Introduction to Meta-Analysis and Systematic Reviews

Noel A. Card

Professor, Dept. of Human Development & Family Studies

noel.card@uconn.edu

2012, Guilford Press
Overview of talk

• Role of literature reviews in the accumulation of scientific knowledge.
• Steps in performing meta-analysis.
• Meta-analysis and other approaches.
Overview of talk

- Role of literature reviews in the accumulation of scientific knowledge.
  - Progress in science.
  - Types of literature review.
- Steps in performing meta-analysis.
- Meta-analysis and other approaches.
Progress in science

• Science is often said to build on prior work in a cumulative manner.

  – “If I have seen further it is by standing upon the shoulders of giants.” - Isaac Newton, 1675
Progress in science

• Science is often said to build on prior work in a cumulative manner.

• **Normal Science** *(Kuhn, 1962 *Structure of Scientific Revolutions*)
  
  – Periods in which scientific knowledge advances by accumulating evidence within the dominant paradigm.
  
  – Sciences (established disciplines) exist primarily in this period, and periods of revolution (paradigm shifts) should be rare.
Progress in science

• Science is often said to build on prior work in a cumulative manner.

• Normal Science (Kuhn, 1962 Structure of Scientific Revolutions)

• Barriers to orderly progression of science:
  – Multiple “blueprints” (i.e., research agendas, theoretical foundations, professional pressures incentivize novelty).
  – Information overload (i.e., increasing numbers of researchers, research production, journals / outlets).
  – Studies differ in samples, methodologies, measures… (i.e., exact, or even deliberate inexact, replication is rare).
Progress in science

- Science is often said to build on prior work in a cumulative manner.
- Normal Science (Kuhn, 1962 *Structure of Scientific Revolutions*)
- Barriers to orderly progression of science
- The solution… literature reviews / meta-analysis.
  - Research questions informed by data (not necessarily researcher’s original intention).
  - Synthesize many studies (as many as you are willing to code).
  - Combine and compare across diverse types of studies.
Types of literature review

• A regular part of scholarship is to review literature.
  – Reading literature to inform practice or policy.
  – Literature reviews for comprehensive exams or “Chapter 1” of dissertation.
  – Writing chapter for edited book.
  – Reviewing literature in Introduction section of manuscript.
Types of literature review

- A regular part of scholarship is to review literature.
- However, few scholars formally trained in performing literature reviews.
  - Common doctoral coursework:
    - 1-2 courses in research methodology.
    - 2-5 courses in data analysis.
    - Additional courses / work developing expertise in conducting primary studies in specialized area.
    - 0 courses / formal training in literature review / MA
Types of literature review

• A regular part of scholarship is to review literature.
• Literature reviews vary along several dimensions (Cooper & Hedges, 1994)
  – **Audience** (e.g., specialized scholars, general scholars, general public).
  – **Perspective** (e.g., neutral versus espousing position).
  – **Coverage** (e.g., central, representative, exhaustive w/ selective citation, exhaustive).
  – **Focus** – *see next slide.*
  – **Method of synthesis** – *see next slide.*
Types of literature review

Super-ordinate category:

Focus:

Method of synthesis:
Recap

• Role of literature reviews in the accumulation of scientific knowledge.
  – Progress in science.
  – Types of literature review.
• Steps in performing meta-analysis.
• Meta-analysis and other approaches.
Roadmap

• Role of literature reviews in the accumulation of scientific knowledge.

• Steps in performing meta-analysis.
  – Research questions appropriate for MA.
  – Searching for literature.
  – Coding the literature.
  – Synthesizing research and drawing conclusions.
  – Presenting findings.

• Meta-analysis and other approaches.
Roadmap

• Role of literature reviews in the accumulation of scientific knowledge.

• Steps in performing meta-analysis.
  – Research questions appropriate for MA.
  – Searching for literature.
  – Coding the literature.
  – Synthesizing research and drawing conclusions.
  – Presenting findings.

