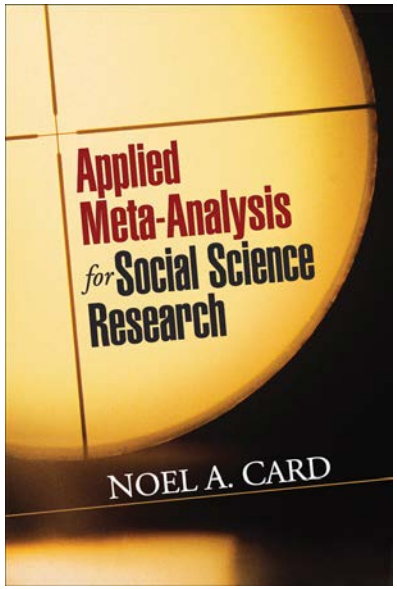
Introduction to Meta-Analysis and Systematic Reviews

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.

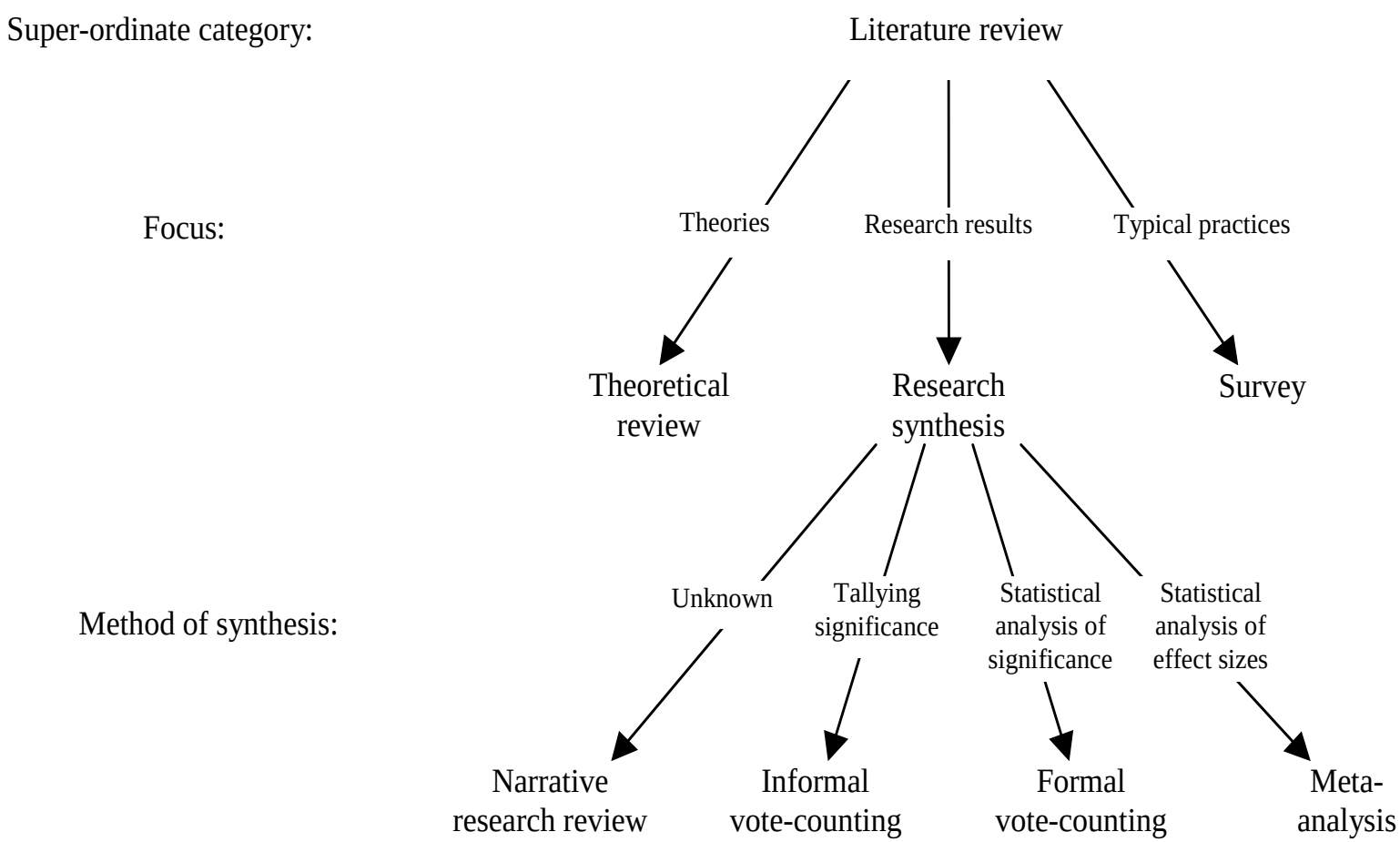
Types of literature review

- 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



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 / listserve requests: Especially for unpublished work.

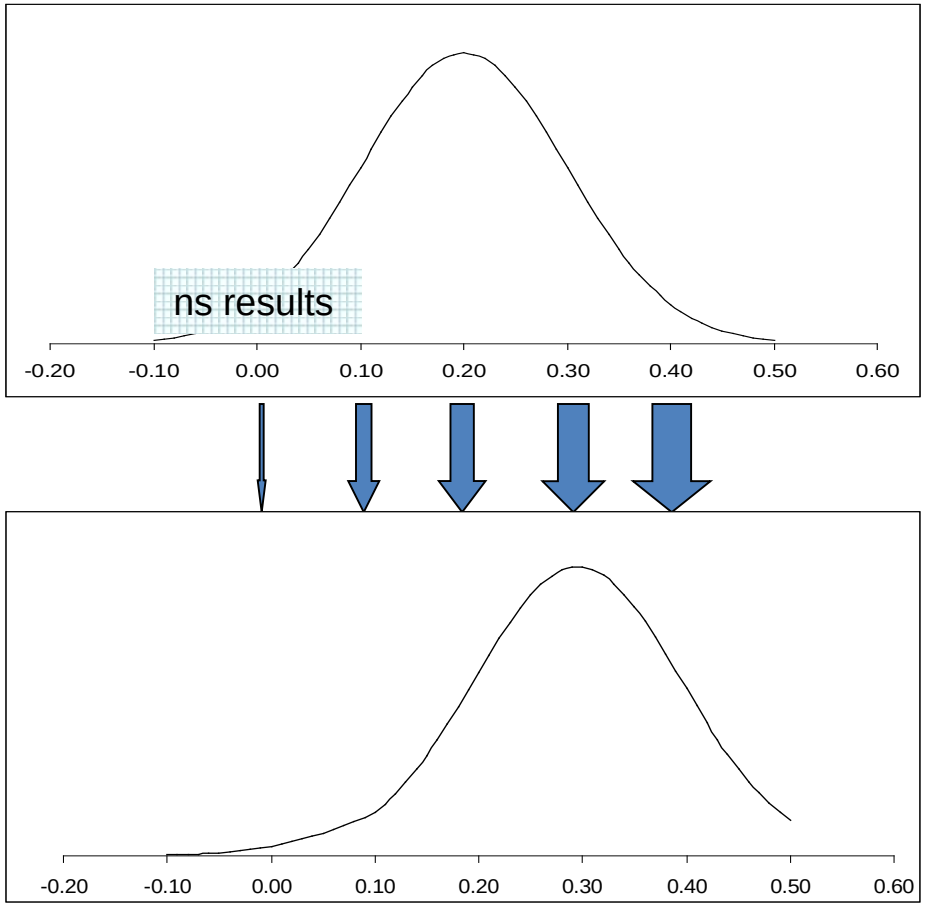
Searching for Literature

- 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 → 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). [CLICK FOR EXAMPLE]

Coding the Literature

- Effect sizes.
  - Effect sizes are not…
    - Author conclusions in Discussion.
    - Significance tests (i.e., *p* values).
  - Effect sizes <u>are…</u>
    - 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.

      [CLICK FOR EXAMPLE]

Table 5.1: Summary of equations for computing effect sizes from results of primary studies.

