$ multivariate meta-analysis model

- **Goal:** Introduce a publication bias method for multilevel/ multivariate meta-analysis model

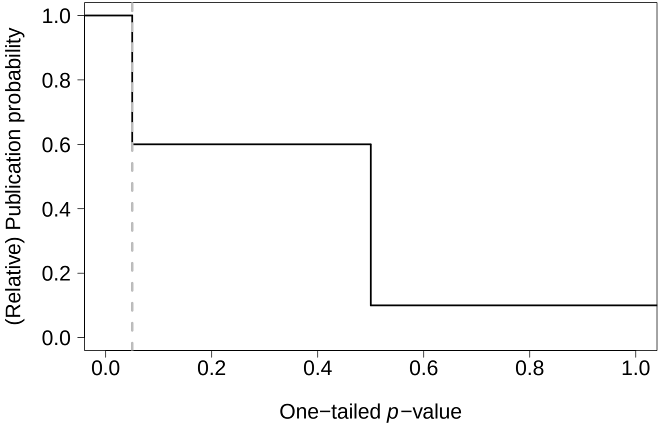
Univariate selection model approach

- Generic term for methods combining effect size model with selection model
- *Effect size model:* distribution of effect sizes in the absence of publication bias
- *Selection model:* mechanism by which effect size estimates are selected to be observed

Univariate selection model approach

- Generic term for methods combining effect size model with selection model
- *Effect size model:* distribution of effect sizes in the absence of publication bias
- *Selection model:* mechanism by which effect size estimates are selected to be observed
- We focus on the step function selection model [5]
- "Steps" of *p*-values need to be set for studies that are assumed to have the same publication probability

Univariate step function selection model



- Steps (0.05, 0.5) $\rightarrow$ selected by the user
- Publication probabilities (1, 0.6, 0.1) $\rightarrow$ estimated or selected by the user

Multivariate step function selection model

- *Effect size model:* Multivariate or multilevel meta-analysis model

- *Selection model* focuses on missing an entire study due to publication bias and **not** on outcome reporting bias

- Selection models using one step at significance threshold:
  - Strict $\rightarrow$ studies with *only* significant outcomes are more likely to be published
  - Relaxed $\rightarrow$ studies with *at least one* significant outcome are more likely to be published
  - ...

Multivariate step function selection model

- Imagine a study with two outcomes:

Outcome 1	Outcome 2	Strict	Relaxed
Significant	Significant	ω_1	ω_1
Significant	Nonsignificant	ω_2	ω_1
Nonsignificant	Significant	ω_2	ω_1
Nonsignificant	Nonsignificant	ω_2	ω_2

- ω_1 and ω_2 are the publication probabilities

Multivariate step function selection model

- Imagine a study with two outcomes:

Outcome 1	Outcome 2	Strict	Relaxed
Significant	Significant	ω_1	ω_1
Significant	Nonsignificant	ω_2	ω_1
Nonsignificant	Significant	ω_2	ω_1
Nonsignificant	Nonsignificant	ω_2	ω_2

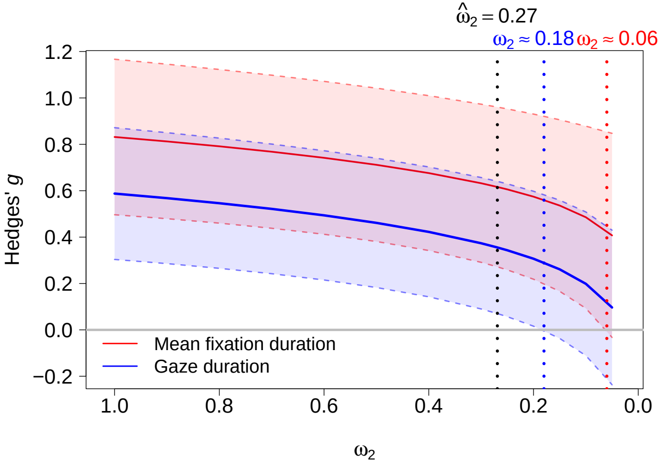
- ω_1 and ω_2 are the publication probabilities

- Implementation:
  - Maximum likelihood estimation and large-sample standard errors for fixed effects
  - Estimating and assuming different publication probabilities in a sensitivity analysis

Example multivariate meta-analysis: Aging and reading

- Moreno et al. (2019) [6] conducted a multivariate meta-analysis with mean fixation and gaze duration as outcomes
- 14 studies were included that all reported both outcomes
- Hedges' $g \rightarrow$ larger g, longer fixation time for older adults
- Both outcomes were significant in 50% of the studies
- The strict selection model was used
- $\omega_1 = 1$ to identify the model

Example multivariate meta-analysis: Aging and reading



- **Conclusion:**
  - Severe publication bias was needed to change the conclusions
  - $\hat{\omega}_2 = 0.27$, 95% CI (0.041;1.618) $\rightarrow$ wide CI!