• Meta-analysis and other approaches.
Searching for Literature

• Inclusion versus exclusion criteria.
  – Set of statements that define that population of studies included in a meta-analysis.
    • Guide decisions to include studies.
    • Communicate to readers.
Searching for Literature

• Inclusion versus exclusion criteria.
  – Set of statements that define that population of studies included in a meta-analysis.
  – Common criteria:
    • Characteristics of intervention.
    • Operational definitions of variables.
    • Research methods (e.g., type of design, placebo control).
    • Sample characteristics (e.g., age range).
    • Cultural range (e.g., publication language).
Searching for Literature

• Search strategy.
  – Specific approaches:
    • **Electronic databases:** Specify keywords, Boolean operations, disciplinary versus general.
    • **Forward searches:** Studies citing seminal works in area.
    • **Backward searches:** Reading literature cited in other reports.
    • **Conference programs:** Captures newer work.
    • **Granting agencies:** Research evaluated before results are known.
    • **Clinical trial / preregistration databases**
    • **Email / listserv requests:** Especially for unpublished work.
Searching for Literature

• Search strategy.
  – Specific approaches
  – Balance between:
    • Recall = theoretical % of documents retrieved to those that should be retrieved.
    • Precision = % retrieved documents that are relevant.
  – Meta-analyses (and other systematic reviews) aim to be exhaustive \(\rightarrow\) maximize recall to reduce threat of publication bias
Searching for Literature

- Publication bias.

Population of effect sizes

Published effect sizes
Roadmap

• Role of literature reviews in the accumulation of scientific knowledge.

• Steps in performing meta-analysis.
  – Research questions appropriate for MA.
  – Searching for literature.
  – Coding the literature.
  – Synthesizing research and drawing conclusions.
  – Presenting findings.

• Meta-analysis and other approaches.
Coding the Literature

- Systematically code information from studies.
  - Study characteristics.
    - Features of sample, methodology, measurement.
    - Used to...
      - Describe state of research.
      - Predict differences in study results.
Coding the Literature

• Systematically code information from studies.
  – Study characteristics.
  – Study results.
  • Coded as Effect Sizes (!!!)
Coding the Literature

• Effect sizes.
  – Effect sizes are not...
    • Author conclusions in Discussion.
    • Significance tests (i.e., \( p \) values).
Coding the Literature

• Effect sizes.
  – Effect sizes are not...
    • Author conclusions in Discussion.
    • Significance tests (i.e., \( p \) values).
  – Effect sizes are...
    • Comparable / combinable across studies.
    • Calculable from a wide range of reported results (e.g., summary statistics, t-tests, ANOVA results, correlations, odds ratios).
    • Understandable / meaningful to readers.
Coding the Literature

• Effect sizes.
  – Many types of effect sizes
    • \( r \) = association between two (continuous) variables.
    • \( d \) or \( g \) = standardized mean difference between two groups.
    • Odds ratio = contingency / association between two dichotomous variables.
Coding the Literature

• Effect sizes.
  – Many types of effect sizes
    • $r$ = association between two (continuous) variables.
    • $d$ or $g$ = standardized mean difference between two groups.
    • Odds ratio = contingency / association between two dichotomous variables.
    • Mean / proportion = central tendency of one variable.
    • Unstandardized mean difference = retains meaningful information about scale.
    • Longitudinal change scores / pre-post differences.
    • Psychometric properties of measures.
Coding the Literature

• Effect sizes.
  – Many types of effect sizes
  – Tasks of meta-analyst:
    • Decide type of study result to extract.
    • Decide type of effect size best suited for this type of result.
    • Compute this effect size from diverse ways results reported.
Coding the Literature

- **Effect sizes.**
  - Many types of effect sizes
  - Tasks of meta-analyst.
  - Computing effect sizes:
    - Formulas reported in books and articles.
    - Use program / online calculator.

![Example Example](image-url)
Roadmap

• Role of literature reviews in the accumulation of scientific knowledge.

• Steps in performing meta-analysis.
  – Research questions appropriate for MA.
  – Searching for literature.
  – Coding the literature.
  – Synthesizing research and drawing conclusions.
  – Presenting findings.

