A Spoonful of Medicine for Meta-Analysis: Introducing Meta-Sen

James G. Field
West Virginia University

Frank A. Bosco, Sven Kepes
Virginia Commonwealth University

SMA PDI | October 11, 2019
Norfolk, VA

All submission materials (e.g., original submission, PowerPoint slides, sample data sets) can be found at:
https://jamiefield.github.io/research/sma2019

Visit https://jamiefield.github.io/research/sma2019 for submission materials

Agenda

- What is meta-analysis and why is so important?
- The current research environment
- Threats to our cumulative scientific knowledge: outliers and publication bias
- What is sensitivity analysis?
- Live demonstration of Meta-Sen
- Recommendations for minimizing the impact of outliers and/or PB
- Discussion/questions/comments from the audience
- Additional slides
  - Review of two outlier assessment methods
  - Review of five publication bias assessment methods
  - Results that illustrate the combined effect of outliers and PB on recently published meta-analytic datasets

What is Meta-Analysis?

- Meta-analysis is a statistical technique by which information from independent studies is assimilated
  - Field, A. P. (2011)

- Meta-analysis is a quantitative method used to combine the quantitative outcomes (effect sizes) of primary research studies.
  - Combines the results from two or more studies
  - Estimates an ‘average’ effect between two constructs

- Meta-analysis is the statistical or data analytic part of a systematic review of a research topic.

What is Meta-Analysis?

- There are two common approaches to meta-analysis
  - The Hunter and Schmidt (2004; 2015) approach, which is most common in organizational research
  - The Hedges and Olkin (1985) approach
- For a description of both approaches, and their differences, please refer to Kepes et al. (2013)
- Important note:
  - We use the Hedges and Olkin (1985) approach as most sensitivity analysis techniques have not been developed for psychometrically-adjusted effect sizes

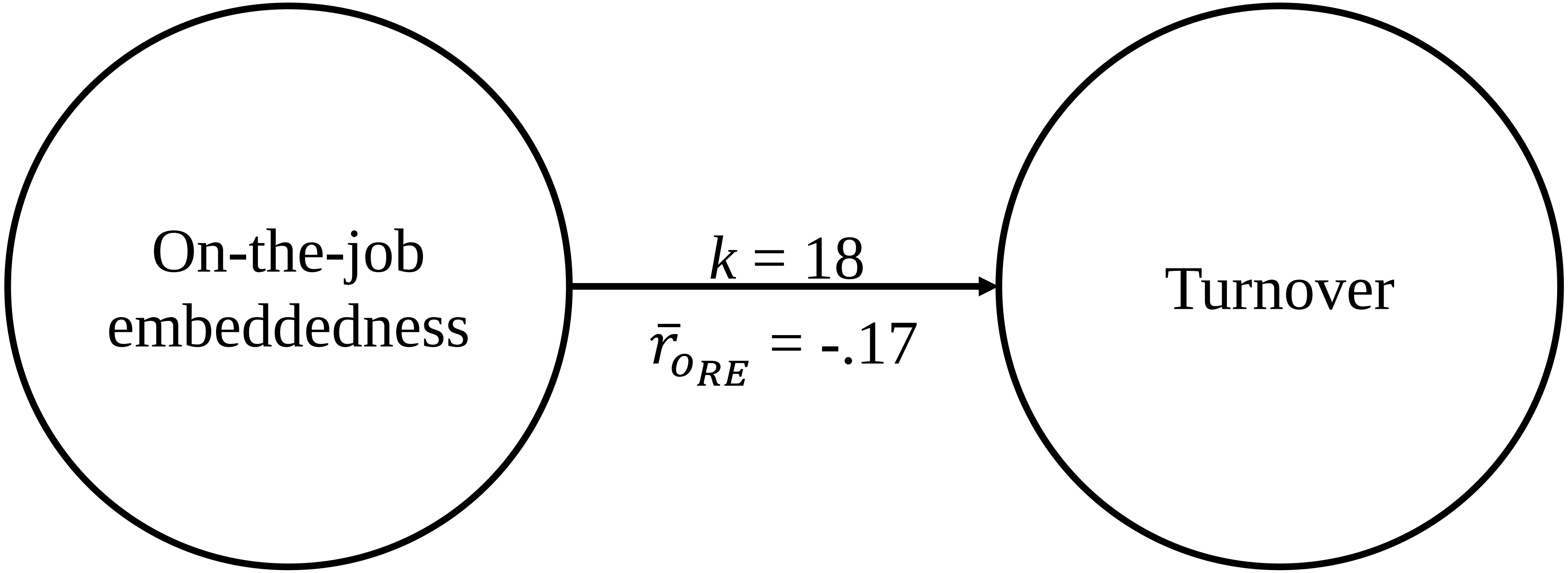
What is Meta-Analysis?

- An example from the published literature

Study ID	Reference	Year	IV	DV	*n*	*r*
1	Mitchell, Holtom, Lee, & Erez	2001	On-the-job embeddedness	Turnover	177	-0.24
2	Crossley, Bennett, Jex, & Burnfield	2007	On-the-job embeddedness	Turnover	306	-0.08
3	Mitchell, Holtom, Lee, & Erez	2001	On-the-job embeddedness	Turnover	208	-0.21
4	Ramesh & Gelfand	2010	On-the-job embeddedness	Turnover	474	-0.13
5	Giosan, Holtom, & Watson	2005	On-the-job embeddedness	Turnover	122	-0.30
6	Mallol, Holtom, & Lee	2007	On-the-job embeddedness	Turnover	164	-0.16
7	Lee, Mitchell, & Holtom	2004	On-the-job embeddedness	Turnover	809	-0.11
8	Kraimer, Shaffer, Harrison, & Ren	2012	On-the-job embeddedness	Turnover	112	-0.17
9	Ramesh & Gelfand	2010	On-the-job embeddedness	Turnover	323	-0.14
10	Harris, Wheeler, & Kacmar	2011	On-the-job embeddedness	Turnover	205	-0.19
11	Allen	2006	On-the-job embeddedness	Turnover	222	-0.23
12	Wheeler, Halbesleben, & Sablynski	2011	On-the-job embeddedness	Turnover	142	-0.26
13	Tanova & Holtom	2008	On-the-job embeddedness	Turnover	9277	-0.08
14	Wheeler, Halbesleben, & Sablynski	2011	On-the-job embeddedness	Turnover	134	-0.19
15	Mallol, Holtom, & Lee	2007	On-the-job embeddedness	Turnover	164	-0.13
16	Tharenou & Caulfield	2010	On-the-job embeddedness	Turnover	546	-0.18
17	Smith, Holtom, & Mitchell	2011	On-the-job embeddedness	Turnover	750	-0.25
18	Smith, Holtom, & Mitchell	2011	On-the-job embeddedness	Turnover	1,089	-0.19

Study ID	Reference	Year	IV	DV	*n*	*r*	sei
1	Mitchell, Holtom, Lee, & Erez	2001	On-the-job embeddedness	Turnover	177	-0.24	0.0758
2	Crossley, Bennett, Jex, & Burnfield	2007	On-the-job embeddedness	Turnover	306	-0.08	0.0574
3	Mitchell, Holtom, Lee, & Erez	2001	On-the-job embeddedness	Turnover	208	-0.21	0.0698
4	Ramesh & Gelfand	2010	On-the-job embeddedness	Turnover	474	-0.13	0.0461
5	Giosan, Holtom, & Watson	2005	On-the-job embeddedness	Turnover	122	-0.3	0.0917
6	Mallol, Holtom, & Lee	2007	On-the-job embeddedness	Turnover	164	-0.16	0.0788
7	Lee, Mitchell, & Holtom	2004	On-the-job embeddedness	Turnover	809	-0.11	0.0352
8	Kraimer, Shaffer, Harrison, & Ren	2012	On-the-job embeddedness	Turnover	112	-0.17	0.0958
9	Ramesh & Gelfand	2010	On-the-job embeddedness	Turnover	323	-0.14	0.0559
10	Harris, Wheeler, & Kacmar	2011	On-the-job embeddedness	Turnover	205	-0.19	0.0704
11	Allen	2006	On-the-job embeddedness	Turnover	222	-0.23	0.0676
12	Wheeler, Halbesleben, & Sablynski	2011	On-the-job embeddedness	Turnover	142	-0.26	0.0848
13	Tanova & Holtom	2008	On-the-job embeddedness	Turnover	9277	-0.08	0.0104
14	Wheeler, Halbesleben, & Sablynski	2011	On-the-job embeddedness	Turnover	134	-0.19	0.0874
15	Mallol, Holtom, & Lee	2007	On-the-job embeddedness	Turnover	164	-0.13	0.0788
16	Tharenou & Caulfield	2010	On-the-job embeddedness	Turnover	546	-0.18	0.0429
17	Smith, Holtom, & Mitchell	2011	On-the-job embeddedness	Turnover	750	-0.25	0.0366
18	Smith, Holtom, & Mitchell	2011	On-the-job embeddedness	Turnover	1,089	-0.19	0.0303

What is Meta-Analysis?