	Pearson r	Hedges' g	Odds ratio
Definitional formula	$\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{(N-1)s_x s_y}$	$\frac{M_1-M_2}{s_{pooled}}$	$\frac{n_{00}n_{11}}{n_{01}n_{10}}$
Independent t-test with unequal group sizes	$\sqrt{\frac{t^2}{t^2+df}}$	$\frac{t\sqrt{n_1+n_2}}{\sqrt{n_1 n_2}}$	N/A
Independent t-test with equal group sizes	"	$\frac{2t}{\sqrt{N}}$	N/A
Independent F-ratio with unequal group sizes	$\sqrt{\frac{F_{(1,df)}}{F_{(1,df)}+df}}$	$\sqrt{\frac{F_{(1,df)}(n_1+n_2)}{n_1 n_2}}$	N/A
Independent F-ratio with equal group sizes	"	$2\sqrt{\frac{F_{(1,df)}}{N}}$	N/A
Dependent (repeated-measures) t-test	$\sqrt{\frac{t^2_{dependent}}{t^2_{dependent}+df}}$	$\frac{t_{dependent}}{\sqrt{N}}$	N/A
Dependent (repeated-measures) F-ratio	$\sqrt{\frac{F_{repeated(1,df)}}{F_{repeated(1,df)}+df}}$	$\sqrt{\frac{F_{repeated(1,df)}}{N}}$	N/A
2 X 2 (i.e., 1 df) contingency χ^2	$\sqrt{\frac{\chi^2_{(1)}}{N}}$	$2\sqrt{\frac{\chi^2_{(1)}}{N-\chi^2_{(1)}}}$	Reconstru contingen table

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.

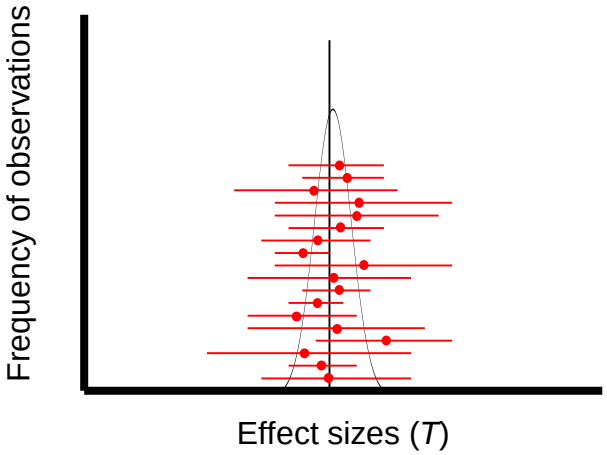
Synthesizing Effect Sizes

- 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…

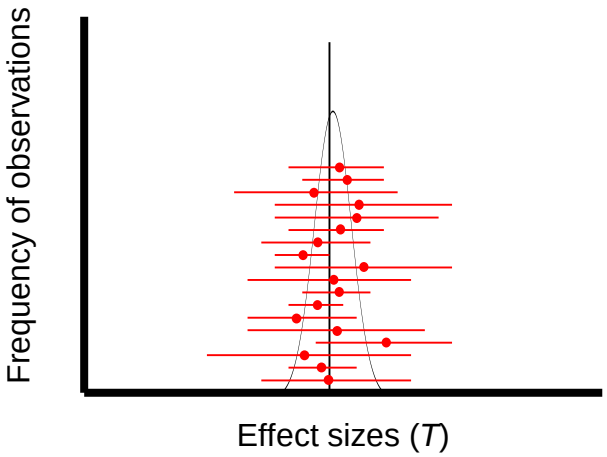
Sampling error only



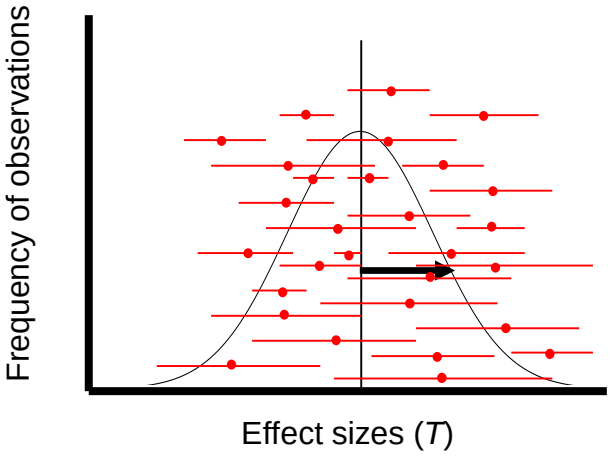
Synthesizing Effect Sizes

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

Sampling error only



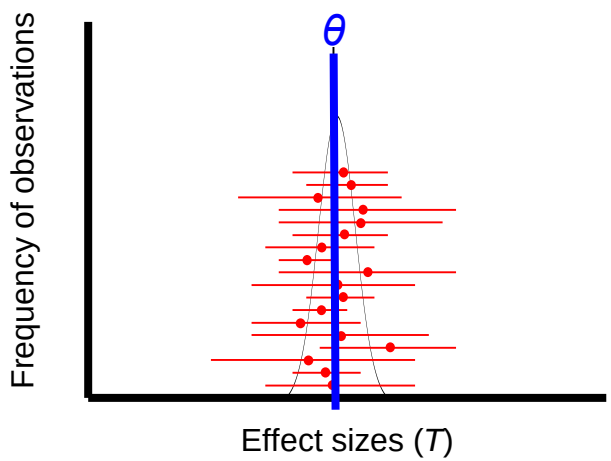
Sampling error + other variability



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

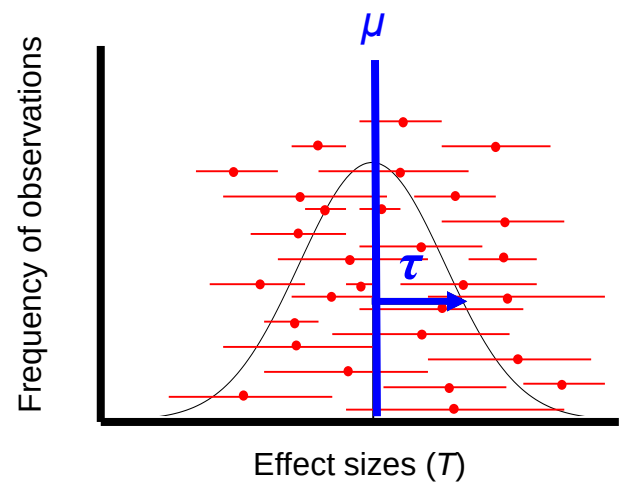
Fixed-effects model



$$T_i = \theta + \varepsilon_i$$

where ε_i is the sampling error of T_i

Random-effects model

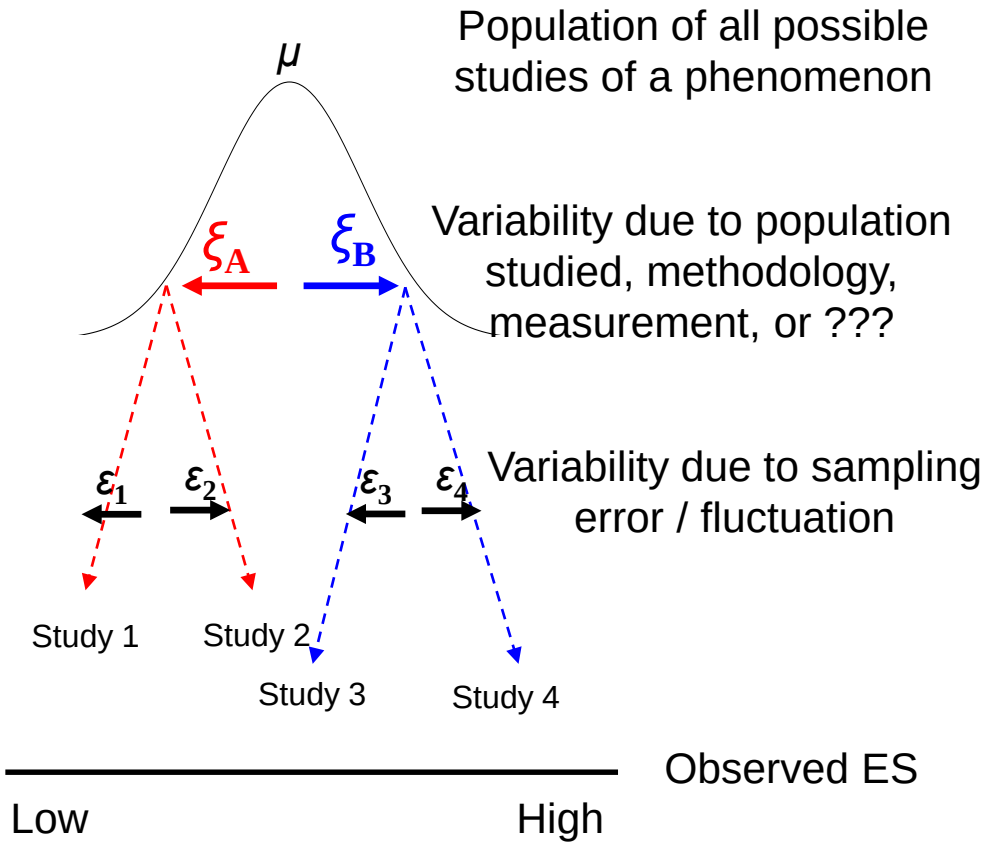


$$T_i = \mu + \xi_i + \varepsilon_i$$

where θ_i is the population effect size for each study, ξ_i is the study-deviation from μ, and ε_i is the sample-sampling error