• Meta-analysis and other approaches.
Synthesizing Effect Sizes

- Two major goals:
  - Combining effect sizes across studies.
    - Overall average effect size.
  - Comparing effect sizes across studies.
    - Coded study characteristics as predictors of effect sizes.
Synthesizing Effect Sizes

• Combining effect sizes across studies.
  – Studies are weighted by the precision of the point estimate of effect size.
  – A common weighting: \( w_i = \frac{1}{SE_i^2} \)

• Where \( SE_i \) is the standard error for study \( i \).
• Larger \( N \) studies receive larger weight
Synthesizing Effect Sizes

• Combining effect sizes across studies.
  – Weighting only by a function of sampling error would only by if…

Sampling error only
Combining effect sizes across studies.
  - Weighting only by a function of sampling error would only be if there is only sampling error variability

Sampling error only

Sampling error + other variability
Synthesizing Effect Sizes

• Combining effect sizes across studies.
  – When there is significant heterogeneity (i.e., other between-study variability), it is necessary to account for both sources of imprecision → Random-effects model

\[ T_i = \theta + \varepsilon_i \]

where \( \varepsilon_i \) is the sampling error of \( T_i \)

\[ T_i = \mu + \xi_i + \varepsilon_i \]

where \( \theta_i \) is the population effect size for each study, \( \xi_i \) is the study-deviation from \( \mu \), and \( \varepsilon_i \) is the sample-sampling error
Synthesizing Effect Sizes

• Combining effect sizes across studies.
  – Another way to think about random-effects models:

Population of all possible studies of a phenomenon

Variability due to population studied, methodology, measurement, or ???

Variability due to sampling error / fluctuation

Study 1  Study 2  Study 3  Study 4
ε₁  ε₂  ε₃  ε₄

Low  High

Observed ES
Synthesizing Effect Sizes

• Combining effect sizes across studies.
  – Studies are weighted by the precision of the point estimate of effect size.
  – A more appropriate (random effects) weighting:

\[ w_i = \frac{1}{SE_i^2 + \tau^2} \]

• Where \( SE_i \) is the standard error for study \( i \).
• \( \tau^2 \) is the estimated between-study variance above and beyond sampling error variance.
Synthesizing Effect Sizes

- Two major goals:
  - Combining effect sizes across studies.
    - Overall average effect size.
  - Comparing effect sizes across studies.
    - Coded study characteristics as predictors of effect sizes.
Synthesizing Effect Sizes

• Comparing effect sizes across studies.
  – Rather than simply modeling heterogeneity as “other between-study variability”…
  – … attempt to predict why studies vary in their effect sizes.
Synthesizing Effect Sizes

- Comparing effect sizes across studies.
  - Two general approaches:
    - ANOVA approach for categorical study characteristics
    - Regression approach for continuous study characteristics
      - Regression approach can also accommodate categorical, multiple predictors, nonlinear, etc.
Synthesizing Effect Sizes

- Comparing effect sizes across studies.
  - ANOVA approach for categorical study characteristics
- Partitions heterogeneity (i.e., variance) into within- and between-group components:

\[ Q_{Total} = Q_{Between} + Q_{Within} \]

\[ Q_{Total} = \sum [w_i(ES_i - GM)^2] = (\sum w_i ES_i^2) - \frac{(\sum w_i ES_i)^2}{\sum w_i} \]

\[ Q_{Between} = \sum [w_j(M_j - GM)^2] = (\sum w_j M_j^2) - \frac{(\sum w_j M_j)^2}{\sum w_j} \]

\[ Q_{within} = \sum^J \sum^I w_i(ES_i - M_j)^2 = Q_1 + Q_2 + \ldots + Q_J \]

- Significance indicated by \( Q_{Between} \)
Synthesizing Effect Sizes

• Comparing effect sizes across studies.
  – Regression approach for continuous study characteristics.