Why are Meta-Analyses so Important?

- Meta-analytic reviews are a primary way to summarize, integrate, and synthesize areas of research
  - Schmidt & Hunter (2015)
- Allows fields to build a cumulative scientific knowledge
  - Kepes & McDaniel (2015)
- Meta-analytic results serve as input for other analytic techniques that allow researchers to test theory
  - E.g., relative importance analysis; meta-analytic structural equation modeling
- Meta-analytic results often are used to inform evidence-based management
  - Banks et al. (2011); Kepes et al. (2014)

The Current Environment

The Current Environment

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

Open Science Collaboration[*,†]
+ See all authors and affiliations

Industrial and Organizational Psychology
PERSPECTIVES ON SCIENCE AND PRACTICE

Industrial and Organizational Psychology, **6** (2013), 252–268.

FOCAL ARTICLE

How Trustworthy Is the Scientific Literature in Industrial and Organizational Psychology?

SCIENCE

Psychology's Replication Crisis Can't Be Wished Away

It has a real and heartbreaking cost.

The Washington Post
Democracy Dies in Darkness

Monkey Cage

Does social science have a replication crisis?

The Current Environment

J Bus Psychol (2011) 26:105–121
DOI 10.1007/s10869-010-9185-2

Meta-analytic Decisions and Reliability: A Serendipitous Case of Three Independent Telecommuting Meta-analyses

Levi R. G. Nieminen · Jessica M. Nicklin · Tara K. McClure · Madhura Chakrabarti

Original Investigation

The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses

JOHN P.A. IOANNIDIS

Science   Contents   News   Careers   Journals

Meta-analyses were supposed to end scientific debates. Often, they only cause more controversy

By Jop de Vrieze | Sep. 18, 2018 , 4:15 PM

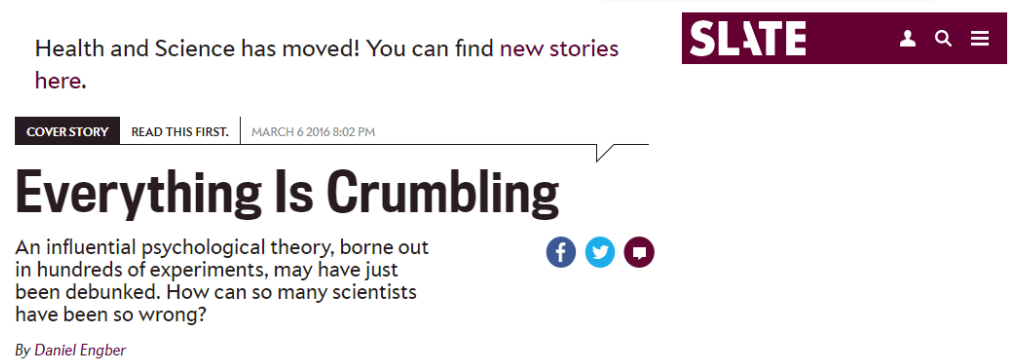
Meta-Analytic Choices and Judgment Calls: Implications for Theory Building and Testing, Obtained Effect Sizes, and Scholarly Impact

Herman Aguinis
Dan R. Dalton
Indiana University
Frank A. Bosco
Charles A. Pierce
University of Memphis
Catherine M. Dalton
Indiana University

Excerpt taken from Engber, D. (March, 2016). Everything is crumbling. *Slate*.

The Current Environment

- Meta-analysis is not immune from scrutiny
  - "All the old methods are in doubt. Even meta-analyses, which once were thought to yield a gold standard for evaluating bodies of research now seem somewhat worthless. "Meta-analyses are f*cked," Inzlicht warned me. If you analyze 200 lousy studies, you'll get a lousy answer in the end. It's garbage in, garbage out."
    - From:


SLATE

Health and Science has moved! You can find new stories here.

COVER STORY READ THIS FIRST. MARCH 6 2016 8:02 PM

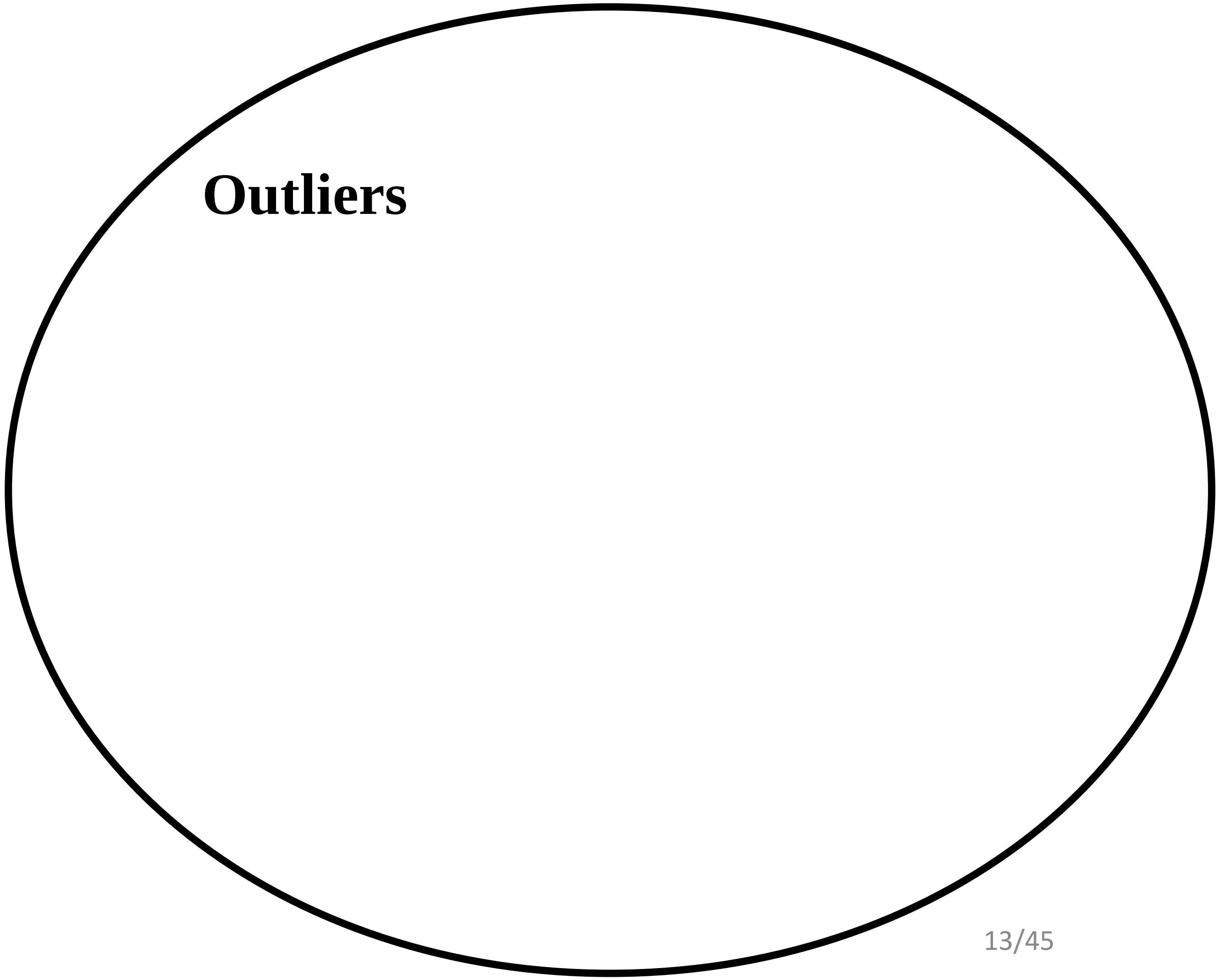
Everything Is Crumbling

An influential psychological theory, borne out in hundreds of experiments, may have just been debunked. How can so many scientists have been so wrong?

By Daniel Engber

- What could be driving opinions like these?

