COMBINING ROBUST VARIANCE ESTIMATION WITH MODELS FOR DEPENDENT EFFECT SIZES

James E. Pustejovsky, UT Austin
joint work with: Beth Tipton, Northwestern University
Ariel Aloe, University of Iowa

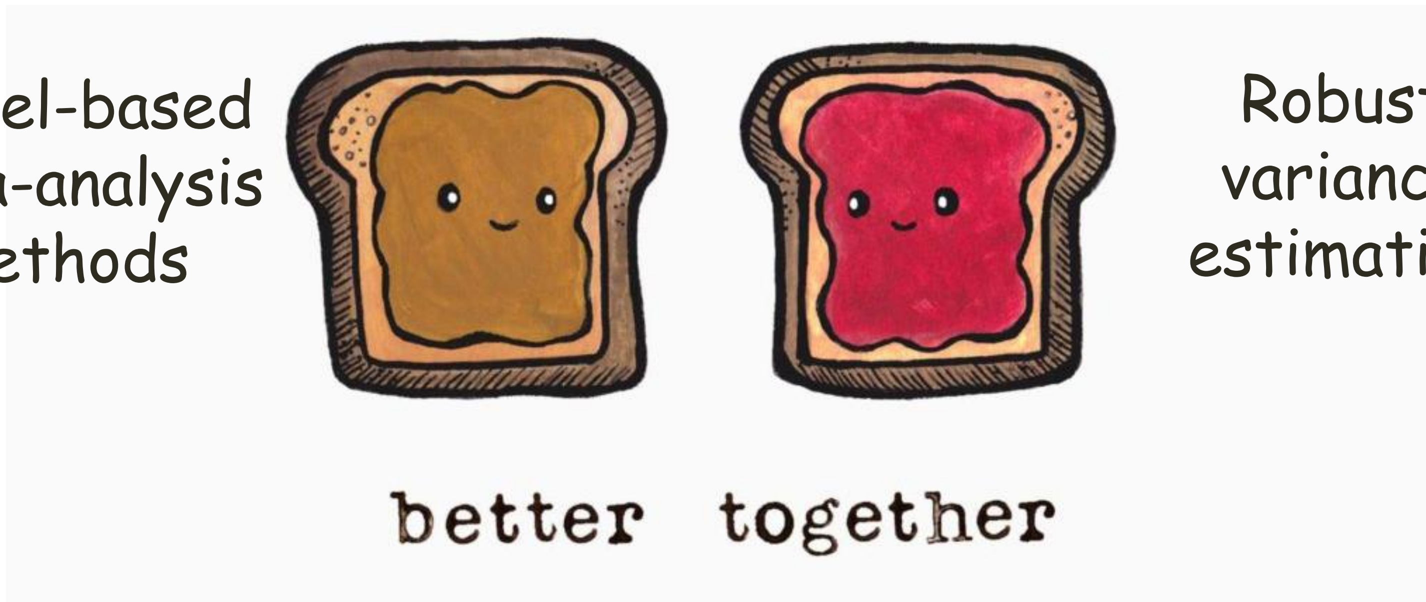
UT Austin
October 1, 2018

pusto@austin.utexas.edu

THESIS

- Meta-analyses often involve statistically dependent effect sizes.
- Many methods available for meta-analyzing dependent effect size estimates.
  - ad hoc methods (Hammering the screws)
  - model-based methods
  - robust variance estimation (RVE)
- Useful to combine RVE with model-based approaches.
  - Addresses limitations of model-based approaches.
  - Addresses limitations of default RVE implementation.