• Use standard software package (e.g., SPSS, SAS, R) to perform weighted regression
Synthesizing Effect Sizes

- Comparing effect sizes across studies.
  - Regression approach for continuous study characteristics.
  - Can extend regression approach to include multiple predictors
- Useful if study characteristics co-occur.
  - E.g., Measures used with different samples.
  - E.g., Methodological quality versus publication status.
Synthesizing Effect Sizes

• Comparing effect sizes across studies.
  – Regression approach for continuous study characteristics.
  – Can extend regression approach to include multiple predictors
  – Can add interactions and powered polynomials
    • Interactions:
      – Are ethnic differences in results more evident with some measures than others?
    • Powered polynomials: Test nonlinear associations of study characteristics.
      – E.g., Is maximum therapy effectiveness reached after a certain number of sessions?
Synthesizing Effect Sizes

• Summary.
  – Methods of analysis are not much more complex than other data analyses.
    • Though there are complex extensions.
  – For full details see e.g.,:
Recap

• Role of literature reviews in the accumulation of scientific knowledge.

• Steps in performing meta-analysis.
  – Research questions appropriate for MA.
  – Searching for literature.
  – Coding the literature.
  – Synthesizing research and drawing conclusions.
  – Presenting findings.

• Meta-analysis and other approaches.
Roadmap

- Role of literature reviews in the accumulation of scientific knowledge.
- Steps in performing meta-analysis.
- Meta-analysis and other approaches.
  - Primary studies.
  - (Non-meta-analytic) systematic reviews.
  - Integrative data analysis.
  - Replication, MA, and accumulation of knowledge.
MA and Other Approaches

- Advantages of Meta-Analysis relative to Primary Studies:
  - Greater statistical power / precision.
  - Broader representation of samples, measures, countries, times, etc.
    - Tests generalizability of an effect
    - Compares effects across samples etc. (moderation)
- Advantages of primary studies:
  - Evaluating complex multivariate models.
  - Absence of previous studies.
MA and Other Approaches

• Advantages of Meta-Analysis over Narrative Reviews:
  – Discipline / guidance for searching, coding, analysis.
  – Handles large amount of information.
  – Conclusions statistically defensible / debatable.
  – Systematic coding and analysis might be more precise.

• Advantages of Narrative Reviews:
  – Possibility of handling study results not easily converted to effect sizes.
  – (Arguably) deeper understanding of individual studies.
MA and Other Approaches

- **Integrative Data Analysis (IDA)** = Combination and analysis of *raw data* from multiple data sets.

- **Advantages of IDA:**
  - Allows analysis of within- and between-study variability. (E.g., analysis of individual-level demographic information rather than aggregating to the study level)
  - Explicit attention to harmonizing and evaluating measurement equivalence across data sets.

- **Advantages of MA over IDA:**
  - Easier (AKA, you can actually do it)
  - As good as IDA for overall average and between-study comparisons.
MA and Other Approaches

• Meta-analysis directly responds to the ‘replication crisis’
  – Cumulative meta-analysis = studies are added one at a time, and results are updated.
  – A new model of programmatic lines of research (versus single-study focus):
MA and Other Approaches

- Meta-analysis directly responds to the ‘replication crisis’
  - Cumulative meta-analysis = studies are added one at a time, and results are updated.
  - A new model of programmatic lines of research (versus single-study focus):