Synthesizing Effect Sizes

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



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*.
- τ^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 <u>why</u> 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:

$$Q_{Total} = Q_{Between} + Q_{Within}$$

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

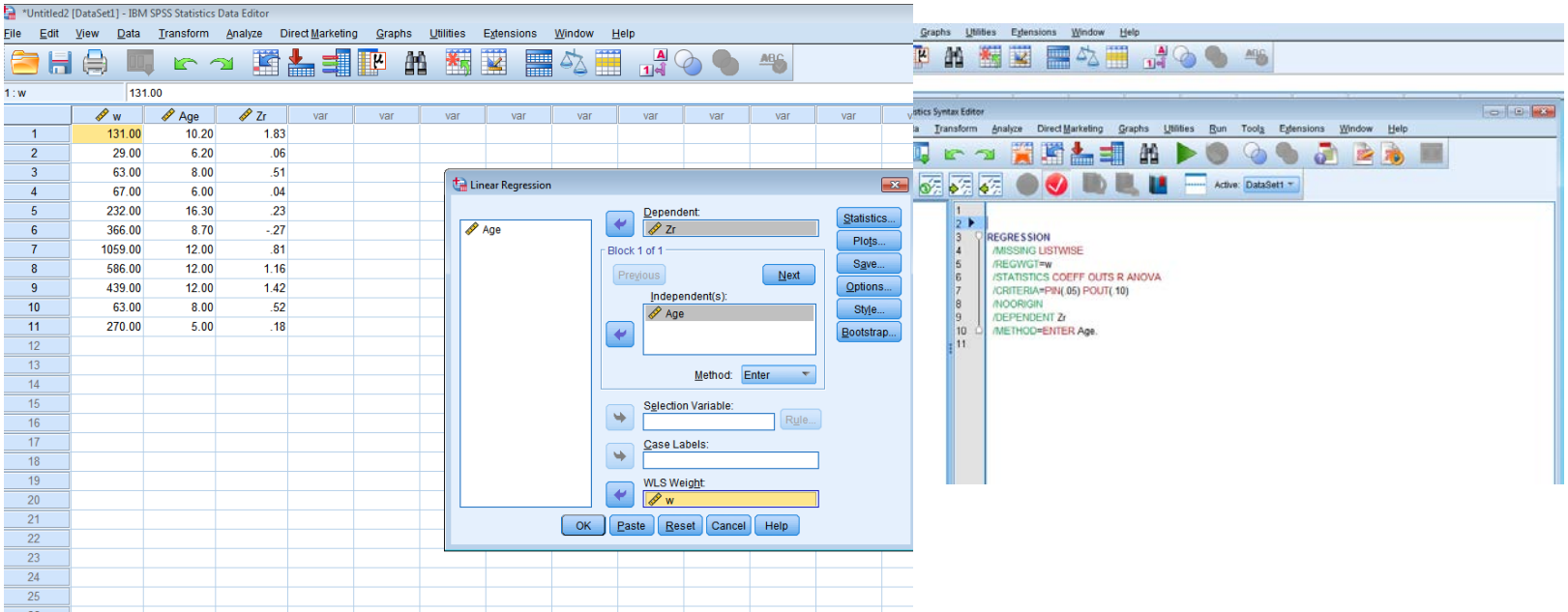
$$Q_{Btwn} = \Sigma\,[w_j(M_j - GM)^2] = (\Sigma w_j M_j^2) - \frac{(\Sigma w_j M_j)^2}{\Sigma w_j}$$

$$Q_{within} = \Sigma^J \Sigma^I\, w_i(ES_i - M_j)^2 = Q_1 + Q_2 + \ldots + Q_J$$

    - Significance indicated by Q_{Between}

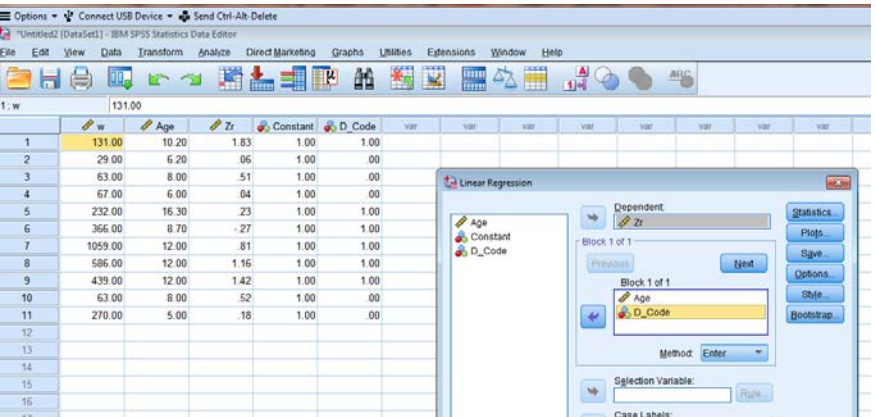
Synthesizing Effect Sizes

- 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.,:
    - Card (2012). *Applied Meta-Analysis for Social Science Research*. Guilford.
    - Lipsey & Wilson (2001). *Practical Meta-Analysis*. Sage.
    - Oriented to CMA software: Borenstein et al. (2009). *Introduction to Meta-Analysis*. Wiley.

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

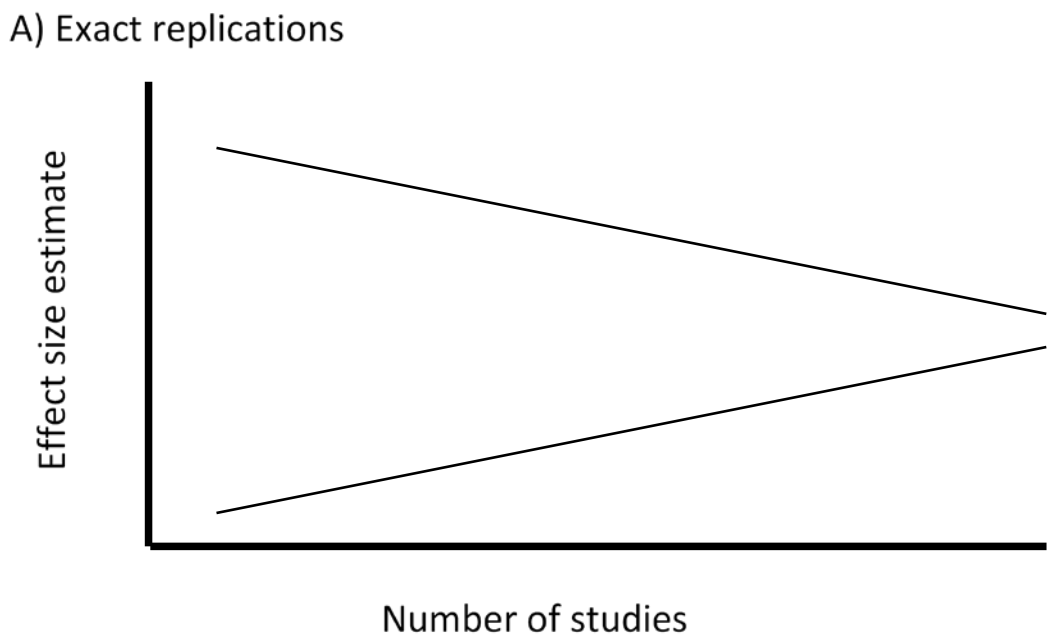
- Integrative Data Analysis (IDA) = Combination and analysis of *raw data* from multiple data sets
- Advantages of IDA:
  - Allows analysis of within- and between-study variability. (E.g., analysis of individual-level demographic information rather than aggregating to the study level)
  - Explicit attention 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):