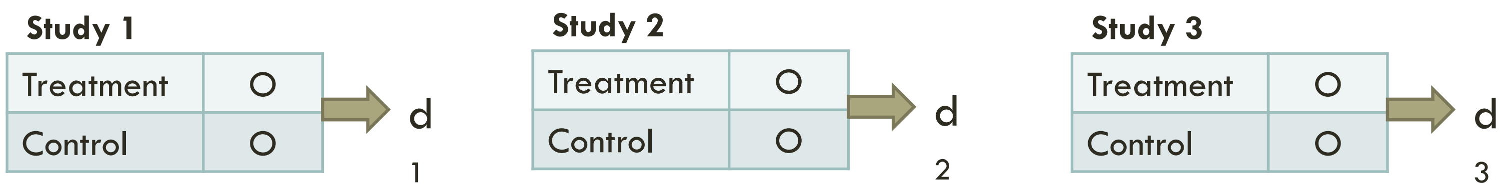
Model-based meta-analysis methods



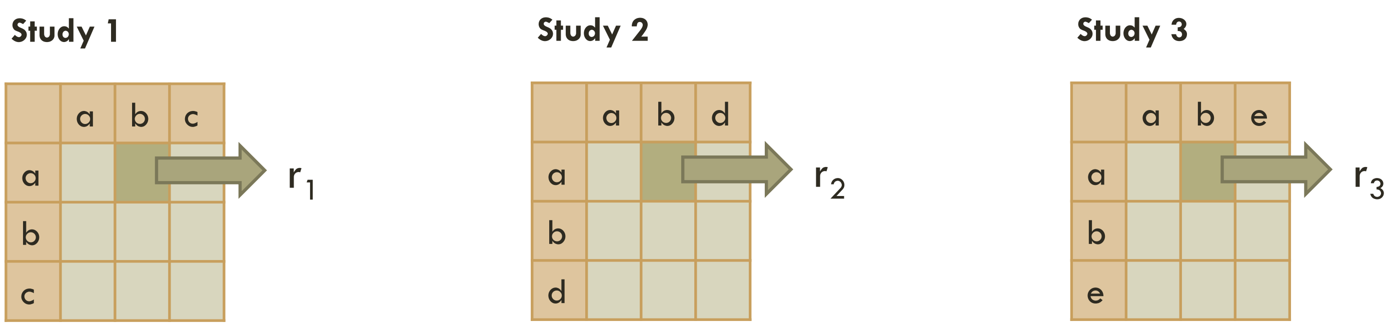
Robust variance estimation

BASIC META-ANALYSIS METHODS ASSUME INDEPENDENT EFFECT SIZES

In a meta-analysis of experiments:



In a meta-analysis of correlations:



DEPENDENT EFFECT SIZES ARE VERY COMMON

Multiple outcomes measured on common set of participants

Treatment	O P Q
Control	O P Q

d_{O1}
d_{P1}
d_{Q1}

Outcome measured at multiple follow-up times

Treatment	O_1	O_2	O_3
Control	O_1	O_2	O_3

d_{12}
d_{22}
d_{32}

Multiple treatment conditions compared to a common control

Treatment T	O
Treatment U	O
Control	O

d_{T3}
d_{U3}

Multiple correlations from a common sample

	a	b	c	d
a				
b				
c				
d				

r_{ab}
r_{ac}
r_{ad}

FRIESE, FRANKENBACH, JOB, & LOSCHELDER (2017). DOES SELF-CONTROL TRAINING IMPROVE SELF-CONTROL: A META-ANALYSIS.

33 experimental studies, 166 effect size estimates (standardized mean differences)

- ✓ 1-52 effect size estimates per study (median = 2)
- ✓ Multiple outcomes (1-13 outcomes per study, median = 2)
- ✓ Multiple follow-up times (immediate post-test and/or later follow-up)
- ✓ Multiple treatment conditions (1-4 treatment conditions per study)
- ✓ Multiple control conditions (active and/or passive control)

LEHTONEN ET AL. (2018). IS BILINGUALISM ASSOCIATED WITH ENHANCED EXECUTIVE FUNCTIONING IN ADULTS?