A) Exact replications

```
<table>
<thead>
<tr>
<th>Effect size estimate</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Role of Lit Rev.  Steps in MA  Other App.
MA and Other Approaches

- Meta-analysis directly responds to the ‘replication crisis”
  - Cumulative meta-analysis = studies are added one at a time, and results are updated.
  - A new model of programmatic lines of research (versus single-study focus):

B) Unplanned inexact replications

![Graph](image)
MA and Other Approaches

- Meta-analysis directly responds to the ‘replication crisis’
  - Cumulative meta-analysis = studies are added one at a time, and results are updated.
  - A new model of programmatic lines of research (versus single-study focus):

C) Planned inexact replications

![Graph showing planned inexact replications](image)
Recap

- Role of literature reviews in the accumulation of scientific knowledge.
- Steps in performing meta-analysis.
- Meta-analysis and other approaches.
  - Primary studies.
  - (Non-meta-analytic) systematic reviews.
  - Integrative data analysis.
  - Replication, MA, and accumulation of knowledge.
Final thoughts

• Literature reviews and meta-analysis play a critical role in accumulating scientific knowledge.
  – Synthesis of existing studies at least as important as creating more studies.
Final thoughts

- Literature reviews and meta-analysis play a critical role in accumulating scientific knowledge.
- The steps of basic meta-analysis are reasonably tractable.
  - One semester course / one week workshop provides working knowledge.
  - Several accessible books.
Final thoughts

• Literature reviews and meta-analysis play a critical role in accumulating scientific knowledge.
• The steps of basic meta-analysis are reasonably tractable.
• Meta-analysis is a valuable piece of (all of our) research agendas.
  – Provides different approach in addition to other efforts
  – Not impeded by IRB issues, inaccessibility of participants, etc.
  – Informs most important next steps for primary studies.
Thanks!!

Questions / Discussion.

Contact: noel.card@uconn.edu
### App. 1  (effect sizes are not significance tests)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>4</td>
<td>S2</td>
</tr>
<tr>
<td>S3</td>
<td>4</td>
<td>S4</td>
</tr>
<tr>
<td>S5</td>
<td>3</td>
<td>S6</td>
</tr>
<tr>
<td>S7</td>
<td>2</td>
<td>S8</td>
</tr>
<tr>
<td>S9</td>
<td>2</td>
<td>S10</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sum</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Mean</td>
<td>3.00</td>
<td>5.00</td>
</tr>
<tr>
<td>S</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

\[
t(8) = \left( \frac{M1 - M2}{\frac{S_{within}}{\sqrt{N}}} \right) \left( \frac{\sqrt{N}}{2} \right) = 3.16, \quad p = .013
\]
<table>
<thead>
<tr>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 = 4</td>
<td>S2 = 6</td>
</tr>
<tr>
<td>S3 = 4</td>
<td>S4 = 6</td>
</tr>
<tr>
<td>S5 = 3</td>
<td>S6 = 5</td>
</tr>
<tr>
<td>S7 = 2</td>
<td>S8 = 4</td>
</tr>
<tr>
<td>S9 = 2</td>
<td>S10 = 4</td>
</tr>
</tbody>
</table>

| N       | 5 | 5 |
| Sum     | 15 | 25 |
| Mean    | 3.00 | 5.00 |
| S       | 1.00 | 1.00 |

\[
t(8) = \left( \frac{M_1 - M_2}{S_{\text{within}}} \right) \left( \sqrt{\frac{N}{2}} \right) = 3.16, \quad p = .013
\]

<table>
<thead>
<tr>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 = 4</td>
<td>S2 = 6</td>
</tr>
<tr>
<td>S3 = 3</td>
<td>S4 = 5</td>
</tr>
<tr>
<td>S5 = 2</td>
<td>S6 = 4</td>
</tr>
</tbody>
</table>

| N       | 3 | 3 |
| Sum     | 9 | 15 |
| Mean    | 3.00 | 5.00 |
| S       | 1.00 | 1.00 |

\[
t(4) = \left( \frac{M_1 - M_2}{S_{\text{within}}} \right) \left( \sqrt{\frac{N}{2}} \right) = 2.45, \quad p = .071
\]
### App. 1  (effect sizes are not significance tests)

<table>
<thead>
<tr>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 = 4</td>
<td>S2 = 6</td>
</tr>
<tr>
<td>S3 = 4</td>
<td>S4 = 6</td>
</tr>
<tr>
<td>S5 = 3</td>