Threats to our Cumulative Knowledge



Threats to our Cumulative Knowledge

Outliers

An observation that appears “to deviate markedly from other members of the sample in which it occurs” (Grubbs, 1969, p. 1)

Threats to our Cumulative Knowledge

Outliers

Outcome-level causes
(e.g., effect size magnitude,
p-value)

Threats to our Cumulative Knowledge

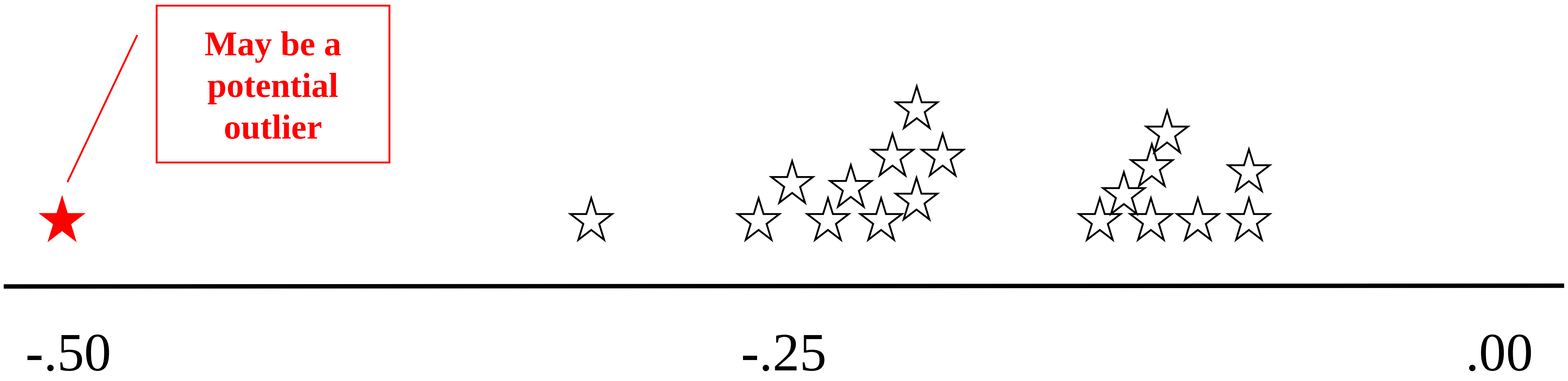
- Outcome-level causes of outliers: Effect size magnitude
  - Samples that have an effect size that diverges from all other samples in the dataset may need to be removed before performing a meta-analysis as they could introduce residual heterogeneity that may threaten its results and conclusions.

Threats to our Cumulative Knowledge

- Outcome-level causes of outliers: Effect size magnitude
  - Each ☆ represents an effect size in the Jiang et al. (2012) dataset (our running example)

Threats to our Cumulative Knowledge

- Outcome-level causes of outliers: Effect size magnitude
  - Each ☆ represents an effect size in the Jiang et al. (2012) dataset (our running example)

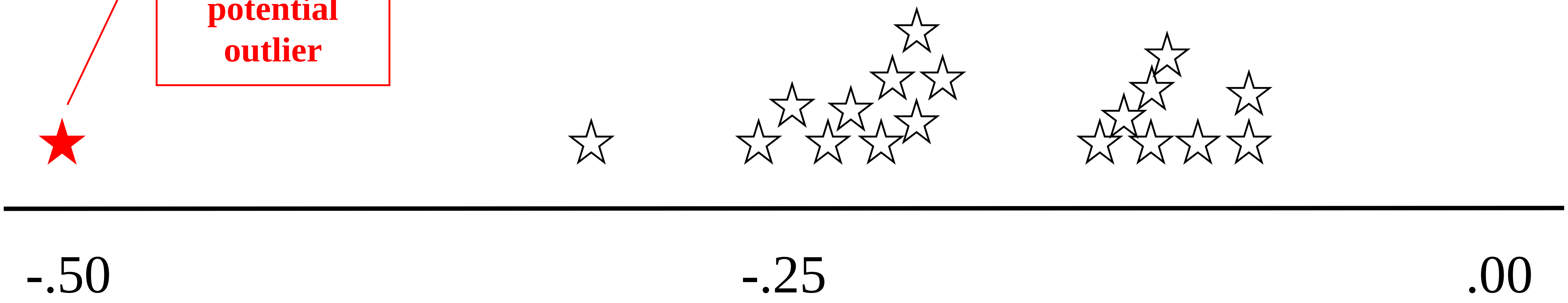


Threats to our Cumulative Knowledge

- Outcome-level causes of outliers: Effect size magnitude
  - Each ☆ represents an effect size in the Jiang et al. (2012) dataset (our running example)

This changes the meta-analytic mean effect size estimate from -.17 to -.22 (|Δ| = .05 or by 29%)

May be a potential outlier



Threats to our Cumulative Knowledge

Outliers

Outcome-level causes
(e.g., effect size magnitude,
p-value)

Sample-level causes
(e.g., sample size,
sample type)

Threats to our Cumulative Knowledge

- Sample-level causes of outliers: Sample size
  - Given that both the Hedges and Olkin (1985; see also Hedges & Olkin, 2014) and Schmidt and Hunter (2015) approaches to meta-analysis estimate the meta-analytic mean by giving more precise studies more weight, relatively large samples can have an undue influence on the meta-analytic mean.

Threats to our Cumulative Knowledge

- Sample-level causes of outliers: Sample size
  - The sample sizes included in Jiang et al.'s (2012) meta-analytic dataset range from 122 → 1,089

Threats to our Cumulative Knowledge

- Sample-level causes of outliers: Sample size
  - The sample sizes included in Jiang et al.'s (2012) meta-analytic dataset range from 122 → 1,089

Study ID	Sample size
1	177
2	306
3	208
4	474
5	122
6	164
7	809
8	112
9	323
10	205
11	222
12	142
13	9277
14	134
15	164
16	546
17	750
18	1,089

Threats to our Cumulative Knowledge

- Sample-level causes of outliers: Sample size
  - The sample sizes included in Jiang et al.'s (2012) meta-analytic dataset range from 122 → 1,089
  - Imagine adding an additional effect size that had a corresponding sample size of 50,000
  - Given that meta-analyses weight by precision, this addition would likely have a noticeable effect on the meta-analytic mean effect size estimate

Study ID	Sample size
1	177
2	306
3	208
4	474
5	122
6	164
7	809
8	112
9	323
10	205
11	222
12	142
13	9277
14	134
15	164
16	546
17	750
18	1,089

Threats to our Cumulative Knowledge

Outliers

Outcome-level causes
(e.g., effect size magnitude, *p*-value)

Sample-level causes
(e.g., sample size, sample type)

Publication bias

A systematic suppression of research findings, which causes the available literature to be unrepresentative of all completed research on a relation of interest

Threats to our Cumulative Knowledge

Outliers

Outcome-level causes
(e.g., effect size magnitude, *p*-value)

Sample-level causes
(e.g., sample size, sample type)

Publication bias

Outcome-level causes
(e.g., author decisions, editorial review process, organizational constraints)

Threats to our Cumulative Knowledge

Outliers

Outcome-level causes
(e.g., effect size magnitude, *p*-value)

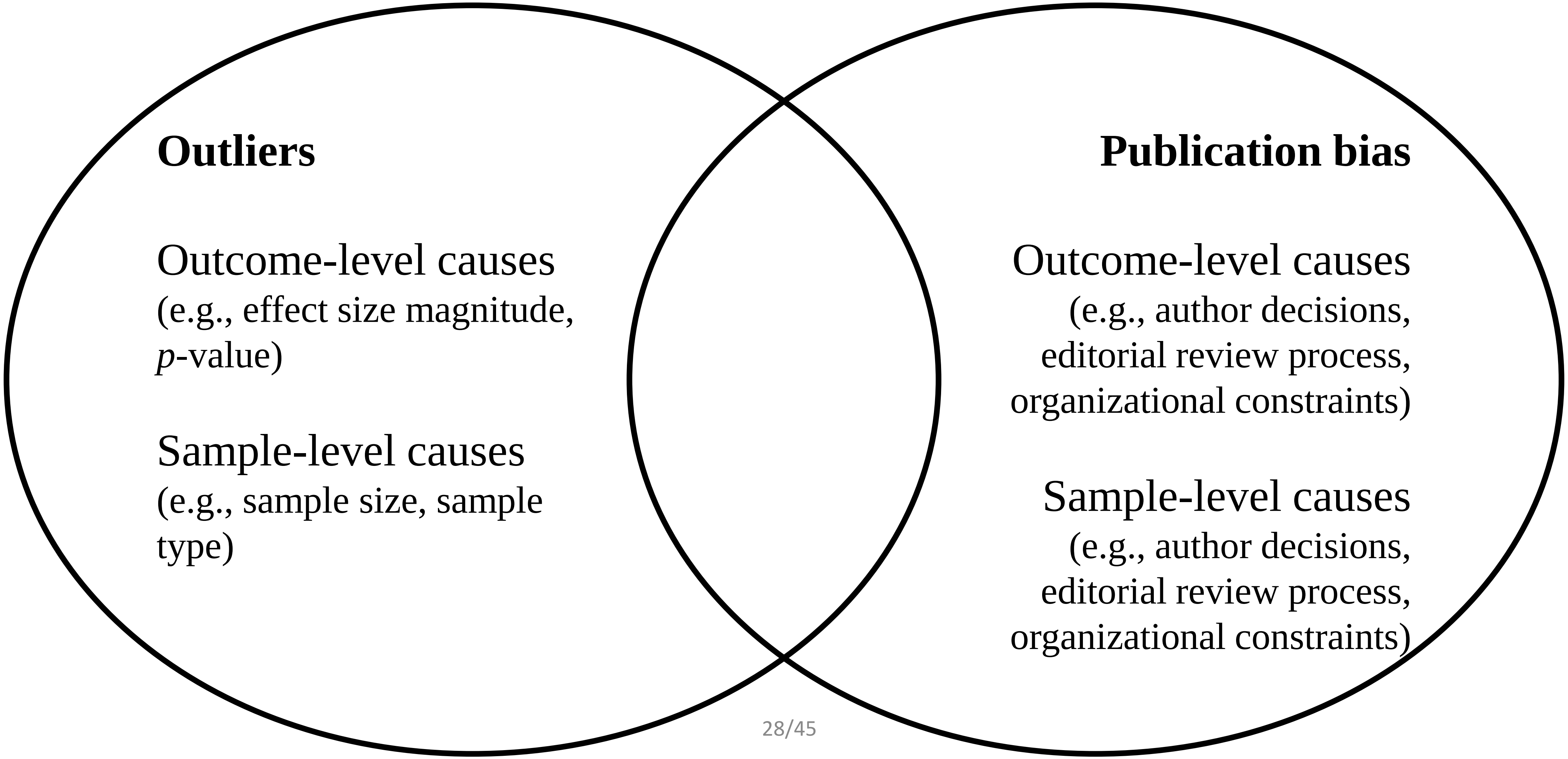
Sample-level causes
(e.g., sample size, sample type)