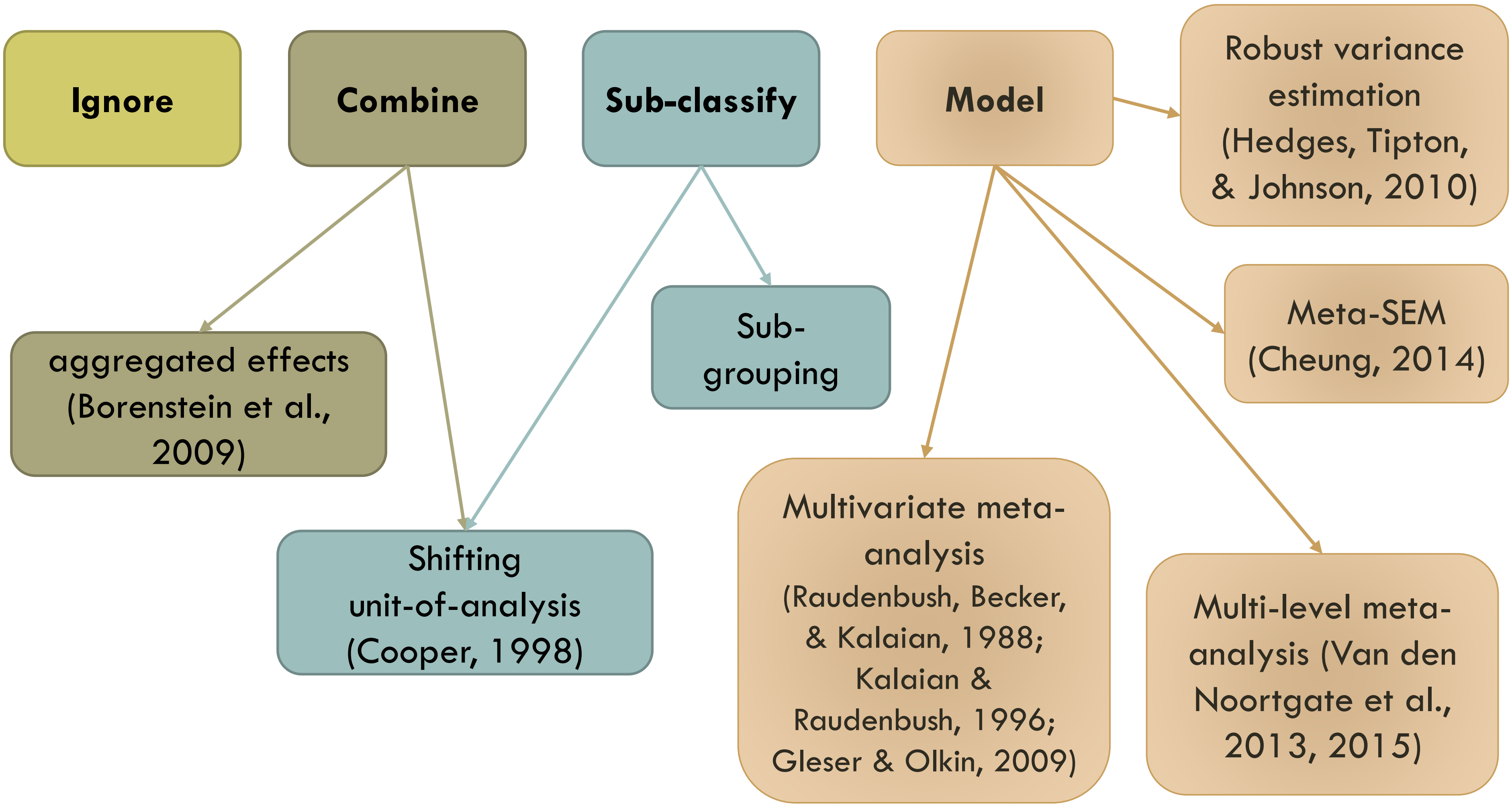
152 studies, 891 effect size estimates (standardized mean differences comparing performance of bilingual and monolingual adults)

- ✓ 1-40 effect size estimates per study (median = 4)
- ✓ Multiple outcomes (1-7 outcomes per study, median = 2)
- ✓ Multiple bilingual groups
- ✓ Multiple monolingual groups

CORRELATIONS BETWEEN ES ESTIMATES ARE OFTEN NOT AVAILABLE

- Multiple treatments compared to common control
  - known formulas (Gleser & Olkin, 2009), easy enough to calculate
- Multiple outcomes/multiple follow-ups
  - known formulas (Gleser & Olkin, 2009)
  - require knowing correlations among outcomes/repeated measures, which are not often reported
- Multiple correlations from common sample
  - known, icky formulas (Olkin & Siotani, 1976)
  - need to know correlations between ALL variables involved

BECKER (2000) DESCRIBED FOUR BROAD STRATEGIES FOR HANDLING DEPENDENCE:



AGGREGATED EFFECTS

- Average estimates to generate single "synthetic" ES per study.