<td>S6 = 5</td>
</tr>
<tr>
<td>S7 = 2</td>
<td>S8 = 4</td>
</tr>
<tr>
<td>S9 = 2</td>
<td>S10 = 4</td>
</tr>
</tbody>
</table>

| N       | 5         | 5         |
| Sum     | 15        | 25        |
| Mean    | 3.00      | 5.00      |
| S       | 1.00      | 1.00      |

\[ t(8) = \left( \frac{M_1 - M_2}{S_{\text{within}}} \right) \left( \frac{\sqrt{N}}{2} \right) = 3.16, \quad p = .013 \]

<table>
<thead>
<tr>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 = 4</td>
<td>S2 = 6</td>
</tr>
<tr>
<td>S3 = 4</td>
<td>S4 = 6</td>
</tr>
<tr>
<td>S5 = 3</td>
<td>S6 = 5</td>
</tr>
<tr>
<td>S7 = 2</td>
<td>S8 = 4</td>
</tr>
<tr>
<td>S9 = 2</td>
<td>S10 = 4</td>
</tr>
<tr>
<td>S11 = 4</td>
<td>S12 = 4</td>
</tr>
<tr>
<td>S13 = 4</td>
<td>S14 = 4</td>
</tr>
<tr>
<td>S15 = 3</td>
<td>S16 = 5</td>
</tr>
<tr>
<td>S17 = 2</td>
<td>S18 = 4</td>
</tr>
<tr>
<td>S19 = 2</td>
<td>S20 = 4</td>
</tr>
</tbody>
</table>

| N       | 10         | 10         |
| Sum     | 30         | 50         |
| Mean    | 3.00       | 5.00       |
| S       | 1.00       | 1.00       |

\[ t = \left( \frac{M_1 - M_2}{S_{\text{within}}} \right) \left( \frac{\sqrt{N}}{2} \right) = 4.74, \quad p = .00016 \]
App. 1 (effect sizes are not significance tests)

• In the three examples, we have three different levels of significance:
  – N=3, $t = 2.45$, $p = .071$
  – N=5, $t = 3.16$, $p = .013$
  – N=10, $t = 4.74$, $p = .00016$

• However, the magnitude of the effect sizes in each sample is identical:
  – The differences in means was 2 standard deviations.
Difference between effect size and statistical significance

- Hypothesis testing = Evaluating whether the null hypothesis can be rejected (at a certain level of improbability, e.g., \( p < .05 \)).
- Effect size = Magnitude of difference between groups (or magnitude of relation between two variables, etc.).
- Relation between hypothesis testing and effect sizes:

\[
\begin{align*}
\text{test of significance} & = \text{effect size} \times \text{size of study} \\
\text{t}_{(\text{independent})} & = \frac{d}{\sqrt{\text{df}}} \\
t & = \frac{r}{\sqrt{(1 - r^2)}} \\
\chi^2(1) & = \frac{\Phi^2}{N}
\end{align*}
\]
App. 2 (Online calculator for ES)

- Calculate effect sizes from two studies
App. 2 (Online calculator for ES)

- Two studies from a review of the effectiveness of teen dating violence prevention programs in K-12 schools

- Note: I will compute different effect sizes (SMD and $r$) for an example. You would usually compute a common effect size for all studies in your MA.
• Standardized mean difference
    • Treatment N = 235
    • Control N = 205
    • Post-treatment comparison on 6 scales, two of which met inclusion criteria of review.
    • For this calculation, focus on “physical conflict tactics”
App. 2 (Online calculator for ES)

- Standardized mean difference
  - Post-treatment comparison…
  - …for this calculation, focus on “physical conflict tactics”

Table 3. Means and Standard Deviations of Relationship Beliefs and Conflict Tactics Scales’ Subscales

<table>
<thead>
<tr>
<th>Conflict tactics subscales</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Reasoning</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Verbal</td>
<td>2.66</td>
<td>1.26</td>
</tr>
<tr>
<td>Physical</td>
<td>1.91</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>.76</td>
<td>1.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relationship belief subscales</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggression beliefs</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Faulty relationship beliefs</td>
<td>3.37</td>
<td>.83</td>
</tr>
<tr>
<td>Realistic relationship beliefs</td>
<td>2.96</td>
<td>.53</td>
</tr>
<tr>
<td></td>
<td>2.89</td>
<td>.47</td>
</tr>
</tbody>
</table>

258
App. 2 (Online calculator for ES)

- Standardized mean difference
  
  - Coded information:
    
    - Treatment N = 235, $M = 0.82$, $SD = 1.33$