Discussion

- Results of a simulation study
  - Small bias of the proposed method when the true publication probability was used
  - Less or comparable bias than uncorrected model was observed when misspecifying the publication probability
  - Small bias when publication probability was estimated $\rightarrow$ standard errors about 1.4 times larger

Discussion

- Results of a simulation study
  - Small bias of the proposed method when the true publication probability was used
  - Less or comparable bias than uncorrected model was observed when misspecifying the publication probability
  - Small bias when publication probability was estimated $\rightarrow$ standard errors about 1.4 times larger

- Corrected estimates of the variance of the random effects and corrected hypothesis tests can also be obtained

Discussion

- Results of a simulation study
  - Small bias of the proposed method when the true publication probability was used
  - Less or comparable bias than uncorrected model was observed when misspecifying the publication probability
  - Small bias when publication probability was estimated $\rightarrow$ standard errors about 1.4 times larger

- Corrected estimates of the variance of the random effects and corrected hypothesis tests can also be obtained

- Future research
  - Direct comparison to univariate publication bias methods
  - Study the flexibility of defining the selection model

Thank you for your attention

Slides are available on my website:

www.robbievanaert.com

Our research group at Tilburg University:

www.metaresearch.nl

Preprint:

van Aert, R. C. M., Riley, R. D., & Jackson, D. (in press). Correcting for publication bias in multivariate and multilevel meta-analysis: A multivariate step function selection model approach. *Psychological Methods.* doi: 10.31234/osf.io/rk4mn_v1

References I

1. Ahn, S., Ames, A. J., & Myers, N. D. (2012). A review of meta-analyses in education: Methodological strengths and weaknesses. *Review of Educational Research*, *82*(4), 436–476. https://doi.org/10.3102/0034654312458162

2. Park, S., & Beretvas, S. N. (2019). Synthesizing effects for multiple outcomes per study using robust variance estimation versus the three-level model. *Behavior Research Methods*, *51*(1), 152–171. https://doi.org/10.3758/s13428-018-1156-y

3. Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019). Current practices in meta-regression in psychology, education, and medicine. *Research Synthesis Methods*, *10*(2), 180–194. https://doi.org/10.1002/jrsm.1339

4. Wu, W., Duan, J., Reed, W. R., & Tipton, E. (2026). What can we learn from 1,000 meta-analyses across 10 different disciplines? *Research Synthesis Methods*, *17*(1), 123–156. https://doi.org/10.1017/rsm.2025.10035

5. Hedges, L. V. (1992). Meta-Analysis. *Journal of Educational Statistics*, *17*(4), 279–296.

6. Moreno, J. D., León, J. A., Arnal, L. A. M., & Botella, J. (2019). Age differences in eye movements during reading: Degenerative problems or compensatory strategy?: A meta-analysis. *European Psychologist*, *24*(4), 297–311. https://doi.org/10.1027/1016-9040/a000344

7. Picho-Kiroga, K., Turnbull, A., & Rodriguez-Leahy, A. (2021). Stereotype threat and its problems: Theory misspecification in research, consequences, and remedies. *Journal of Advanced Academics*, *32*(2), 231–264. https://doi.org/10.1177/1932202X20986161

Marginal model of multilevel meta-analysis model

$$\boldsymbol{Y}_i \sim N(\mu\mathbf{1}, \boldsymbol{S}_i^{ML} + \boldsymbol{\Sigma}^{ML})$$

In case a study contributes two effect sizes:

$$\boldsymbol{S}_i^{ML} = \begin{pmatrix} \sigma_{i1}^2 & 0 \\ 0 & \sigma_{i2}^2 \end{pmatrix}$$

$$\boldsymbol{\Sigma}^{ML} = \begin{pmatrix} \sigma_B^2 + \sigma_W^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_B^2 + \sigma_W^2 \end{pmatrix}$$

- $\boldsymbol{Y}_i$: vector containing the d effect sizes of the i^{th} study
- $\mathbf{1}$: column vector of length k where all entries are 1
- $\boldsymbol{S}_i^{ML}$: within-study covariance matrix
- $\boldsymbol{\Sigma}^{ML}$: between-study covariance matrix

Marginal model of multivariate meta-analysis model

$$\boldsymbol{Y}_i \sim N(\boldsymbol{\mu}, \boldsymbol{S}_i^{MV} + \boldsymbol{\Sigma}^{MV})$$