Publication bias

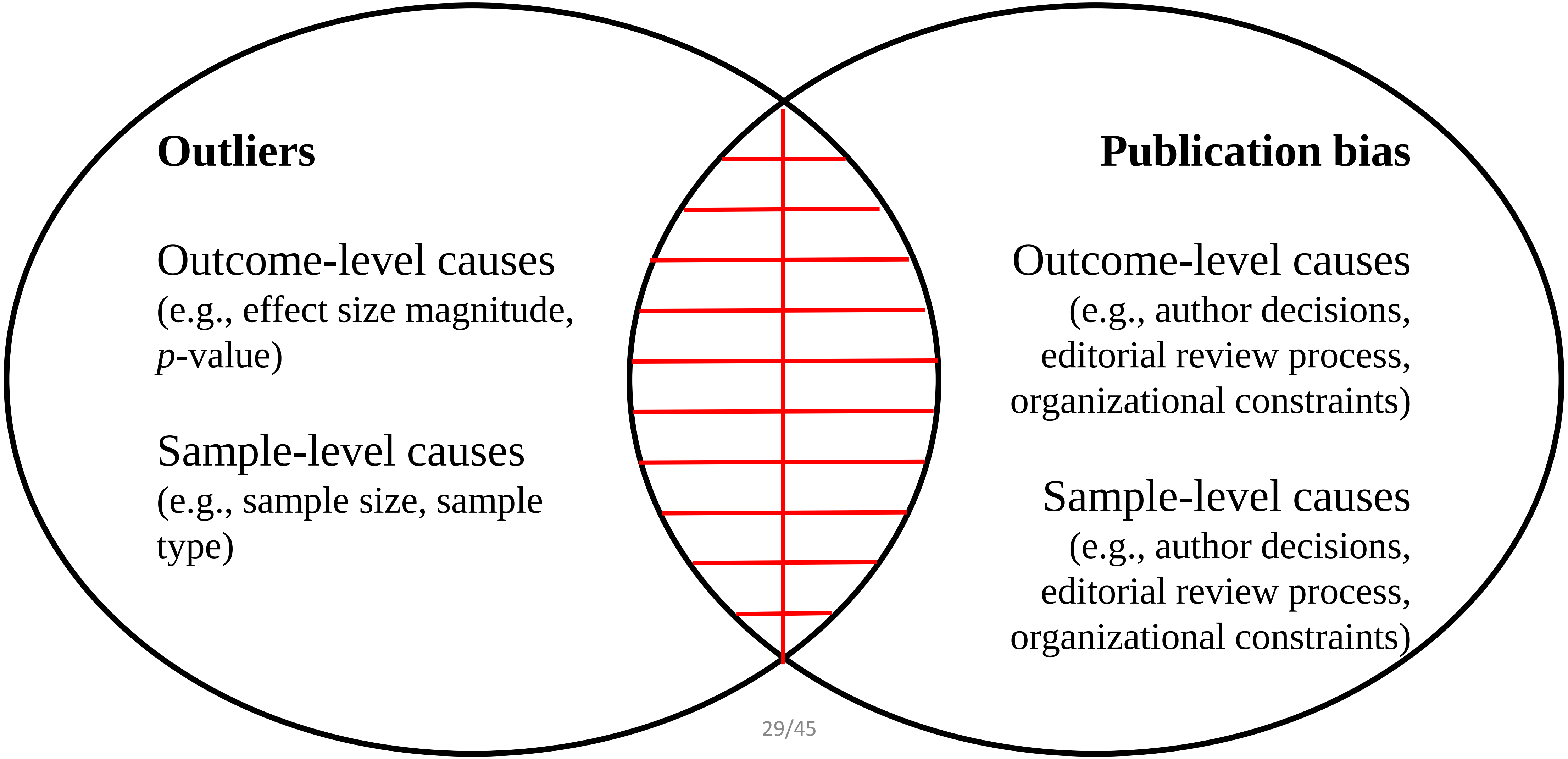
Outcome-level causes
(e.g., author decisions, editorial review process, organizational constraints)

Sample-level causes
(e.g., author decisions, editorial review process, organizational constraints)

Threats to our Cumulative Knowledge



Threats to our Cumulative Knowledge

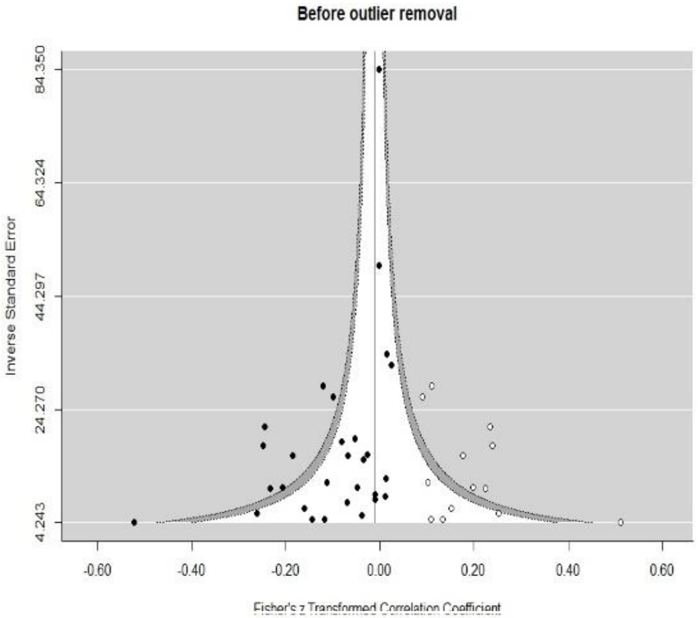


Threats to our Cumulative Knowledge

- Combined outlier *and* publication bias effect
  - Rarely tested!
  - However, outlier-induced heterogeneity may limit the efficacy of publication bias detection methods (Kepes & McDaniel, 2015; Peters et al., 2007; Terrin et al., 2003).
    - Some scholars have started to examine the possibility of a combined effect
      - (Benjamin et al.; in press; Kepes & McDaniel, 2015; Kepes et al., 2017; and Kepes & Thomas, 2018).

Threats to our Cumulative Knowledge

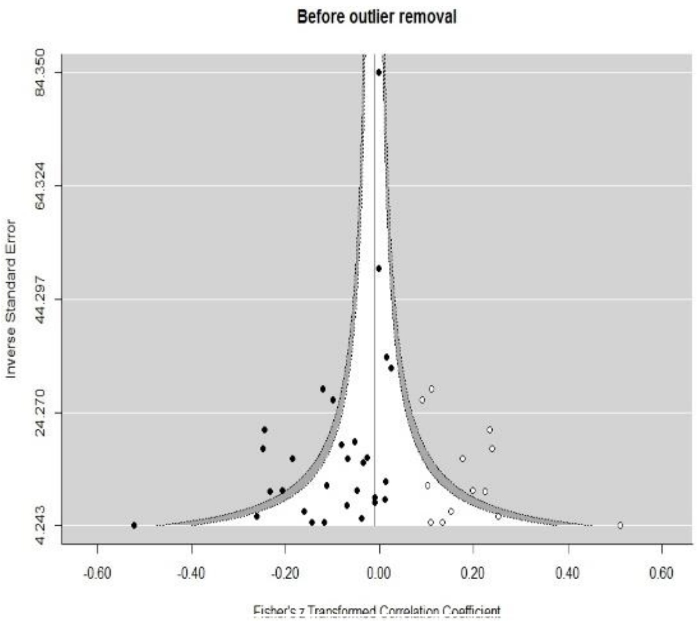
- Combined outlier *and* publication bias effect
  - k = 29
  - $\bar{r}_{o_{RE}}$ = -.08