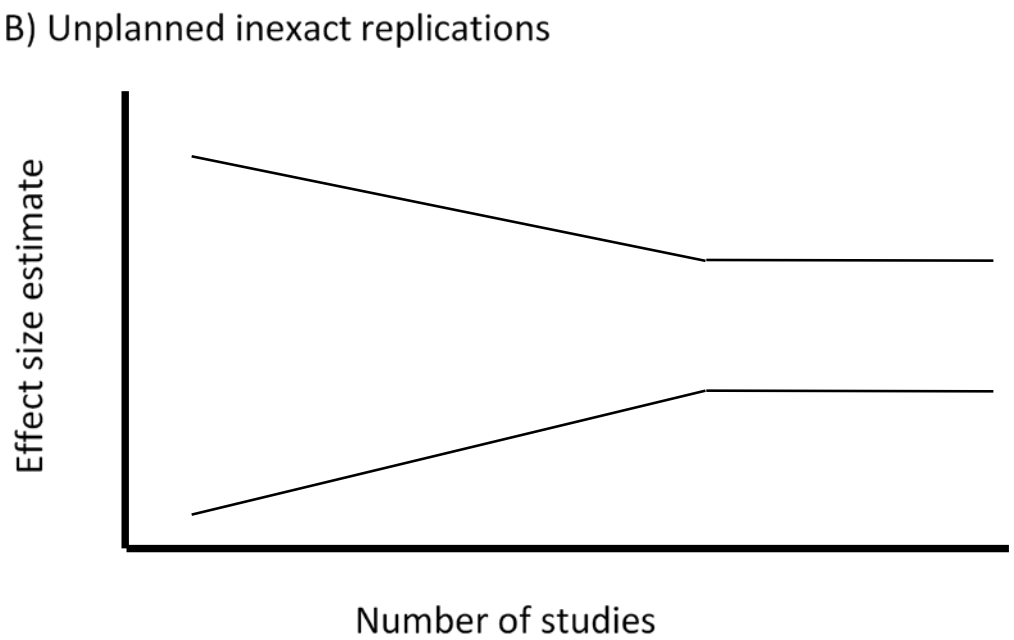
MA and Other Approaches

- 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):



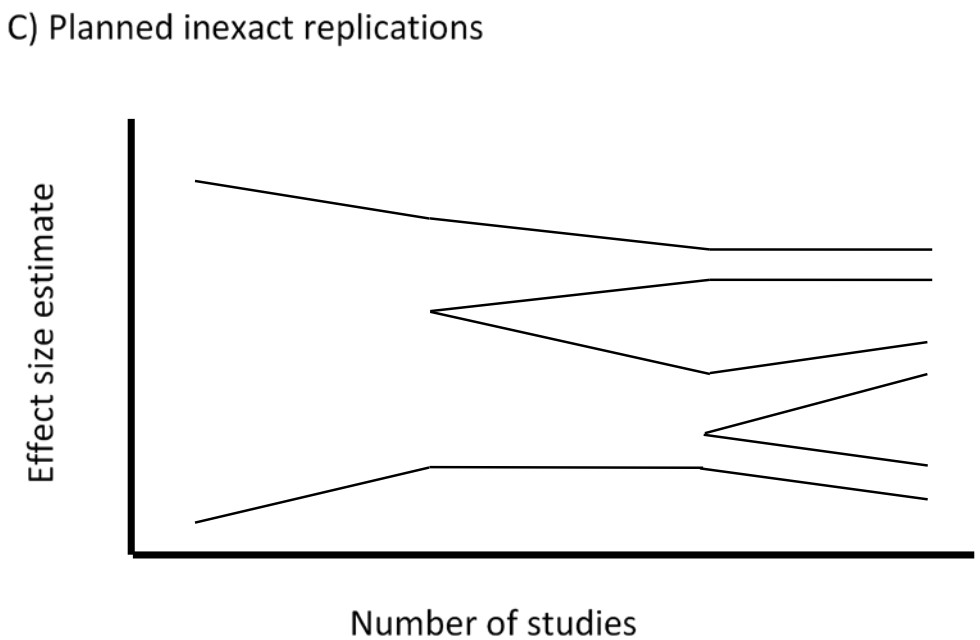
MA and Other Approaches

- 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):



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.
    - Card (2012). *Applied Meta-Analysis for Social Science Research*. Guilford.
    - Lipsey & Wilson (2001). *Practical Meta-Analysis*. Sage.
    - Oriented to CMA software: Borenstein et al. (2009). *Introduction to Meta-Analysis*. Wiley.

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)

	Control	Experimental
	S1 = 4	S2 = 6
	S3 = 4	S4 = 6
	S5 = 3	S6 = 5
	S7 = 2	S8 = 4
	S9 = 2	S10 = 4
N	5	5
Sum	15	25
Mean	3.00	5.00
S	1.00	1.00

$$t(8) = \left(\frac{M1 - M2}{S_{within}}\right)\left(\frac{\sqrt{N}}{2}\right) = \mathbf{3.16,} \quad \boldsymbol{p = .013}$$

App. 1 (effect sizes are not significance tests)

	Control	Experimental
	S1 = 4	S2 = 6
	S3 = 4	S4 = 6
	S5 = 3	S6 = 5
	S7 = 2	S8 = 4
	S9 = 2	S10 = 4
N	5	5
Sum	15	25
Mean	3.00	5.00
S	1.00	1.00

$$t(8) = \left(\frac{M1 - M2}{S_{within}}\right)\left(\frac{\sqrt{N}}{2}\right) = \mathbf{3.16,\quad \textit{p} = .013}$$

	Control	Experimental
	S1 = 4	S2 = 6
	S3 = 3	S4 = 5
	S5 = 2	S6 = 4
N	3	3
Sum	9	15
Mean	3.00	5.00
S	1.00	1.00

$$t(4) = \left(\frac{M1 - M2}{S_{within}}\right)\left(\frac{\sqrt{N}}{2}\right) = \mathbf{2.45,\quad \textit{p} = .071}$$

App. 1 (effect sizes are not significance tests)

	Control	Experimental
	S1 = 4	S2 = 6
	S3 = 4	S4 = 6
	S5 = 3	S6 = 5
	S7 = 2	S8 = 4
	S9 = 2	S10 = 4
N	5	5
Sum	15	25
Mean	3.00	5.00
S	1.00	1.00

$$t(8) = \left(\frac{M1 - M2}{S_{within}}\right)\left(\frac{\sqrt{N}}{2}\right) = \mathbf{3.16,\quad \textit{p} = .013}$$

	Control	Experimental
	S1 = 4	S2 = 6
	S3 = 4	S4 = 6
	S5 = 3	S6 = 5
	S7 = 2	S8 = 4
	S9 = 2	S10 = 4
	S11 = 4	S12 = 6
	S13 = 4	S14 = 6
	S15 = 3	S16 = 5
	S17 = 2	S18 = 4
	S19 = 2	S20 = 4
N	10	10
Sum	30	50
Mean	3.00	5.00
S	1.00	1.00

$$t = \left(\frac{M1 - M2}{S_{within}}\right)\left(\frac{\sqrt{N}}{2}\right) = \mathbf{4.74,\quad p = .00016}$$

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.

