META-ANALYSIS OF DEPENDENT EFFECT SIZES:
A REVIEW AND CONSOLIDATION OF METHODS

James E. Pustejovsky, UT Austin
Beth Tipton, Columbia University
Ariel Aloe, University of Iowa

AERA NYC
April 15, 2018
pusto@austin.utexas.edu
BASIC META-ANALYSIS METHODS ASSUME INDEPENDENT EFFECT SIZES

In a meta-analysis of experiments:

\[
\begin{array}{c|c}
\text{Treatment} & O \\
\hline
\text{Control} & O \\
\end{array}
\]
\[d_1\]

\[
\begin{array}{c|c}
\text{Treatment} & O \\
\hline
\text{Control} & O \\
\end{array}
\]
\[d_2\]

\[
\begin{array}{c|c}
\text{Treatment} & O \\
\hline
\text{Control} & O \\
\end{array}
\]
\[d_3\]

In a meta-analysis of correlations:

\[
\begin{array}{ccc}
\text{Study 1} & \text{Study 2} & \text{Study 3} \\
\hline
\text{a} & \text{b} & \text{c} \\
\text{a} & \text{b} & \text{d} \\
\text{a} & \text{b} & \text{e} \\
\end{array}
\]
\[r_1\]

\[
\begin{array}{ccc}
\text{Study 1} & \text{Study 2} & \text{Study 3} \\
\hline
\text{a} & \text{b} & \text{c} \\
\text{a} & \text{b} & \text{d} \\
\text{a} & \text{b} & \text{e} \\
\end{array}
\]
\[r_2\]

\[
\begin{array}{ccc}
\text{Study 1} & \text{Study 2} & \text{Study 3} \\
\hline
\text{a} & \text{b} & \text{c} \\
\text{a} & \text{b} & \text{d} \\
\text{a} & \text{b} & \text{e} \\
\end{array}
\]
\[r_3\]
BUT DEPENDENT EFFECT SIZES ARE VERY COMMON IN PRACTICE

Multiple outcomes measured on common set of participants

<table>
<thead>
<tr>
<th>Treatment</th>
<th>O</th>
<th>P</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>O</td>
<td>P</td>
<td>Q</td>
</tr>
</tbody>
</table>

\[
d_{O1} \quad d_{P1} \quad d_{Q1}
\]

Multiple treatment conditions compared to a common control

<table>
<thead>
<tr>
<th>Treatment T</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment U</td>
<td>O</td>
</tr>
<tr>
<td>Control</td>
<td>O</td>
</tr>
</tbody>
</table>

\[
d_{T3} \quad d_{U3}
\]

Outcome measured at multiple follow-up times

<table>
<thead>
<tr>
<th>Treatment</th>
<th>(O_1)</th>
<th>(O_2)</th>
<th>(O_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>(O_1)</td>
<td>(O_2)</td>
<td>(O_3)</td>
</tr>
</tbody>
</table>

\[
d_{12} \quad d_{22} \quad d_{32}
\]

Multiple correlations from a common sample

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
r_{ab} \quad r_{ac} \quad r_{ad}
\]

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

✓ 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)
✓ 1-52 effect size estimates per study (median = 2)
**CORRELATIONS BETWEEN ES ESTIMATES**

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 (often not available)

Multiple correlations from common sample
- known, icky formulas (Steiger, 1980)
- need to know correlations between ALL variables involved
METHODS FOR HANDLING DEPENDENCE

Becker (2000) described four broad strategies:

- **Ignore**
  - Aggregated effects (Borenstein et al., 2009)

- **Combine**
  - Shifting unit-of-analysis (Cooper, 1998)

- **Sub-classify**

- **Model**
  - Robust variance estimation (Hedges, Tipton, & Johnson, 2010)
  - Multivariate meta-analysis (Raudenbush, Becker, & Kalaian, 1988; Kalaian & Raudenbush, 1996)
  - Multi-level meta-analysis (Van den Noortgate et al., 2013, 2015)