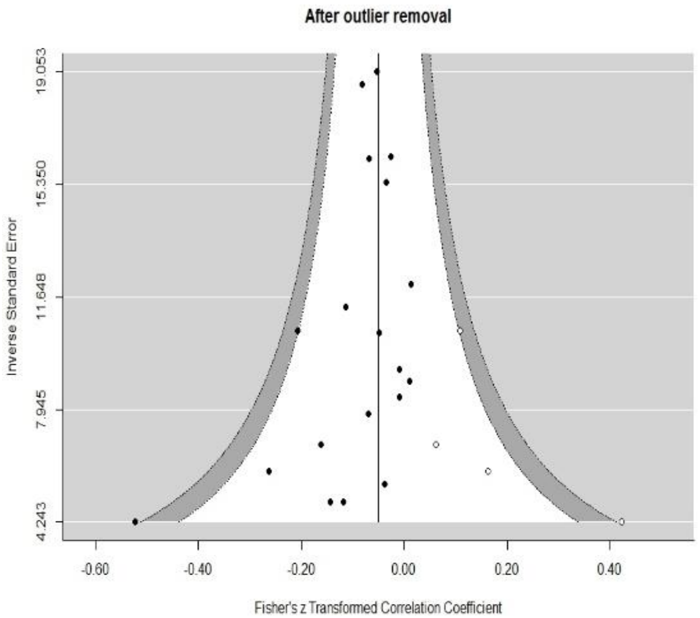
k = 29; ik = 13; t&f$_{FE}$ $\bar{r}_o$ = -.01

Threats to our Cumulative Knowledge

- Combined outlier *and* publication bias effect
  - k = 29
  - $\bar{r}_{o_{RE}}$ = -.08



k = 29; ik = 13; t&f$_{FE}\,\bar{r}_o$ = -.01



k = 19; ik = 4; t&f$_{FE}\,\bar{r}_o$ = -.05

  - Suggests that the publication bias detection result overestimates the distorting effect of publication bias!

What is Sensitivity Analysis?

- A sensitivity analysis examines the extent to which results and conclusions are altered as a result of changes in the data or analysis approach
  - Greenhouse & Iyengar (2009)

- If the conclusions do not change as a result of the sensitivity analysis, one can state that the conclusions are robust and one can have greater confidence in the conclusions.

What is Sensitivity Analysis?

- “Sensitivity analysis is the most powerful tool we have for assessing the influence of the specific choices made by the researchers”
  - Aytug, Rothstein, Zhou, & Kern (2012, p. 118)

What is Sensitivity Analysis?

- Sensitivity analyses are rarely conducted in meta-analyses in the organizational sciences
  - Kepes, McDaniel, Brannick, & Banks (2013)

- Because meta-analyses have a strong impact on our literatures, sensitivity analyses need to become much more common (and reported) in meta-analyses.

Live Meta-Sen Demonstration

- https://metasen.shinyapps.io/gen1/
- You can find some additional sample data files here: https://jamiefield.github.io/research/sma2019

How to minimize the impact of outliers & PB

- Ultimately, it is always best to report the range of results

- The effect of PB can be reduced by
  - Conducting extremely thorough literature reviews
  - Using research registries
  - Changing the journal review process
  - Altering author and organization norms
  - Obsessing less about theoretical contributions
  - Supporting data repositories like metaBUS

How to minimize the impact of outliers & PB

Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7

- Step 1: Conduct a meta-analysis on original dataset

How to minimize the impact of outliers & PB

Step 1 → Step 2 → Step 3 → Step 4 → Step 5 → Step 6 → Step 7

- Step 1: Conduct a meta-analysis on original dataset
- Step 2: Perform osr and recommended PB analyses

How to minimize the impact of outliers & PB

Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7

- Step 1: Conduct a meta-analysis on original dataset
- Step 2: Perform osr and recommended PB analyses
- Step 3: Perform Viechtbauer and Cheung's (2010) influence diagnostics

How to minimize the impact of outliers & PB

- Step 1: Conduct a meta-analysis on original dataset
- Step 2: Perform osr and recommended PB analyses
- Step 3: Perform Viechtbauer and Cheung's (2010) influence diagnostics
- Step 4: If detected, remove outliers and repeat Steps 1 and 2. If outliers are not detected in Step 3, proceed directly to Step 5

How to minimize the impact of outliers & PB

Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7

- Step 1: Conduct a meta-analysis on original dataset
- Step 2: Perform osr and recommended PB analyses
- Step 3: Perform Viechtbauer and Cheung's (2010) influence diagnostics
- Step 4: If detected, remove outliers and repeat Steps 1 and 2. If outliers are not detected in Step 3, proceed directly to Step 5
- Step 5: Report BRE and MRE (see Kepes et al., 2012)**

How to minimize the impact of outliers & PB

Step 1 → Step 2 → Step 3 → Step 4 → Step 5 → Step 6 → Step 7

- Step 1: Conduct a meta-analysis on original dataset
- Step 2: Perform osr and recommended PB analyses
- Step 3: Perform Viechtbauer and Cheung's (2010) influence diagnostics
- Step 4: If detected, remove outliers and repeat Steps 1 and 2. If outliers are not detected in Step 3, proceed directly to Step 5
- Step 5: Report BRE and MRE (see Kepes et al., 2012)**
- Step 6: Visually inspect the range of results before and after outlier removal

How to minimize the impact of outliers & PB

Step 1 → Step 2 → Step 3 → Step 4 → Step 5 → Step 6 → Step 7

- Step 1: Conduct a meta-analysis on original dataset
- Step 2: Perform osr and recommended PB analyses
- Step 3: Perform Viechtbauer and Cheung's (2010) influence diagnostics
- Step 4: If detected, remove outliers and repeat Steps 1 and 2. If outliers are not detected in Step 3, proceed directly to Step 5
- Step 5: Report BRE and MRE (see Kepes et al., 2012)**
- Step 6: Visually inspect the range of results before and after outlier removal
- Step 7: Assess the robustness of recommendations for practice

Questions?

Thank you for attending today!

Feel free to follow up with me

james.field2@mail.wvu.edu

In the following slides…

- I review two outlier and five publication bias assessment methods used by the Meta-Sen app
- I present results that illustrate an outlier and PB effect, as well as a combined effect of these phenomena, on meta-analytic findings on employee turnover

Review of Two Outlier Assessment Methods

- One form of sensitivity analysis is to conduct meta-analyses with and without outliers

- Only 3% of meta-analyses conduct outlier analyses (Aguinis et al., 2011)
  - Effect size outlier (large or small)
    - Graphical methods and statistical tests for outliers (e.g., SAMD statistic; Beal, Corey, & Dunlap, 2002)
  - Sample size outlier (large)
    - Sample sizes influence effect size weights in meta-analyses.

Review of Two Outlier Assessment Methods

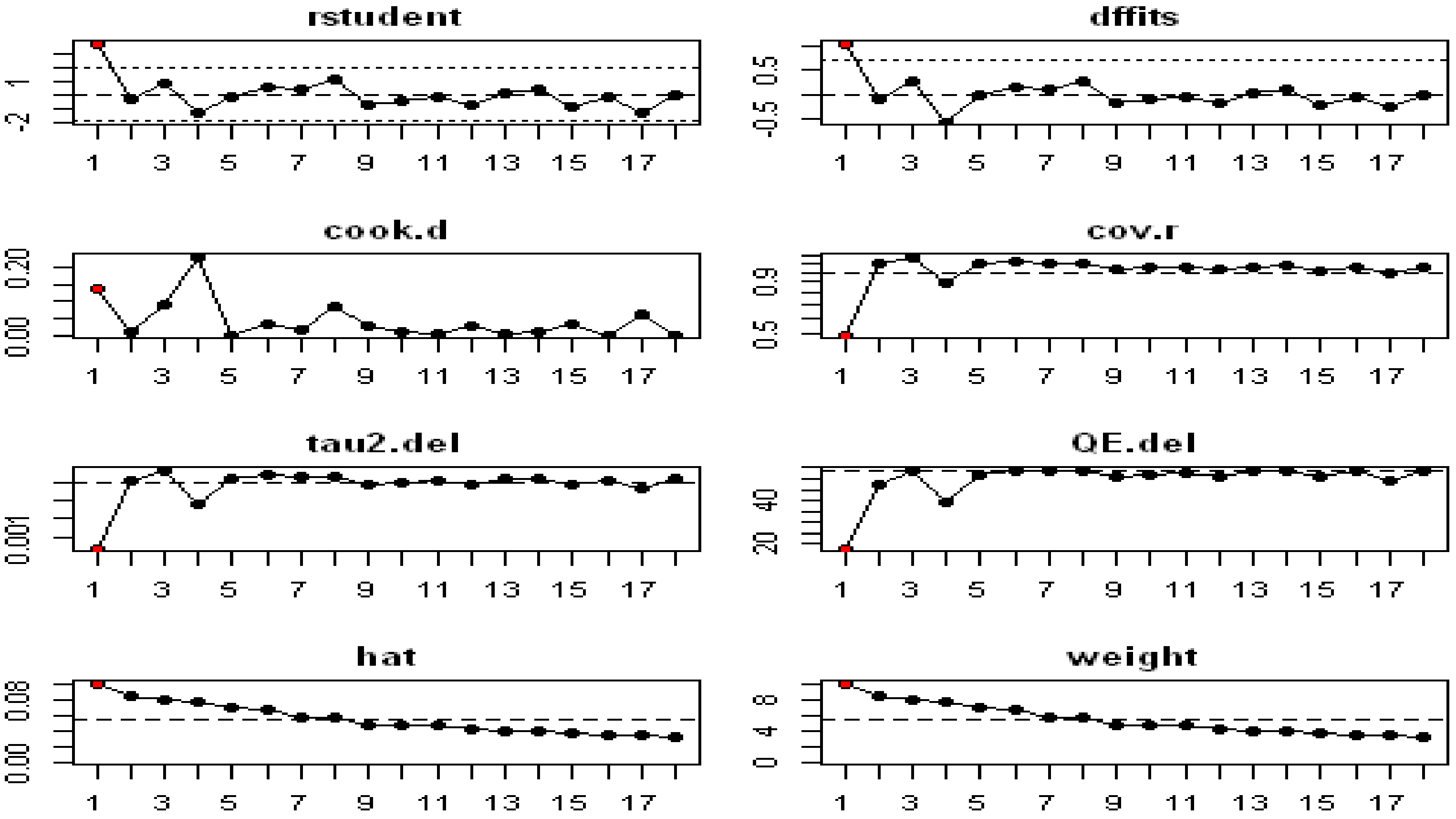
- One sample removed analysis:
  - Individual samples are removed one-by-one from the dataset and the point estimate is recalculated after each removal.
  - Thus, a one-sample removed analysis, yields *k*-1 meta-analytic mean estimates.
  - Given the Jiang, et al. (2012) dataset included 18 effect sizes, the one-sample removed analysis will produce 17 estimates of the meta-analytic mean
  - Important questions to ask:
    - How much does the distribution mean change when a given sample is excluded from the analysis?
    - Are the results due to a small number of influential samples?