App. 1 (effect sizes are not significance tests)

[Return from example]

- 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:

test of significance	=	effect size	X	size of study
$t_{(independent)}$	=	d	X	$\frac{\sqrt{df}}{2}$
t	=	$\frac{r}{\sqrt{(1 - r^2)}}$	X	$\sqrt{df}$
$\chi^2(1)$	=	Φ^2	X	N

App. 2 (Online calculator for ES)

- Calculate effect sizes from two studies
- Use online calculator at
  http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

App. 2 (Online calculator for ES)

- Two studies from a review of the effectiveness of teen dating violence prevention programs in K-12 schools
  - Adler-Baeder, F., et al. (2007). The impact of relationship education on adolescents of diverse backgrounds. *Family Relations, 56*, 291-303.
  - Wolfe, D.A., et al. (2009). A school-based program to prevent adolescent dating violence. *Archives of Pediatric Medicine, 163*, 692-699.
- 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.

App. 2 (Online calculator for ES)

- Standardized mean difference
  - Adler-Baeder, F., et al. (2007). The impact of relationship education on adolescents of diverse backgrounds. *Family Relations, 56*, 291-303.
    - 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”

298 *Family Relations • Volume 56, Number 3 • July 2007*

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

	Experimental Group				Control Group			
	Pretest		Posttest		Pretest		Posttest	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Conflict tactics subscales								
Reasoning	2.66	1.26	2.73	1.32	2.77	1.55	2.90	1.27
Verbal	1.91	1.45	1.75	1.42	2.05	1.52	2.29	1.61
Physical	.76	1.21	.82	1.33	.99	1.42	1.17	1.67
Relationship belief subscales								
Aggression beliefs	3.37	.83	3.37	.86	3.19	.99	3.26	.94
Faulty relationship beliefs	2.96	.53	3.03	.58	2.97	.50	3.02	.52
Realistic relationship beliefs	2.89	.47	3.01	.52	2.89	.50	2.85	.58

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

Effect Size Calculator

www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

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

- HOME
- EFFECT SIZE TYPE
- + Standardized Mean Difference (d)
- + Correlation Coefficient (r)
- + Odds-ratio (OR) and Risk Ratio (RR)
- FORMULAS

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.

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 Calculator

www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD-main.php

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

- HOME
- EFFECT SIZE TYPE
- + Standardized Mean Difference (d)
  - Means and standard deviations
  - t-test, unequal sample sizes
  - t-test, equal sample sizes
  - F-test, 2-group, unequal sample sizes
  - F-test, 2-group, equal sample sizes
  - t-test p-value, equal sample sizes
  - t-test p-value, unequal sample sizes
  - Means and standard errors
  - 2 by 2 frequency table
  - Binary proportions
  - Point-biserial correlation, equal Ns
  - Point-biserial correlation, unequal Ns
  - Point-biserial correlation p-value, equal Ns
  - Point-biserial correlation p-value, unequal Ns
  - Phi-coefficient
  - Phi-coefficient p-value
  - Chi-square
  - Chi-square p-value
  - Frequency distribution
  - Frequency distribution (proportions)
  - Unstandardized regression coefficient
  - Standardized regression coefficient
  - Means and full sample standard deviation
  - Mean gains scores and gain score SDs
  - Mean gain scores, pre and post SDs, and paired t-tests
  - Mean gain scores, pre and post SDs, and pre-post r
  - Means and standard deviations with subgroups
  - F-test, 3 or more groups
  - Means and ANCOVA
  - Two-way ANOVA
- + Correlation Coefficient (r)
- + Odds-ratio (OR) and Risk Ratio (RR)
- FORMULAS

The standardized mean-difference effect size (*d*) is designed for contrasting two groups on a continuous dependent variable. It can be computed from means and standard deviations, a t-test, and a one-way ANOVA. Methods have also be developed for estimating *d* based on a dichotomous dependent variable.

From the menu, select the type of data available for computing the effect size.

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 Calculator

www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD1.php

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

- HOME
- EFFECT SIZE TYPE
- + Standardized Mean Difference (d)
  - Means and standard deviations
  - t-test, unequal sample sizes
  - t-test, equal sample sizes
  - F-test, 2-group, unequal sample sizes
  - F-test, 2-group, equal sample sizes
  - t-test p-value, equal sample sizes
  - t-test p-value, unequal sample sizes
  - Means and standard errors
  - 2 by 2 frequency table
  - Binary proportions
  - Point-biserial correlation, equal Ns
  - Point-biserial correlation, unequal Ns
  - Point-biserial correlation p-value, equal Ns
  - Point-biserial correlation p-value, unequal Ns
  - Phi-coefficient
  - Phi-coefficient p-value
  - Chi-square
  - Chi-square p-value
  - Frequency distribution
  - Frequency distribution (proportions)
  - Unstandardized regression coefficient
  - Standardized regression coefficient
  - Means and full sample standard deviation
  - Mean gains scores and gain score SDs
  - Mean gain scores, pre and post SDs, and paired t-tests
  - Mean gain scores, pre and post SDs, and pre-post r
  - Means and standard deviations with subgroups
  - F-test, 3 or more groups
  - Means and ANCOVA
  - Two-way ANOVA
- + Correlation Coefficient (r)
- + Odds-ratio (OR) and Risk Ratio (RR)
- FORMULAS