Study	ES	V	Predictors
A	ES_{A1}	V_{A1}	X_{A1}
A	ES_{A2}	V_{A2}	X_{A2}
A	ES_{A3}	V_{A3}	X_{A3}
B	ES_{B1}	V_{B1}	X_{B1}
C	ES_{C1}	V_{C1}	X_{C1}
C	ES_{C2}	V_{C2}	X_{C2}

Study	ES	V	Predictors
A	$\overline{ES}_A$	$\overline{V}_A$	$\overline{X}_A$
B	$\overline{ES}_B$	$\overline{V}_B$	$\overline{X}_B$
C	$\overline{ES}_C$	$\overline{V}_C$	$\overline{X}_C$

- Estimating variance of synthetic ES requires correlations among component ES (Borenstein et al., 2009).
- Limits moderator/meta-regression analyses to between-study predictors.

SUB-GROUPS/SHIFTING UNIT-OF-ANALYSIS

- Classify ES into sub-groups where each study contributes $\leq$ 1 ES.
- If there are still multiple ES per sub-group, aggregate (Cooper, 1998).

Study	ES	V	Category
A	ES_{A1}	V_{A1}	1
A	ES_{A2}	V_{A2}	2
A	ES_{A3}	V_{A3}	2
B	ES_{B1}	V_{B1}	1
C	ES_{C1}	V_{C1}	1
C	ES_{C2}	V_{C2}	2

Study	Category 1		Category 2	
A	ES_{A1}	V_{A1}		
A			$\overline{ES}_{A2}$	$\overline{V}_{A2}$
B	ES_{B1}	V_{B1}		
C	ES_{C1}	V_{C1}		
C			ES_{C2}	V_{C2}

- Use univariate meta-analysis within sub-groups.

PROBLEMS WITH SHIFTING UNIT-OF-ANALYSIS

- Average effects by sub-group are not independent.
  - How to make comparisons between average effects by sub-group?
- Different ES estimates for each moderator analysis.
  - How to do meta-regression with multiple predictors?

Study	ES	V	Category
A	ES_{A1}	V_{A1}	1
A	ES_{A2}	V_{A2}	2
A	ES_{A3}	V_{A3}	2
B	ES_{B1}	V_{B1}	1
C	ES_{C1}	V_{C1}	1
C	ES_{C2}	V_{C2}	2

Study	Category 1		Category 2	
A	ES_{A1}	V_{A1}		
A			$\overline{ES}_{A2}$	$\overline{V}_{A2}$
B	ES_{B1}	V_{B1}		
C	ES_{C1}	V_{C1}		
C			ES_{C2}	V_{C2}

MULTIVARIATE META-ANALYSIS

(Raudenbush, Becker, & Kalaian,1988; Kalaian & Raudenbush, 1996)

- Hierarchical model for component ES estimates nested within studies

$$T_{ij} = \mathbf{x}_{ij}\boldsymbol{\beta} + u_j + v_{ij} + e_{ij}$$

where $u_j \sim N(0, \tau^2)$, $v_{ij} \sim N(0, \omega^2)$, $e_{ij} \sim N(0, s_{ij}^2)$, and
$\boxed{\mathrm{Cov}\left(e_{hj}, e_{ij}\right) = r_{hij} s_{hj} s_{ij}}$.

- Requires estimates/assumptions about ES correlations r_{hij}.
- Allows for modeling of between- and within-study variation in the ES.
- Makes use of between- and within-study variation in predictors.

MULTI-LEVEL META-ANALYSIS

(Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013, 2015)

- Use multi-level model to account for dependence between ES estimates within studies, ***ignoring the sampling correlations***:

$$T_{ij} = \mathbf{x}_{ij}\boldsymbol{\beta} + u_j + v_{ij} + e_{ij}$$

where $u_j \sim N(0, \tau^2)$, $v_{ij} \sim N(0, \omega^2)$, $e_{ij} \sim N(0, s_{ij}^2)$, $\boxed{\mathrm{Cov}\left(e_{hj}, e_{ij}\right) = 0.}$

- Simulation evidence indicates that this approach can be "robust" to mis-specified correlation structure.
- But unclear whether robustness holds generally.