Review of Two Outlier Assessment Methods

- Viechtbauer and Chueng’s (2010; Viechtbauer [2015]) multivariate, multidimensional influence diagnostics:
  - A framework that calculates leave-one-out analyses for
    - externally standardized residuals
    - DFFITS value,
    - Cook’s distance,
    - covariance ratio,
    - the leave-one-out amount of heterogeneity,
    - the leave-one-out test statistic for the test of heterogeneity, and
    - DFBETAS values.
  - In addition, an inspection of the hat matrix is examined for highly influential observations.

Review of Two Outlier Assessment Methods

- Viechtbauer and Chueng's (2010; Viechtbauer [2015]) multivariate, multidimensional influence diagnostics:



!!IMPORTANT!!

This is an iterative process that must be performed until *all* identified outliers are removed

Review of Five PB Assessment Methods

- Symmetry-based methods

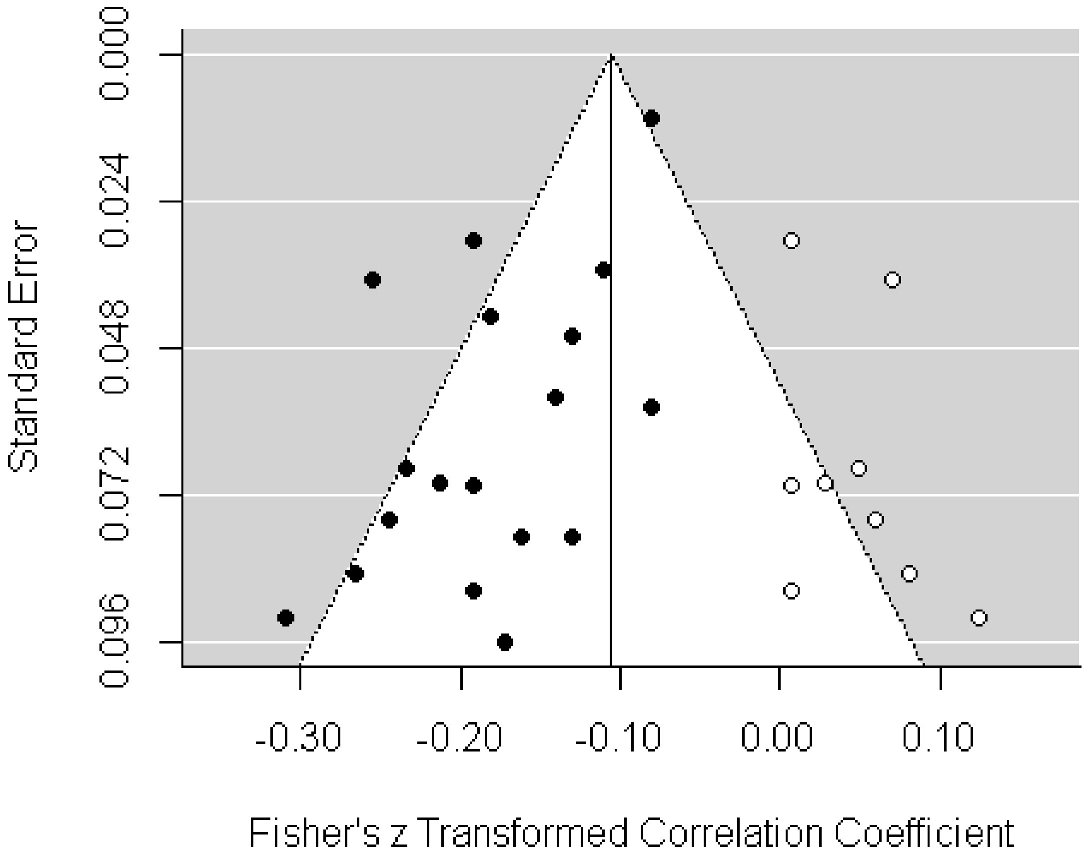
  - When sampling error is the sole source of variance, and the sampling distribution is symmetrical, then a funnel plot can be examined for symmetry.

  - A funnel plot is a plot of effect sizes by precision (1/standard error).

  - Examples of symmetry-based methods include (1) trim and fill models and (2) contour-enhanced funnel plot

Review of Five PB Assessment Methods

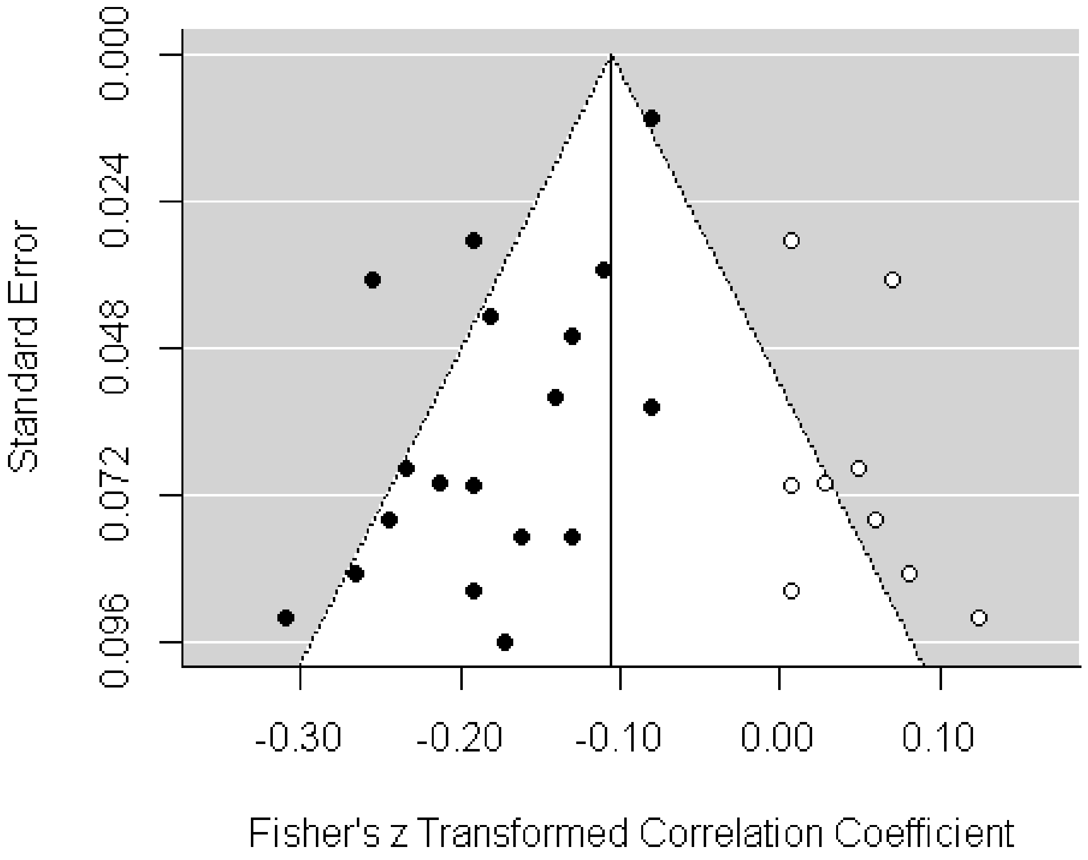
- Trim and fill models
  - The trim and fill method is probably the most useful symmetry based method in that it estimates what the population distribution would be if the missing studies were located
  - Analyses are re-conducted on the distribution containing both the observed data and the imputed data