Which one should I use?
## RE-ANALYSIS OF SELF-CONTROL TRAINING STUDIES

<table>
<thead>
<tr>
<th></th>
<th>(1) Aggregated effects</th>
<th>(2) Shifting unit-of-analysis</th>
<th>(3) Multivariate meta-analysis</th>
<th>(4) Multi-level meta-analysis</th>
<th>(5) Robust variance estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Average ES</td>
<td>0.281*** [0.059]</td>
<td>0.261*** [0.052]</td>
<td>0.263*** [0.054]</td>
<td>0.289*** [0.060]</td>
<td></td>
</tr>
<tr>
<td>(k = 33, N = 166)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-study SD</td>
<td>0.207</td>
<td>0.202</td>
<td>0.254</td>
<td>0.289</td>
<td></td>
</tr>
<tr>
<td>Within-study SD</td>
<td></td>
<td>0.143</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Moderator analysis by type of outcome

<table>
<thead>
<tr>
<th></th>
<th>(1) Aggregated effects</th>
<th>(2) Shifting unit-of-analysis</th>
<th>(3) Multivariate meta-analysis</th>
<th>(4) Multi-level meta-analysis</th>
<th>(5) Robust variance estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stamina (k = 16, N = 31)</td>
<td>0.579*** [0.157]</td>
<td>0.413*** [0.093]</td>
<td>0.359*** [0.077]</td>
<td>0.351*** [0.071]</td>
<td>0.579*** [0.123]</td>
</tr>
<tr>
<td>Strength (k = 28, N = 135)</td>
<td>0.199** [0.071]</td>
<td>0.171** [0.064]</td>
<td>0.236*** [0.054]</td>
<td>0.238*** [0.055]</td>
<td>0.203** [0.065]</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.380* [0.185]</td>
<td>-0.243* [0.113]</td>
<td>-0.123 [0.072]</td>
<td>-0.112 [0.059]</td>
<td>-0.376* [0.136]</td>
</tr>
</tbody>
</table>
AGGREGATED EFFECTS

- Average estimates to generate single “synthetic” ES per study.

- Estimating variance of synthetic ES requires correlations among component ES (Borenstein et al., 2009).
  - Common to use a rough approximation assuming $r \approx 1$.

- Limits moderator/meta-regression analyses to between-study predictors.
SUB-GROUPS/SHIFTING UNIT-OF-ANALYSIS

- If ES can be classified into sub-groups where each study contributes ≤ 1 ES estimate, then univariate meta-analysis can be conducted within sub-groups.

- If there are still multiple ES per sub-group, aggregate (Cooper, 1998).
  - Need correlations between effects within sub-group in order to get variances of aggregated effects.

- 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?
**MULTIVARIATE META-ANALYSIS**

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

- Hierarchical model for component ES estimates nested within studies

\[ T_{ij} = x_{ij} \beta + u_j + v_{ij} + e_{ij} \]

where \( u_j \sim N(0, \tau^2) \), \( v_{ij} \sim N(0, \omega^2) \), \( e_{ij} \sim N(0, s_{ij}^2) \), \( \text{Cov}(e_{hj}, e_{ij}) = r_{hij} s_{hj} s_{ij} \).

- Requires estimates/assumptions about ES correlations \( r_{hij} \).
  - In the example, I calculated \( r \) for multiple T-common C studies, assumed \( r = 0.17 \) for multiple outcomes/time-points.

- 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

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

\[ T_{ij} = x_{ij} \beta + u_j + v_{ij} + e_{ij} \]

where \( u_j \sim N(0, \tau^2) \), \( v_{ij} \sim N(0, \omega^2) \), \( e_{ij} \sim N(0, s_{ij}^2) \), \( \text{Cov}(e_{hj}, e_{ij}) = 0 \).

- Simulation evidence indicates that this approach can be “robust” to mis-specified correlation structure.

- But unclear whether robustness holds generally.
ROBUST VARIANCE ESTIMATION
(Hedges, Tipton, & Johnson, 2010)

- Meta-analysis/meta-regression using “sandwich” variance estimation methods
  - robust to mis-specified/unknown correlations between ES within studies.
  - sandwich estimation methods apply to very general class of models.