Means, Standard Deviations, and Sample Sizes

	Mean	SD	N
Treatment			
Control			

Calculate Reset

d =

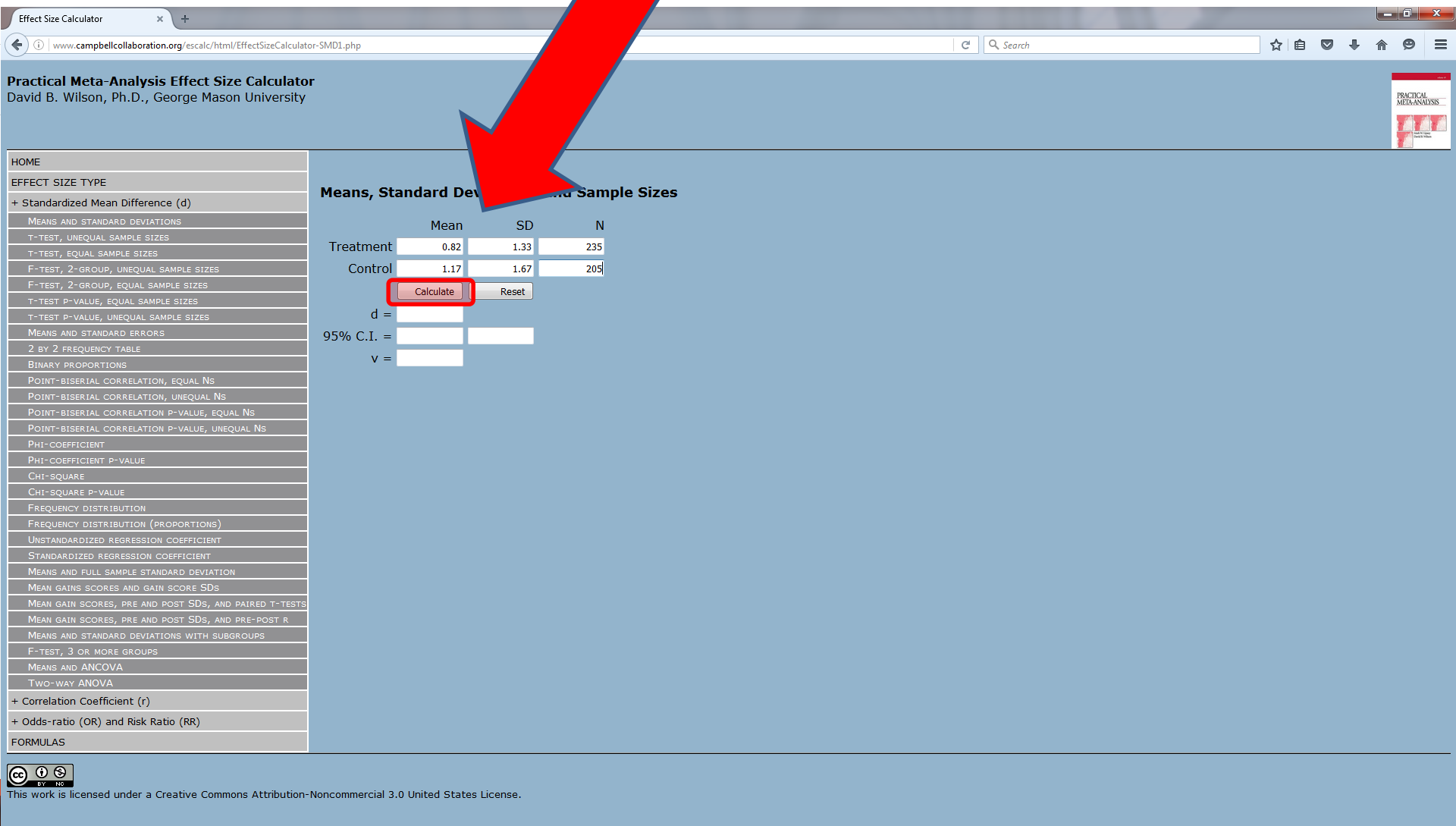
95% C.I. =

v =

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 Calculator

www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD1.php

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

- HOME
- EFFECT SIZE TYPE
- + Standardized Mean Difference (d)
  - Means and standard deviations
  - t-test, unequal sample sizes
  - t-test, equal sample sizes
  - F-test, 2-group, unequal sample sizes
  - F-test, 2-group, equal sample sizes
  - t-test p-value, equal sample sizes
  - t-test p-value, unequal sample sizes
  - Means and standard errors
  - 2 by 2 frequency table
  - Binary proportions
  - Point-biserial correlation, equal Ns
  - Point-biserial correlation, unequal Ns
  - Point-biserial correlation p-value, equal Ns
  - Point-biserial correlation p-value, unequal Ns
  - Phi-coefficient
  - Phi-coefficient p-value
  - Chi-square
  - Chi-square p-value
  - Frequency distribution
  - Frequency distribution (proportions)
  - Unstandardized regression coefficient
  - Standardized regression coefficient
  - Means and full sample standard deviation
  - Mean gains scores and gain score SDs
  - Mean gain scores, pre and post SDs, and paired t-tests
  - Mean gain scores, pre and post SDs, and pre-post r
  - Means and standard deviations with subgroups
  - F-test, 3 or more groups
  - Means and ANCOVA
  - Two-way ANOVA
- + Correlation Coefficient (r)
- + Odds-ratio (OR) and Risk Ratio (RR)
- FORMULAS

Means, Standard Deviations, and Sample Sizes

	Mean	SD	N
Treatment	0.82	1.33	235
Control	1.17	1.67	205

Calculate | Reset

d =	-0.2336	
95% C.I. =	-0.4216	-0.0457
v =	0.0092	

Effect size = *d* = -0.2336

(optional) CI for forest plot or table

V = SE^2 = 0.0092 (for weighting)

[Return from example]

App. 2 (Online calculator for ES)

- Second Example:
- Correlation (*r*) between treatment and dating violence
  - Wolfe, D.A., et al. (2009). A school-based program to prevent adolescent dating violence. *Archives of Pediatric Medicine, 163*, 692-699.

App. 2 (Online calculator for ES)

- Correlation (*r*) between treatment and dating violence
  - Wolfe, D.A., et al. (2009). A school-based program to prevent adolescent dating violence. *Archives of Pediatric Medicine, 163*, 692-699.

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

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

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
  - Wolfe, D.A., et al. (2009). A school-based program to prevent adolescent dating violence. *Archives of Pediatric Medicine, 163*, 692-699.

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

	Students With PDV, No./Total No. (%)					
	Control Group	Intervention Group	ICC	OR (95% CI)	*t* Test	*P* Value
All students	74/754 (9.8)	72/968 (7.4)	0.02			
Unadjusted OR[a]				1.42 (0.87-2.33)	1.49	.15
Adjusted OR[b]				2.42 (1.00-6.02)	2.06	.05
Students who dated in the past 12 mo	[illegible]	[illegible]	[illegible]			
Unadjusted OR[a]				1.37 (0.89-2.13)	1.53	.14
Adjusted OR[b]				2.13 (0.81-5.66)	1.65	.12

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
  - Wolfe, D.A., et al. (2009). A school-based program to prevent adolescent dating violence. *Archives of Pediatric Medicine, 163*, 692-699.

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

	Students With PDV, No./Total No. (%)					
	Control Group	Intervention Group	ICC	OR (95% CI)	*t* Test	*P* Value
All students	74/754 (9.8)	72/968 (7.4)	0.02			
Unadjusted OR[a]				1.42 (0.87-2.33)	1.49	.15
Adjusted OR[b]				2.42 (1.00-6.02)	2.06	.05
Students who dated in the past 12 mo	[illegible]	[illegible]	[illegible]			
Unadjusted OR[a]				1.37 (0.89-2.13)	1.53	.14
Adjusted OR[b]				2.13 (0.81-5.66)	1.65	.12

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
  - Wolfe, D.A., et al. (2009). A school-based program to prevent adolescent dating violence. *Archives of Pediatric Medicine, 163*, 692-699.

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

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

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

ce Reported in the Past Year at 2.5-Year Follow

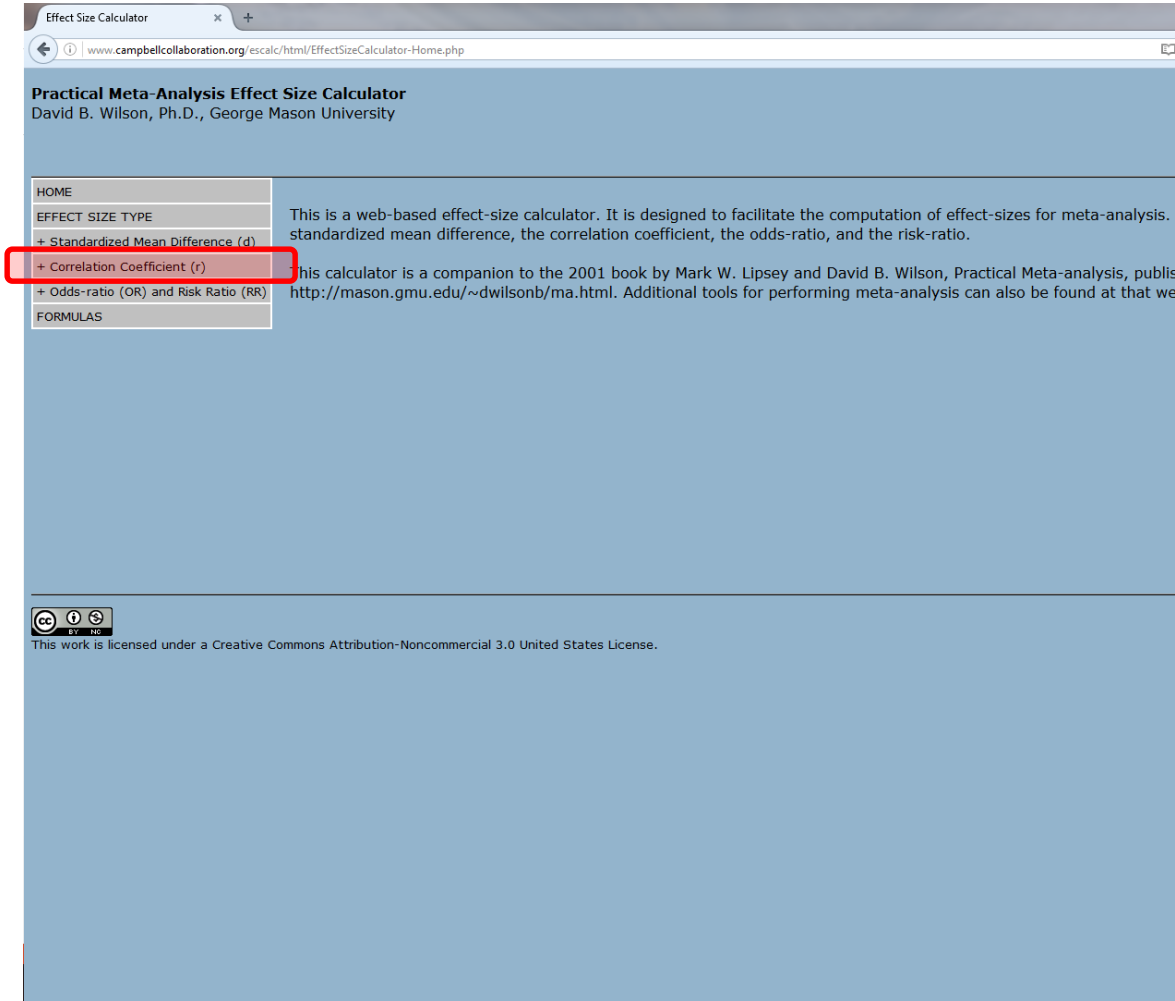
Students With PDV, No./Total No. (%)	
Control Group	Intervention Group
74/754 (9.8)	72/968 (7.4)