RE-ANALYSIS OF BILINGUALISM STUDIES

	(1) Aggregated effects	(2) Shifting unit-of-analysis	(3) Multivariate meta-analysis	(4) Multi-level meta-analysis
Overall Average ES (152 studies, 869 ES)	0.028 [0.026]		0.047* [0.022]	0.055* [0.023]
Between-study SD	0.168		0.157	0.254
Within-study SD			0.289	0.232

BILINGUALISM EFFECTS BY DOMAIN

	(1) Aggregated effects	(2) Shifting unit-of-analysis	(3) Multivariate meta-analysis	(4) Multi-level meta-analysis
Inhibition (95 studies, 212 ES)	0.077 [0.003]	0.114** [0.037]	0.106*** [0.031]	0.115*** [0.032]
Monitoring (81 studies, 184 ES)	0.003 [0.100]	0.077 [0.039]	0.058 [0.033]	0.065 [0.034]
Shifting (37 studies, 79 ES)	0.147 [0.127]	0.147** [0.056]	0.141** [0.046]	0.148** [0.047]
Attention (18 studies, 53 ES)	0.230 [0.193]	-0.013 [0.080]	-0.031 [0.058]	-0.021 [0.058]
Working Memory (73 studies, 243 ES)	0.045 [0.059]	0.058 [0.042]	0.057 [0.032]	0.064* [0.033]
Fluency (28 studies, 98 ES)	-0.313** [0.106]	-0.260*** [0.066]	-0.211*** [0.045]	-0.196*** [0.045]
Between-study SD	0.155	0.224	0.150	0.249
Within-study SD			0.276	0.217

COMPARISON

Method	Requires making assumptions about ES covariances
Aggregated effects	✓
Sub-grouping	✓
Shifting unit-of-analysis	✓
Multivariate meta-analysis	✓
Multi-level meta-analysis	✓

ROBUST VARIANCE ESTIMATION
(Hedges, Tipton, & Johnson, 2010)

- Meta-analysis/meta-regression method using "sandwich" variance estimators (a.k.a., "clustered" SEs)
  - Robust to mis-specified assumptions about variance-covariance structure within independent studies.

- Sandwich methods work with very general classes of models, including any of the other methods for handling dependent effects.
  - Proof: See Hedges et al. (2010, Appendix A).

- Conventional sandwich estimators require large number of studies.
  - But small-sample corrections are available (Tipton, 2015; Tipton & Pustejovsky, 2015).

ROBUST VARIANCE ESTIMATION THEORY

- A generic meta-regression model (in matrix form):

$$\mathbf{T}_j = \mathbf{X}_j \boldsymbol{\beta} + \mathbf{e}_j$$

where $\mathrm{E}(\mathbf{e}_j) = \mathbf{0}$ and $\mathrm{Var}(\mathbf{e}_j) = \boldsymbol{\Omega}_{\mathbf{j}}$, for $\boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{m}$.

- Estimate $\boldsymbol{\beta}$ using weighted least squares for some weight matrices $\mathbf{W}_{\boldsymbol{j}}$:

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{j=1}^{J} \mathbf{X}_j' \mathbf{W}_j \mathbf{X}_j \right)^{-1} \sum_{j=1}^{J} \mathbf{X}_j' \mathbf{W}_j \mathbf{T}_j \,.$$

HOW TO ESTIMATE $\mathrm{VAR}(\widehat{\boldsymbol{\beta}})$?

- The true variance of $\widehat{\boldsymbol{\beta}}$:

$$Var(\widehat{\boldsymbol{\beta}}) = \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1} \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\boldsymbol{\Omega}_j\mathbf{W}_j\mathbf{X}_j\right)\left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1}$$

- ***Model-based variance estimation*** assumes a correct model for $\boldsymbol{\Omega}_j$:

$$\mathbf{V}^{model} = \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1} \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\widehat{\boldsymbol{\Omega}}_j\mathbf{W}_j\mathbf{X}_j\right)\left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1}.$$

HOW TO ESTIMATE $\mathrm{VAR}(\widehat{\boldsymbol{\beta}})$?