    - Control N = 205, $M = 1.17$, $SD = 1.67$
    - All necessary information to compute SMD
Standardized mean difference

Coded information: Treatment $N = 235, M = 0.82, SD = 1.33$
Control $N = 205, M = 1.17, SD = 1.67$

Practical Meta-Analysis Effect Size Calculator
David B. Wilson, Ph.D., George Mason University

This is a web-based effect-size calculator. It is designed to facilitate the computation of effect-sizes for meta-analysis. Four effect-size types can be computed from various input data: the standardized mean difference, the correlation coefficient, the odds ratio, and the risk-ratio.

This calculator is a companion to the 2001 book by Mark W. Lipsey and David B. Wilson, Practical Meta-analysis, published by Sage. An older Excel based version of the calculator can be found at http://mason.gmu.edu/~dwilsonb/ma.html. Additional tools for performing meta-analysis can also be found at that web address.

This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.
Standardized mean difference

Coded information: Treatment N = 235, $M = 0.82$, $SD = 1.33$
Control N = 205, $M = 1.17$, $SD = 1.67$
Standardized mean difference

Coded information:

Treatment N = 235, $M = 0.82$, $SD = 1.33$

Control N = 205, $M = 1.17$, $SD = 1.67$
Standardized mean difference

Coded information:

Treatment N = 235, $M = 0.82$, $SD = 1.33$

Control N = 205, $M = 1.17$, $SD = 1.67$
Standardized mean difference

Coded information: Treatment $N = 235, M = 0.82, SD = 1.33$
Control $N = 205, M = 1.17, SD = 1.67$

Effect size = $d = -0.2336$

$V = SE^2 = 0.0092$ (for weighting)
Second Example:
Correlation ($r$) between treatment and dating violence
• Correlation \( (r) \) between treatment and dating violence
  

### Table 2. Physical Dating Violence Reported in the Past Year at 2.5-Year Follow-up According to Experimental Group

<table>
<thead>
<tr>
<th>Students With PDV, No./Total No. (%)</th>
<th>Control Group</th>
<th>Intervention Group</th>
<th>ICC</th>
<th>OR (95% CI)</th>
<th>t Test</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted OR(^a)</td>
<td>74/754 (9.8)</td>
<td>72/968 (7.4)</td>
<td>0.02</td>
<td>1.42 (0.87-2.33)</td>
<td>1.49</td>
<td>.15</td>
</tr>
<tr>
<td>Adjusted OR(^b)</td>
<td></td>
<td></td>
<td></td>
<td>2.42 (1.00-6.02)</td>
<td>2.06</td>
<td>.05</td>
</tr>
<tr>
<td>Students who dated in the past 12 mo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted OR(^a)</td>
<td>71/369 (19.2)</td>
<td>72/480 (15.0)</td>
<td>0.02</td>
<td>1.37 (0.89-2.13)</td>
<td>1.53</td>
<td>.14</td>
</tr>
<tr>
<td>Adjusted OR(^b)</td>
<td></td>
<td></td>
<td></td>
<td>2.13 (0.81-5.66)</td>
<td>1.65</td>
<td>.12</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; OR, odds ratio; PDV, physical dating violence.

\(^a\) Odds ratios of the intervention effect from the multilevel model.

\(^b\) Odds ratios were adjusted for baseline behavior, stratifying variables, and sex (\(n=1722\) in the full model; \(n=1041\) in the model restricted to the dating sample).
App. 2  (Online calculator for ES)

- Correlation ($r$) between treatment and dating violence

<table>
<thead>
<tr>
<th>Table 2. Physical Dating Violence Reported in the Past Year at 2.5-Year Follow-up According to Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>All students</td>
</tr>
<tr>
<td>Unadjusted OR$^a$</td>
</tr>
<tr>
<td>Adjusted OR$^b$</td>
</tr>
<tr>
<td>Students who dated in the past 12 mo</td>
</tr>
<tr>
<td>Unadjusted OR$^a$</td>
</tr>
<tr>
<td>Adjusted OR$^b$</td>
</tr>
</tbody>
</table>

Subgroup analysis not relevant for this review

Abbreviations: CI, confidence interval; ICC, intracluster correlation coefficient; OR, odds ratio; PDV, physical dating violence.

$^a$Odds ratios of the intervention effect from the multilevel model.

$^b$Odds ratios were adjusted for baseline behavior, stratifying variables, and sex (n=1722 in the full model; n=1041 in the model restricted to the dating sample).