App. 2 (Online calculator for ES)

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

ce Reported in the Past Year at 2.5-Year Follow

Students With PDV, No./Total No. (%)	
Control Group	Intervention Group
74/754 (9.8)	72/968 (7.4)

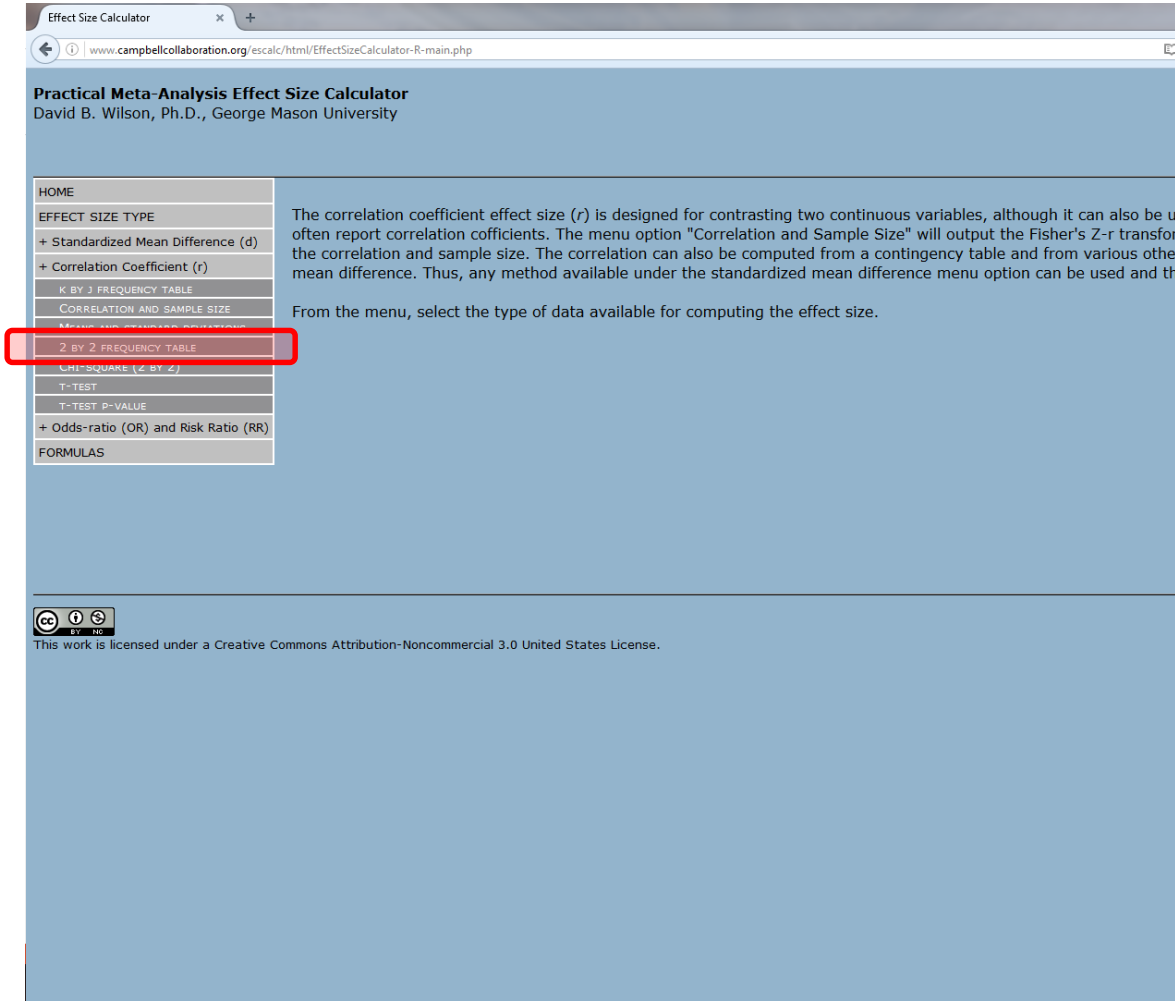


App. 2 (Online calculator for ES)

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

ce Reported in the Past Year at 2.5-Year Follow

	Students With PDV, No./Total No. (%)	
	Control Group	Intervention Group
	74/754 (9.8)	72/968 (7.4)



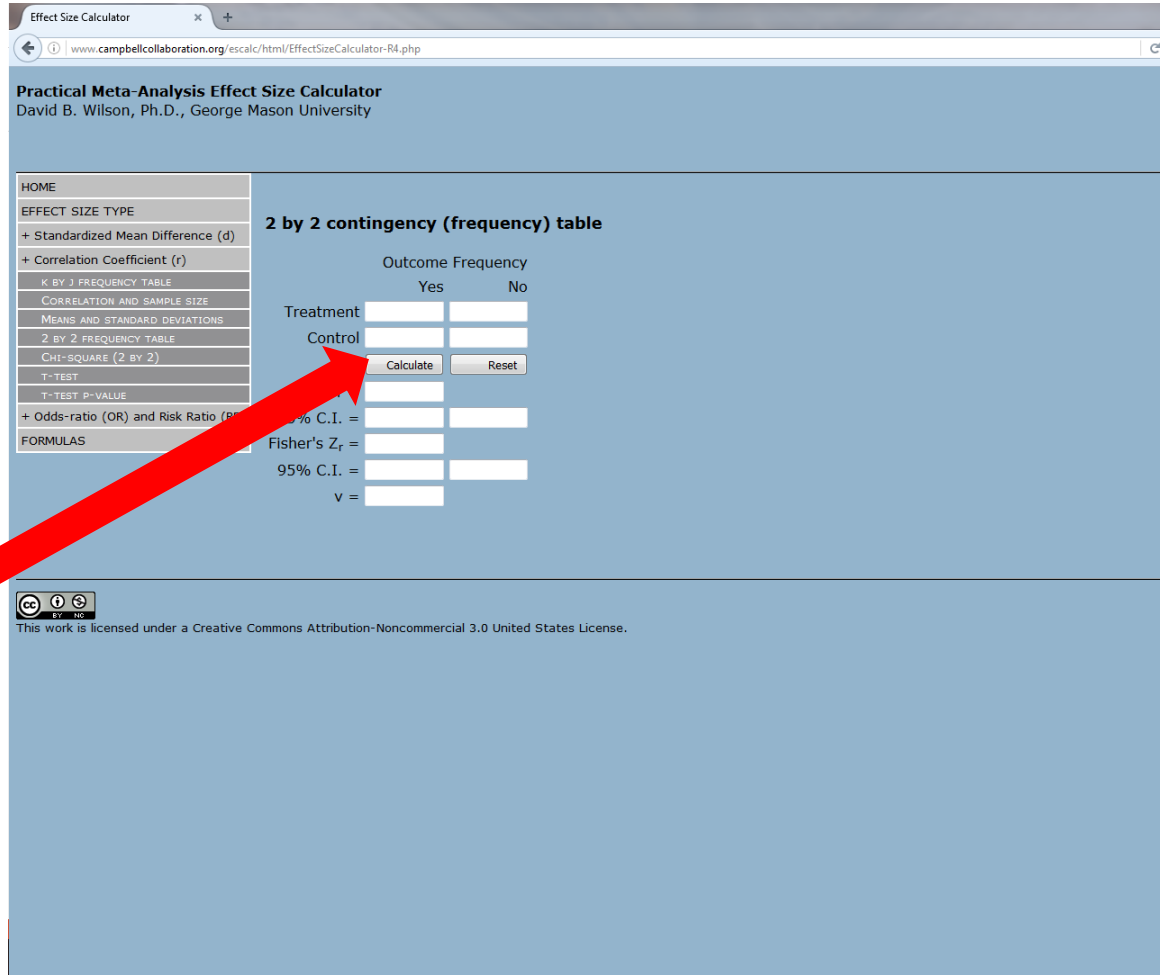
App. 2 (Online calculator for ES)

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

ce Reported in the Past Year at 2.5-Year Follow

	Students With PDV, No./Total No. (%)	
	Control Group	Intervention Group
	74/754 (9.8)	72/968 (7.4)

How do we go from information reported to a 2 X 2 contingency table??