FE trim and fill model of Jiang et al.'s meta-analytic distribution

Review of Five PB Assessment Methods

- Trim and fill models
  - It is unwise to consider this distribution of observed and imputed data as the “true” distribution
  - More reasonable to compare the observed mean with the trim and fill adjusted mean
  - If the mean drops from .45 to .15, one should worry about publication bias
  - But, one should not assume that .15 is the best estimate of the population mean



FE trim and fill model of Jiang et al.’s meta-analytic distribution

Review of Five PB Assessment Methods

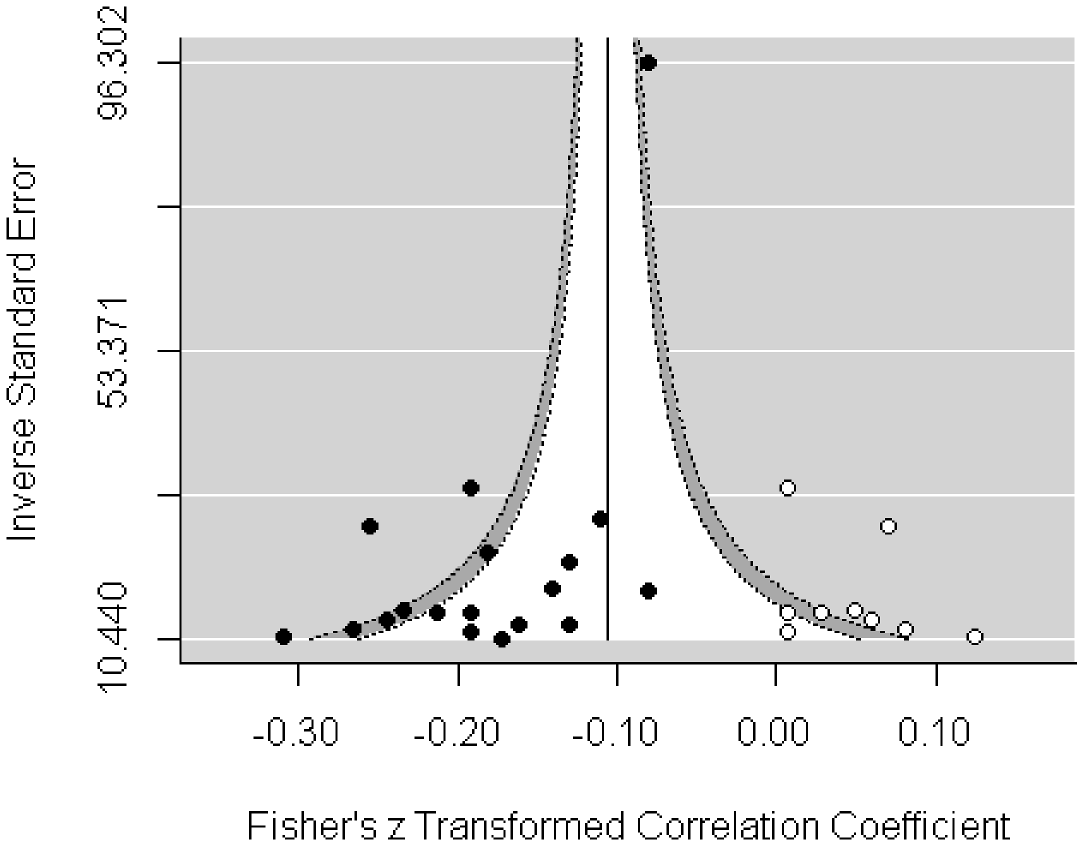
- Some asymmetry is not due to publication bias but to “small sample effects.”

  - A medicine may work best with the sickest (small N) patients and work less well with moderately sick (larger N) patients.

  - Small sample studies may yield larger effects due to better measures that are more difficult to collect in larger samples.

Review of Five PB Assessment Methods

- Contour-enhanced funnel plots
    - Related to the funnel plot and trim and fill is the contour-enhanced funnel plot, which displays graphically whether the imputed samples are a function of statistical significance (Peters et al., 2008).
        - Helps separate publication bias effects from "small sample effects."



Contour enhanced funnel plot of Jiang et al.'s meta-analytic distribution

Review of Five PB Assessment Methods

- A priori selection models
  - Selection models, also called weight-function models, originated in econometrics to estimate missing data at the item level.
  - Hedges and Vevea introduced the method to the publication bias literature
    - Hedges (1992)
    - Vevea and Hedges (1995)
  - Relatively robust to heterogeneity
    - Vevea and Woods (2005)

Review of Five PB Assessment Methods

- A priori selection models
    - As with trim and fill, selection models estimate what the population distribution would be if the missing studies were located and included in the meta-analytic distribution
    - When one is conducting a meta-analysis without regard to suppressed studies, one is implicitly assuming that one has 100% of the completed studies
        - This assumption is unlikely to be valid
            - Vevea and Woods (2005)
- Selection models permit you to make other assumptions

Review of Five PB Assessment Methods

- A priori selection models

  - Selection models assume that the probability that an effect size is included in a distribution is a function of a characteristic of that effect size
    - This characteristic is usually the level of statistical significance

  - Consider an *a priori* assumed selection model

Review of Five PB Assessment Methods

- A priori selection models
  - Selection models assume that the probability that an effect size is included in a distribution is a function of a characteristic of that effect size
    - This characteristic is usually the level of statistical significance
  - Consider an *a priori* assumed selection model

Significance level	**Probability of being in the distribution**
$p \leq .001$	100%
$.001 < p \leq .05$	90%
$.005 < p \leq .10$	70%
$p > .10$	30%

Review of Five PB Assessment Methods

- Cumulative meta-analysis by precision
  - Sort samples by sample size or precision
  - Conduct a meta-analysis starting with one effect size (the most precise effect) and add an additional effect size (with increasingly less precision) with each iteration of the meta-analysis
  - Inspect the meta-analytic means for drift

Review of Five PB Assessment Methods

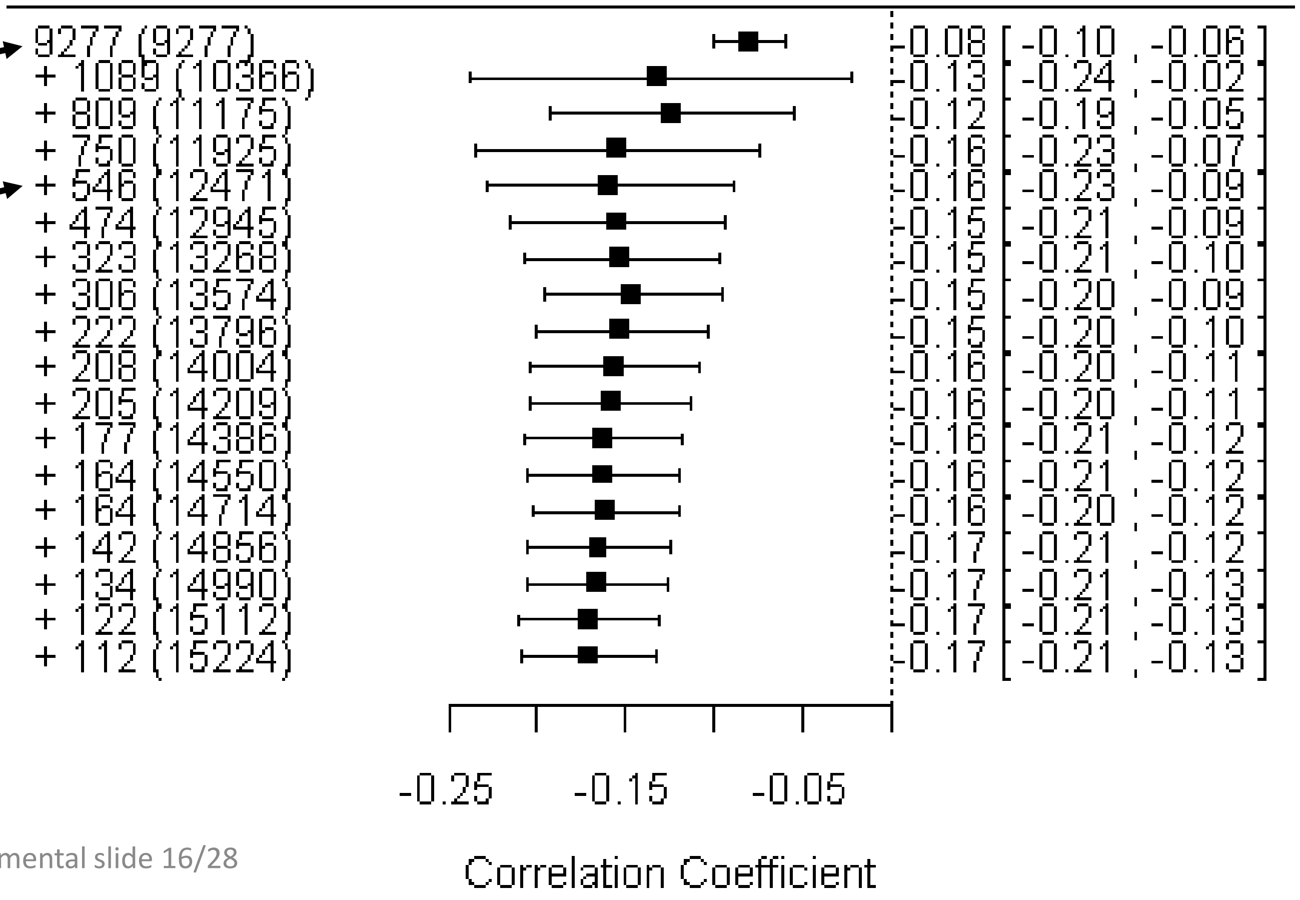
- Cumulative meta-analysis by precision

CMA by precision of Jiang et al.'s meta-analytic distribution

The most precise sample (N = 9,277), has an effect size of -.08.