In case of a bivariate meta-analysis:

$$\boldsymbol{S}_i^{MV} = \begin{pmatrix} \sigma_{i1}^2 & r_i \sigma_{i1} \sigma_{i2} \\ r_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{pmatrix}$$

$$\boldsymbol{\Sigma}^{MV} = \begin{pmatrix} \tau_1^2 & \rho \tau_1 \tau_2 \\ \rho \tau_1 \tau_2 & \tau_2^2 \end{pmatrix}$$

- $\boldsymbol{Y}_i$: vector containing the d effect sizes of the i^{th} study
- $\boldsymbol{\mu}$: the vector containing the pooled true effect sizes
- $\boldsymbol{S}_i^{MV}$: within-study covariance matrix
- $\boldsymbol{\Sigma}^{MV}$: between-study covariance matrix

PDF of univariate step function selection model

$$f(Y_i;\mu,\boldsymbol{\omega},s_i,\sigma_i) = \frac{s_i^{-1} w(Y_i,\sigma_i)\phi\left(\frac{Y_i-\mu}{s_i}\right)}{\int s_i^{-1} w(Y_i,\sigma_i)\phi\left(\frac{Y_i-\mu}{s_i}\right) dY_i} \tag{1}$$

$$w(Y_i,\sigma_i) = \begin{cases} \omega_1 & \text{if } -\sigma_i\Phi^{-1}(a_1) < Y_i < \infty; \\ \omega_m & \text{if } -\sigma_i\Phi^{-1}(a_m) < Y_i \leq -\sigma_i\Phi^{-1}(a_{m-1}); \\ \omega_M & \text{if } -\infty < Y_i \leq -\sigma_i\Phi^{-1}(a_{M-1}) \end{cases} \tag{2}$$

- Y_i: observed effect size estimate
- σ_i: square root of the within-study sampling variance
- s_i: square root of the total variance
- ϕ: PDF of standard normal distribution
- m: index for the intervals
- M: total number of intervals
- $a_1, \cdots, a_M$: cut points on the p-value scale

PDF of multivariate step function selection model

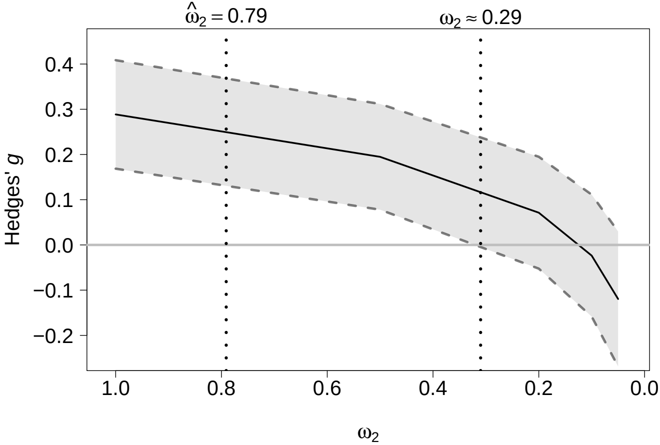
$$f(\boldsymbol{Y}_i, \boldsymbol{\theta}, \Sigma_i, \boldsymbol{\omega}, \boldsymbol{\sigma}_i) = \frac{w(\boldsymbol{Y}_i, \boldsymbol{\sigma}_i)\phi_{d_i}(\boldsymbol{Y}_i, \boldsymbol{\theta}, \Sigma_i)}{\int w(\boldsymbol{Y}_i, \boldsymbol{\sigma}_i)\phi_{d_i}(\boldsymbol{Y}_i; \boldsymbol{\theta}, \Sigma_i) d^{d_i} \boldsymbol{Y}_i} \tag{3}$$

- $\boldsymbol{Y}_i$: the vector containing the d outcomes from the i^{th} study
- $\boldsymbol{\sigma}_i = (\sigma_{i1}, \cdots, \sigma_{id})$
- $\boldsymbol{\theta}$: vector with average true effect sizes
- $\boldsymbol{\Sigma}_i$: total covariance matrix
- ϕ_{d_i}: d_i dimensional multivariate normal density

Example multilevel meta-analysis: Stereotype threat

- Stereotype threat refers to girls' underperformance on math tests when primed with a negative stereotype
- Picho-Kiroga et al. (2021) [7] conducted a three-level meta-analysis $\rightarrow$ 100 effect sizes nested in 52 studies
- Hedges' $g \rightarrow$ larger g, more evidence for stereotype threat
- 50% of the studies contained at least one significant effect
- The relaxed selection model was used
- $\omega_1 = 1$ to identify the model

Example multilevel meta-analysis: Stereotype threat



- **Conclusion:**
  - Results are not robust if publication probability of studies without significant effect sizes is $\omega_2 \approx 0.29$ or lower
  - $\hat{\omega}_2 = 0.79$, 95% CI (0.349;1.782) $\rightarrow$ wide CI!