- The true variance of $\widehat{\boldsymbol{\beta}}$:

$$Var(\widehat{\boldsymbol{\beta}}) = \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1} \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\boldsymbol{\Omega}_j\mathbf{W}_j\mathbf{X}_j\right)\left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1}$$

- ***Robust variance estimation*** avoids relying on a model for $\boldsymbol{\Omega}_{\mathrm{j}}$ by using the regression residuals:

$$\mathbf{V}^{robust} = \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1} \left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{A}_j\hat{\mathbf{e}}_j\hat{\mathbf{e}}_j'\mathbf{A}_j\mathbf{W}_j\mathbf{X}_j\right)\left(\sum_{j=1}^{J} \mathbf{X}_j'\mathbf{W}_j\mathbf{X}_j\right)^{-1}$$

- Residuals are lousy estimates of specific $\boldsymbol{\Omega}_{\mathrm{j}}$, but they ***work well on average.***

COMPARISON

Method	Requires making assumptions about ES covariances	Robustness to assumptions about ES covariances
Aggregated effects	✓	Robust*
Sub-grouping	✓	Robust*
Shifting unit-of-analysis	✓	Robust*
Multivariate meta-analysis	✓	Robust*
Multi-level meta-analysis	✓	Robust*
Robust variance estimation	✓ (Working model)	Robust

* When combined with robust (sandwich) variance estimation

BILINGUALISM EFFECTS BY DOMAIN [ROBUST SE]

	(1) Aggregated effects	(2) Shifting unit-of-analysis	(3) Multivariate meta-analysis	(4) Multi-level meta-analysis
Inhibition (95 studies, 212 ES)	0.077 [0.080]	0.114*** [0.033]	0.106** [0.035]	0.115*** [0.036]
Monitoring (81 studies, 184 ES)	0.003 [0.111]	0.077 [0.041]	0.058 [0.036]	0.065 [0.037]
Shifting (37 studies, 79 ES)	0.147 [0.117]	0.147* [0.059]	0.141** [0.057]	0.148** [0.058]
Attention (18 studies, 53 ES)	0.230 [0.173]	-0.013 [0.076]	-0.031 [0.087]	-0.021 [0.085]
Working Memory (73 studies, 243 ES)	0.045 [0.073]	0.058 [0.043]	0.057 [0.036]	0.064* [0.037]
Fluency (28 studies, 98 ES)	-0.313** [0.127]	-0.260*** [0.071]	-0.211** [0.057]	-0.196*** [0.057]
Between-study SD	0.155	0.224	0.150	0.249
Within-study SD			0.276	0.217

DEFAULT RVE IMPLEMENTATION HAS LIMITATIONS

- Implementation in **robumeta** packages for R and Stata.
- Limited to two "working models": correlated effects or hierarchical effects.
- Uses semi-efficient diagonal weights:

$$w_{ij} = \frac{1}{n_j\left(\bar{s}_j^2 + \hat{\tau}^2\right)}, \qquad \text{where } \bar{s}_j^2 = \frac{1}{n_j}\sum_{i=1}^{n_j} s_{ij}^2$$

- Studies contributing more effects get less weight in meta-regressions that have within-study predictors.
  - Similar to meta-regression after aggregating to the study level.

BILINGUALISM EFFECTS BY DOMAIN [ROBUST SE]

	(1) Aggregated effects	(2) Shifting unit-of-analysis	(3) Multivariate meta-analysis	(4) Multi-level meta-analysis	(5) Default RVE (HIER weights)
Inhibition (95 studies, 212 ES)	0.077 [0.080]	0.114*** [0.033]	0.106** [0.035]	0.115*** [0.036]	0.103*** [0.032]
Monitoring (81 studies, 184 ES)	0.003 [0.111]	0.077 [0.041]	0.058 [0.036]	0.065 [0.037]	0.061* [0.036]
Shifting (37 studies, 79 ES)	0.147 [0.117]	0.147* [0.059]	0.141** [0.057]	0.148** [0.058]	0.135** [0.061]
Attention (18 studies, 53 ES)	0.230 [0.173]	-0.013 [0.076]	-0.031 [0.087]	-0.021 [0.085]	0.015 [0.103]
Working Memory (73 studies, 243 ES)	0.045 [0.073]	0.058 [0.043]	0.057 [0.036]	0.064* [0.037]	0.072 [0.048]
Fluency (28 studies, 98 ES)	-0.313** [0.127]	-0.260*** [0.071]	-0.211** [0.057]	-0.196*** [0.057]	-0.222*** [0.056]
Between-study SD	0.155	0.224	0.150	0.249	0.220
Within-study SD			0.276	0.217	0.211