With five studies, the cumulative sample size is 12,471 and the mean effect size is -.16

By the time one gets to 18 studies (N = 15,224), the mean effect size is -.17



Review of Five PB Assessment Methods

- Cumulative meta-analysis by precision
  - Gives similar results to that obtained in symmetry based methods
  - When symmetry analyses suggest small effects are suppressed, cumulative meta-analysis will show a drift toward larger effects
  - When symmetry analyses suggest larger effects are suppressed, cumulative meta-analysis will show a drift toward smaller effects.

Review of Five PB Assessment Methods

- Precision-effect test-precision effect estimate with standard error analysis (PET-PEESE)

  - A relatively new PB detection technique (Stanley and Doucouliagos, 2014)

  - This method is a combination of two regression models (PET and PEESE)
    - Conditional decision rule that determines which of the two models should be used

  - PET → Observed effect sizes are regressed on their corresponding standard errors using meta-regression techniques

  - PEESE → Observed effect sizes are regressed on their corresponding *squared* SE

Evidence of Combined Outlier and PB Effect

- Comprehensive sensitivity analyses were conducted on 84 recently published meta-analytic distributions on employee turnover
    - Examined the trustworthiness of these distributions
    - Does a greater threat to the trustworthiness arise from outliers or publication bias?
    - Assessed if meta-analytic and PB results changed after outlier removal
    - Examined whether or not recommendations for practice were robust to outliers and/or PB

- We assessed the generalizability of our results to other management topics by performing comprehensive sensitivity analyses on an additional 103 meta-analytic distributions
    - These data were taken from the metaBUS database (see metaBUS.org)

Evidence of Combined Outlier and PB Effect

- **How trustworthy is our cumulative scientific knowledge on turnover?**
  - 92% (77/84) of the turnover distributions were misestimated by a “noticeable” amount (i.e., > 20%; Kepes et al., 2012)
  - 96% (99/103) of the metaBUS distributions were misestimated by a “noticeable” amount

Evidence of Combined Outlier and PB Effect

- **Does a greater threat to the trustworthiness arise from outliers or publication bias?**

  - PB was the source of non-robustness in 92% (77/84) of the turnover meta-analytic mean effect size estimates
  - Outliers was the source of the non-robustness in 48% (40/84) of the turnover meta-analytic mean effect size estimates
  - Therefore, PB > outliers

  - A combined outlier *and* PB effect was observed in 48% (40/84) of the turnover distributions

Evidence of Combined Outlier and PB Effect

- **Does a greater threat to the trustworthiness arise from outliers or publication bias?**

  - PB was the source of non-robustness in 96% (99/103) of the metaBUS meta-analytic mean effect size estimates
  - Outliers was the source of the non-robustness in 74% (76/103) of the metaBUS meta-analytic mean effect size estimates
  - Therefore, and similar to the turnover results, PB > outliers

  - A combined outlier *and* PB effect was observed in 96% (99/84) distributions

Evidence of Combined Outlier and PB Effect

- **Do outliers distort meta-analytic results?**
  - 52% (44/84) of the turnover distributions contained at least one outlier
    - Nine had $k < 10$ and, thus, could not be reanalyzed after outlier removal
  - For the 35 that could be compared, our results suggest that 86% (30/35) of the meta-analytic mean effect size estimates changed after outlier removal
    - 34% (12/30) were misestimated by more than 20%
  - Similar results were observed for the metaBUS distributions
    - 81% (83/103) distributions had at least one outlier
    - The meta-analytic mean effect size changed in 68% (54/79) of the cases
      - 11% (9/79) were misestimated by at least 20%

Evidence of Combined Outlier and PB Effect

- **Do outliers distort publication bias results?**
  - For the 35 turnover distributions that could be compared::
    - t&f$_{FE}$ $\bar{r}_o$ remained unchanged following outlier removal in only 31% (11/35) of the cases
    - t&f$_{RE}$ $\bar{r}_o$ remained unchanged following outlier removal in only 20% (7/35) of the cases
    - *pr* $\bar{r}_o$remained unchanged following outlier removal in only 40% (14/35) of the cases
    - *pp* $\bar{r}_o$ remained unchanged following outlier removal in only 37% (13/35) of the cases
    - sm$_m$ $\bar{r}_o$ remained unchanged following outlier removal in only 6% (2/35) of the cases

  - For the 79 metaBUS distributions that could be compared:
    - t&f$_{FE}$ $\bar{r}_o$ remained unchanged following outlier removal in only 27% (21/79) of the cases
    - t&f$_{RE}$ $\bar{r}_o$ remained unchanged following outlier removal in only 19% (15/79) of the cases
    - *pr* $\bar{r}_o$remained unchanged following outlier removal in only 71% (56/79) of the cases
    - *pp* $\bar{r}_o$ remained unchanged following outlier removal in only 18% (14/79) of the cases
    - sm$_m$ $\bar{r}_o$ remained unchanged following outlier removal in only 38% (30/79) of the cases

Evidence of Combined Outlier and PB Effect

- **Do recommendations for practice change after accounting for outliers and PB?**
  - Hancock et al. (2013) recommended that most organizations should increase their investments in reducing turnover
  - Estimated that a one SD decrease in turnover would be associated with a \$352 million increase in profits for *Fortune* 1,000 companies
  - However, out results suggest that this may be dramatically overestimated
  - Our FE trim and fill mean estimate following outlier removal (k = 46, t&f$_{FE}$ $\bar{r}_o$ = -.02 suggests that a one SD decrease in turnover would be associated with a \$101 million increase in profits
    - Suggests that the originally estimated financial benefit of a reduction in turnover may be overestimated by \$251 million (or 249%)

Evidence of Combined Outlier and PB Effect

- **Do recommendations for practice change after accounting for outliers and PB?**
  - We found that 75% (12/15) of the recommendations for practice were *not* robust to outliers and publication bias

  - Specifically, at least one of the following three occurred after taking into account the effect of outliers and/or PB
    - The direction of the meta-analytic mean used to justify the recommendation changed
    - The magnitude of he meta-analytic mean used to justify the recommendation changed by at least 20%
    - A moderating effect used to justify the recommendation disappeared

Convergence of PB Detection Methods

Convergence Rates Regarding Practical Differences Before and After Outlier Removal for 35 Turnover Distributions

	Before outlier removal								After outlier removal		
PB method	Negligible	Moderate	Severe	1.	2.	3.	4.	5.	Negligible	Moderate	Severe
1. t&f$_{FE}$ $\bar{r}_o$	4 (11%)	6 (17%)	25 (71%)	-	21 (60%)	12 (34%)	26 (74%)	25 (73%)	11 (31%)	6 (17%)	18 (51%)
2. t&f$_{RE}$ $\bar{r}_o$	15 (43%)	8 (23%)	12 (34%)	16 (46%)	-	20 (57%)	21 (60%)	15 (43%)	16 (47%)	9 (26%)	10 (29%)
3. sm$_{\mathrm{m}}$ $\bar{r}_o$	16 (47%)	13 (37%)	6 (17%)	10 (29%)	14 (40%)	-	12 (34%)	9 (26%)	25 (71%)	9 (26%)	1 (3%)
4. *pr* $\bar{r}_o$	8 (23%)	6 (17%)	21 (60%)	23 (66%)	19 (54%)	14 (40%)	-	18 (51%)	12 (34%)	10 (29%)	13 (37%)
5. *pp* $\bar{r}_o$	2 (6%)	4 (11%)	29 (83%)	29 (83%)	15 (43%)	9 (26%)	21 (60%)	-	6 (17%)	6 (17%)	23 (66%)

Convergence of PB Detection Methods

- Based on our turnover and metaBUS results we recommend that future meta-analysts use the following to triangulate the potentially most robust estimate of the "true" meta-analytic effect size
  - FE trim and fill model
  - CMA by precision
  - PET-PEESE analysis

- For outlier detection, we recommend Viechtbauer and Cheung's (2010; Viechtbauer 2015) influence diagnostics procedure due to its statistical rigor