App. 2 (Online calculator for ES)

- Correlation ($r$) between treatment and dating violence

Table 2. Physical Dating Violence Reported in the Past Year at 2.5-Year Follow-up According to Experimental Group

<table>
<thead>
<tr>
<th>Students With PDV, No./Total No. (%)</th>
<th>Control Group</th>
<th>Intervention Group</th>
<th>ICC</th>
<th>OR (95% CI)</th>
<th>t Test</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td>74/754 (9.8)</td>
<td>72/968 (7.4)</td>
<td>0.02</td>
<td>1.42 (0.87-2.33)</td>
<td>1.49</td>
<td>.15</td>
</tr>
<tr>
<td>Unadjusted OR$^a$</td>
<td></td>
<td></td>
<td></td>
<td>2.42 (1.00-6.02)</td>
<td>2.06</td>
<td>.05</td>
</tr>
<tr>
<td>Adjusted OR$^b$</td>
<td></td>
<td></td>
<td></td>
<td>1.37 (0.89-2.13)</td>
<td>1.53</td>
<td>.14</td>
</tr>
<tr>
<td>Students who dated in the past 12 mo</td>
<td>54/629 (8.6)</td>
<td>58/864 (6.7)</td>
<td></td>
<td>2.13 (0.81-5.66)</td>
<td>1.65</td>
<td>.12</td>
</tr>
<tr>
<td>Unadjusted OR$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted OR$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subgroup analysis not relevant for this review

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; OR, odds ratio; PDV, physical dating violence.

$^a$Odds ratios of the intervention effect from the multilevel model.

$^b$Odds ratios were adjusted for baseline behavior, stratifying variables, and sex (n=1722 in the full model; n=1041 in the model restricted to the dating sample).
App. 2 (Online calculator for ES)

- Correlation ($r$) between treatment and dating violence

<table>
<thead>
<tr>
<th>Table 2. Physical Dating Violence Reported in the Past Year at 2.5-Year Follow-up According to Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Students With PDV, No./Total No. (%)</strong></td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td><strong>Control Group</strong></td>
</tr>
<tr>
<td>All students</td>
</tr>
<tr>
<td>Unadjusted OR$^a$</td>
</tr>
<tr>
<td>Adjusted OR$^b$</td>
</tr>
<tr>
<td>Students who dated in the past 12 mo</td>
</tr>
<tr>
<td>Unadjusted OR$^a$</td>
</tr>
<tr>
<td>Adjusted OR$^b$</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI, confidence interval; ICC, intraclass correlation coefficient; OR, odds ratio; PDV, physical dating violence.

$^a$Odds ratios of the intervention effect from the multilevel model.

$^b$Odds ratios were adjusted for baseline behavior, stratifying variables, and sex (n=1722 in the full model; n=1041 in the model restricted to the dating sample).
Correlation \( (r) \) between treatment and dating violence

- Correlation \( (r) \) between treatment and dating violence

<table>
<thead>
<tr>
<th></th>
<th>Control Group</th>
<th>Intervention Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students With PDV, No./Total No. (%)</td>
<td>74/754 (9.8)</td>
<td>72/968 (7.4)</td>
</tr>
</tbody>
</table>
App. 2 (Online calculator for ES)

- Correlation ($r$) between treatment and dating violence
Correlation ($r$) between treatment and dating violence
- Correlation \((r)\) between treatment and dating violence

How do we go from information reported to a 2 X 2 contingency table??
App. 2 (Online calculator for ES)

- Correlation \((r)\) between treatment and dating violence

Intervention: yes = 72
no = 968 – 72 = 896

Control: yes = 74
no = 754 – 74 = 680
Correlation ($r$) between treatment and dating violence

Transformed Effect size = $Z_r = -0.0423$

(Optional) CI for forest plot or table

$V = SE^2 = 0.0006$ (for weighting)
Computed two different effect sizes ($d$ and $r$) for didactic purposes

- In a real situation, would convert all studies results to a common effect size.

Conclusion:

- Online calculator very helpful
  - Reduces time, coder burden, and probably error
- Must be sure you understand the reported results, and the options of the online calculator
  - Don’t guess