RE-ANALYSIS OF SELF-CONTROL TRAINING STUDIES

	(1) Aggregated effects	(2) Shifting unit-of-analysis	(3) Multivariate meta-analysis	(4) Multi-level meta-analysis	(5) Robust variance estimation
Overall Average ES (33 studies, 166 ES)	0.281*** [0.059]		0.261*** [0.052]	0.263*** [0.054]	0.289*** [0.060]
Between-study SD	0.207		0.202	0.254	0.289
Within-study SD			0.143	0.027	
Moderator analysis by type of outcome					
Stamina (16 studies, 31 ES)	0.579*** [0.157]	0.413** [0.093]	0.359*** [0.077]	0.351*** [0.071]	0.579*** [0.123]
Strength (28 studies, 135 ES)	0.199** [0.071]	0.171** [0.064]	0.236*** [0.054]	0.238*** [0.055]	0.203** [0.065]
Difference	-0.380* [0.185]	-0.243* [0.113]	-0.123 [0.072]	-0.112 [0.059]	-0.376* [0.136]

DISCUSSION

- Robust "sandwich" variance estimation can be used with ***any*** of the available methods for handling dependence.
  - R packages metafor + clubSandwich.
- Default RVE should not be used for meta-regression with predictors that vary within study.
- Meta-analysts need to pay more attention to within- versus between-study variation in moderators.
- Improve software to make multivariate meta-analysis easier to implement.
- Outstanding problem: methods for examining ***publication/outcome reporting bias*** while handling dependent effects.

THANKS

James E. Pustejovsky
pusto@austin.utexas.edu
https://jepusto.com

REFERENCES

Becker, B. J. (2000). Multivariate Meta-analysis. In S. D. Brown & H. E. A. Tinsley (Eds.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (pp. 499–525). San Diego, CA: Academic Press.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons, Ltd.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. http://doi.org/10.1037/a0032968

Cooper, H. M. (1998). *Synthesizing Research: A Guide for Literature Reviews* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Friese, M., Frankenbach, J., Job, V., & Loschelder, D. D. (2017). Does Self-Control Training Improve Self-Control? A Meta-Analysis. *Perspectives on Psychological Science*, *12*(6), 1077–1099. http://doi.org/10.1177/1745691617697076

Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed., pp. 357–376). New York, NY: Russell Sage Foundation.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. http://doi.org/10.1002/jrsm.5

Hong, C., D. Riley, R., & Chen, Y. (2018). An improved method for bivariate meta-analysis when within-study correlations are unknown. *Research synthesis methods*, *9*(1), 73-88.

Kalaian, H. A., & Raudenbush, S. W. (1996). A multivariate mixed linear model for meta-analysis. *Psychological Methods*, *1*(3), 227–235. http://doi.org/10.1037/1082-989X.1.3.227

MORE REFERENCES

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, 144(4), 394-425 .

Olkin, I., & Siotani, M. (1976). Asymptotic distribution of functions of a correlation matrix. In S. Ikeda (Ed.), *Essays in Probability and Statistics*. Tokyo, Japan: Shinko Tsusho Co., Ltd., p. 235-251.

Raudenbush, S. W., Becker, B. J., & Kalaian, H. a. (1988). Modeling multivariate effect sizes. *Psychological Bulletin*, *103*(1), 111–120. http://doi.org/10.1037/0033-2909.103.1.111

Riley, R. D. (2009). Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(4), 789-811.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, *20*(3), 375–393. http://doi.org/10.1037/met0000011

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*(6), 604–634. http://doi.org/10.3102/1076998615606099

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*(2), 576–594. http://doi.org/10.3758/s13428-012-0261-6

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294. http://doi.org/10.3758/s13428-014-0527-2