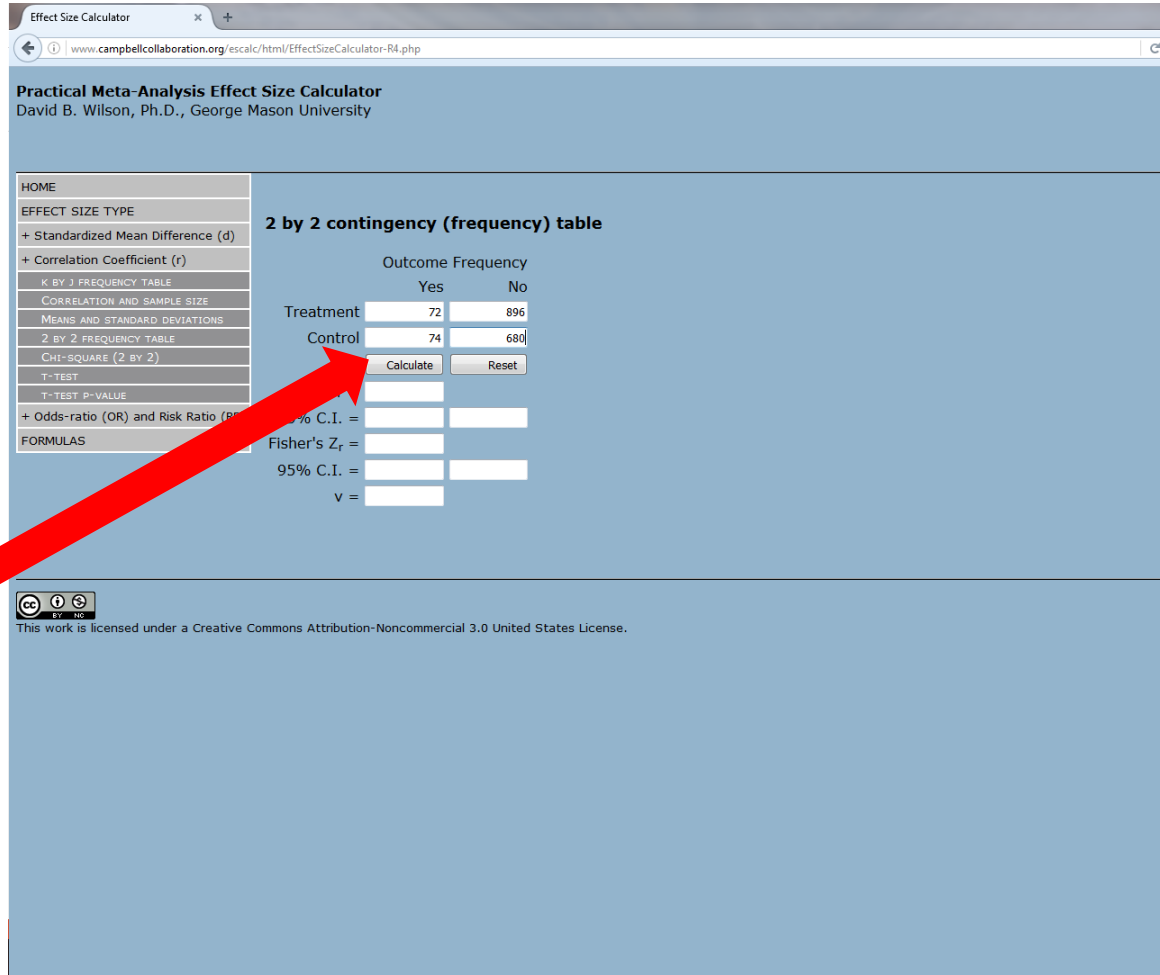
App. 2 (Online calculator for ES)

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

ce Reported in the Past Year at 2.5-Year Follow

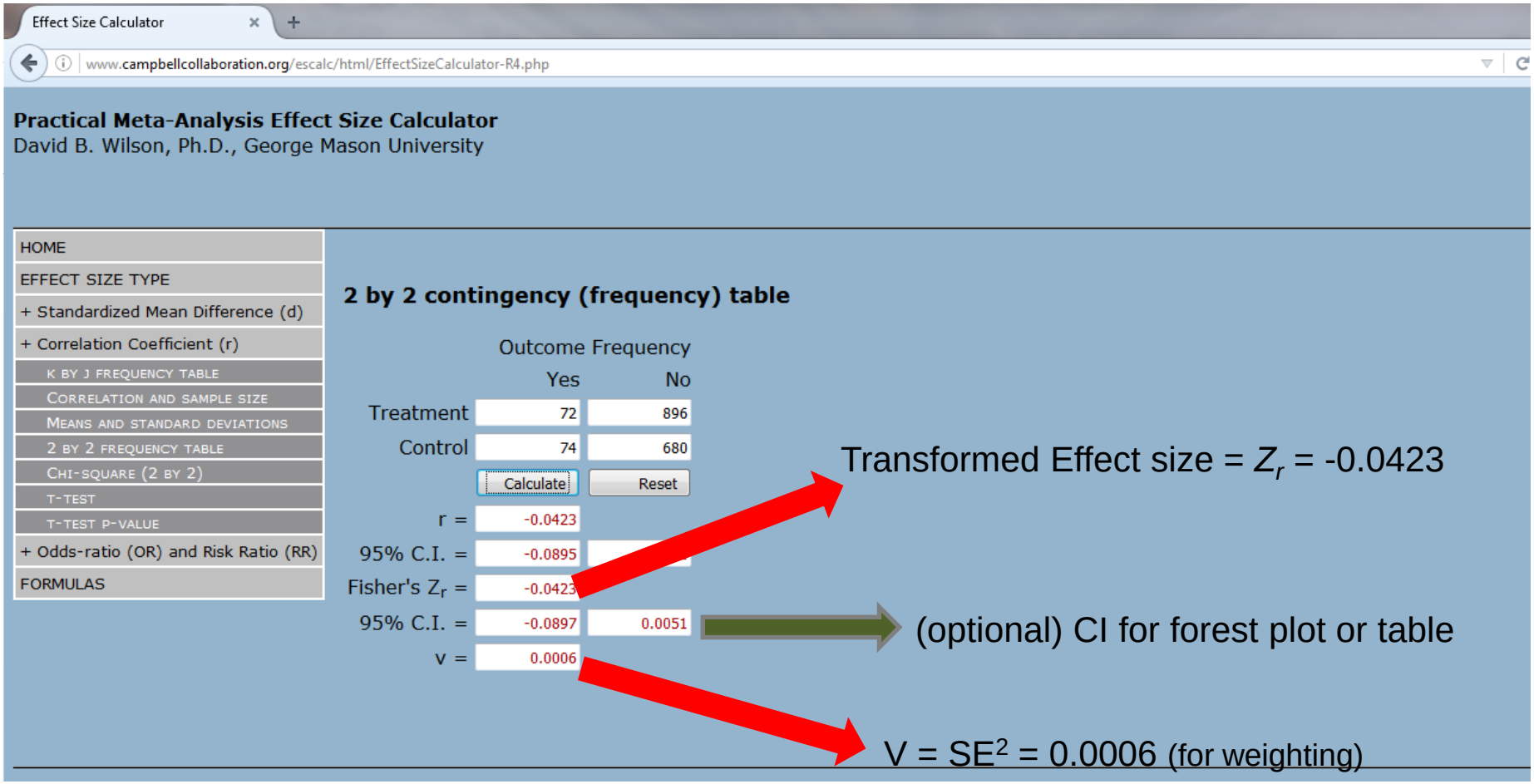
Students With PDV, No./Total No. (%)	
Control Group	Intervention Group
74/754 (9.8)	72/968 (7.4)

Intervention: yes = 72
no = 968 – 72 = 896
Control: yes = 74
no = 754 – 74 = 680



App. 2 (Online calculator for ES)

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



App. 2 (Online calculator for ES)

- 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

[Return from example]