- RVE implementation involves
  - choosing between “correlated effects” or “hierarchical effects” working models.
  - making “working” assumption about correlation between ES estimates.

- Uses semi-efficient diagonal weights:
  \[ w_{ij} = \frac{1}{n_j(\overline{s}_j^2 + \hat{\tau}^2)}, \quad \text{where} \quad \overline{s}_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} s_{ij}^2 \]

- Studies with more effects will get less weight in meta-regressions that have within-study predictors.
## RE-ANALYSIS OF SELF-CONTROL TRAINING STUDIES

<table>
<thead>
<tr>
<th></th>
<th>(1) Aggregated effects</th>
<th>(2) Shifting unit-of-analysis</th>
<th>(3) Multivariate meta-analysis</th>
<th>(4) Multi-level meta-analysis</th>
<th>(5) Robust variance estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall Average ES</strong></td>
<td>0.281*** [0.059]</td>
<td>0.261*** [0.052]</td>
<td>0.263*** [0.054]</td>
<td>0.289*** [0.060]</td>
<td></td>
</tr>
<tr>
<td>(k = 33, N = 166)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Between-study SD</strong></td>
<td>0.207</td>
<td>0.202</td>
<td>0.254</td>
<td>0.289</td>
<td></td>
</tr>
<tr>
<td><strong>Within-study SD</strong></td>
<td></td>
<td>0.143</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Moderator analysis by type of outcome

<table>
<thead>
<tr>
<th>Type</th>
<th>(1) Aggregated effects</th>
<th>(2) Shifting unit-of-analysis</th>
<th>(3) Multivariate meta-analysis</th>
<th>(4) Multi-level meta-analysis</th>
<th>(5) Robust variance estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stamina</strong></td>
<td>0.579*** [0.157]</td>
<td>0.413*** [0.093]</td>
<td>0.359*** [0.077]</td>
<td>0.351*** [0.071]</td>
<td>0.579*** [0.123]</td>
</tr>
<tr>
<td>(k = 16, N = 31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Strength</strong></td>
<td>0.199** [0.071]</td>
<td>0.171** [0.064]</td>
<td>0.236*** [0.054]</td>
<td>0.238*** [0.055]</td>
<td>0.203** [0.065]</td>
</tr>
<tr>
<td>(k = 28, N = 135)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Difference</strong></td>
<td>-0.380* [0.185]</td>
<td>-0.243* [0.113]</td>
<td>-0.123 [0.072]</td>
<td>-0.112 [0.059]</td>
<td>-0.376* [0.136]</td>
</tr>
</tbody>
</table>

**Overall Average ES**

- Overall average effect size of self-control training studies.

**Between-study SD**

- Standard deviation between studies.

**Within-study SD**

- Standard deviation within studies.

**Stamina**

- Effect size for stamina.

**Strength**

- Effect size for strength.

**Difference**

- Difference in effect size compared to the overall average.
# COMPARISON

<table>
<thead>
<tr>
<th></th>
<th>Aggregated effects</th>
<th>Shifting unit-of-analysis</th>
<th>Multivariate meta-analysis</th>
<th>Multi-level meta-analysis</th>
<th>Robust variance estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires making “working” assumption about correlations</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Robustness to correlation assumptions</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>Robust</td>
</tr>
<tr>
<td>Meta-regression specification</td>
<td>Limited</td>
<td>Limited</td>
<td>Flexible</td>
<td>Flexible</td>
<td>Flexible</td>
</tr>
<tr>
<td>Random effects specification</td>
<td>Limited</td>
<td>Somewhat limited</td>
<td>Flexible</td>
<td>Flexible</td>
<td>Limited</td>
</tr>
</tbody>
</table>
CONSOLIDATION

- Robust “sandwich” variance estimation can be used with *any* of the methods.
- Default RVE weights should not be used for meta-regression with predictors that vary within study.
- Multi-level meta-analysis = multi-variate meta-analysis assuming \( r = 0 \).
- More attention to within- versus between-study variation in moderators.
- Improve computational tools to make multivariate meta-analysis easier to implement.
REFERENCES


