Meta-analyse: Beperkingen en potentieel

Robbie C.M. van Aert

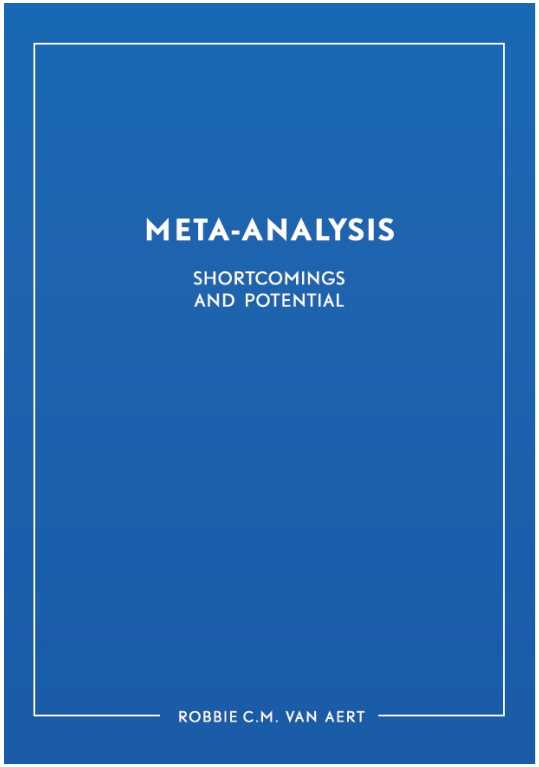
29 september, 2020

European Research Council

Established by the European Commission

Introduction

- Research Master at Tilburg University
- Completed PhD in 2018 also at Tilburg University
- Title of PhD thesis: "Meta-analysis: Shortcomings and potential"
- Currently, postdoctoral researcher working on meta-analysis methods



Outline

1. History of meta-analysis
2. Why do we need meta-analysis?
3. Stages of a meta-analysis
4. Meta-analysis models
5. Issues with meta-analysis
6. Publication bias
7. Snapshot Bayesian Hybrid Meta-Analysis Method

History of meta-analysis

- **Overarching aim:** Summarizing the available evidence
- **Prior to 1990s:** Narrative literature review where an expert reads the literature and answers a research question
- When using vote counting there are three options:
  - Significant in right direction
  - Nonsignificant
  - Significant in wrong direction

$\rightarrow$ Modal category is the winner!

- Inadequate method because power of studies is ignored

History of meta-analysis

- **Now:** Systematic review and meta-analysis

- Starting point is that the research synthesis should be *scientific*, *systematic*, *transparent*, and *replicable/reproducible*

History of meta-analysis

- **Now:** Systematic review and meta-analysis
- Starting point is that the research synthesis should be *scientific*, *systematic*, *transparent*, and *replicable/reproducible*
- Systematic review: clear set of rules that are specified in advance with respect to inclusion or exclusion of studies
- Meta-analysis: "the statistical synthesis of the data from separate but similar studies leading to a quantitative summary" (Last, 2001)

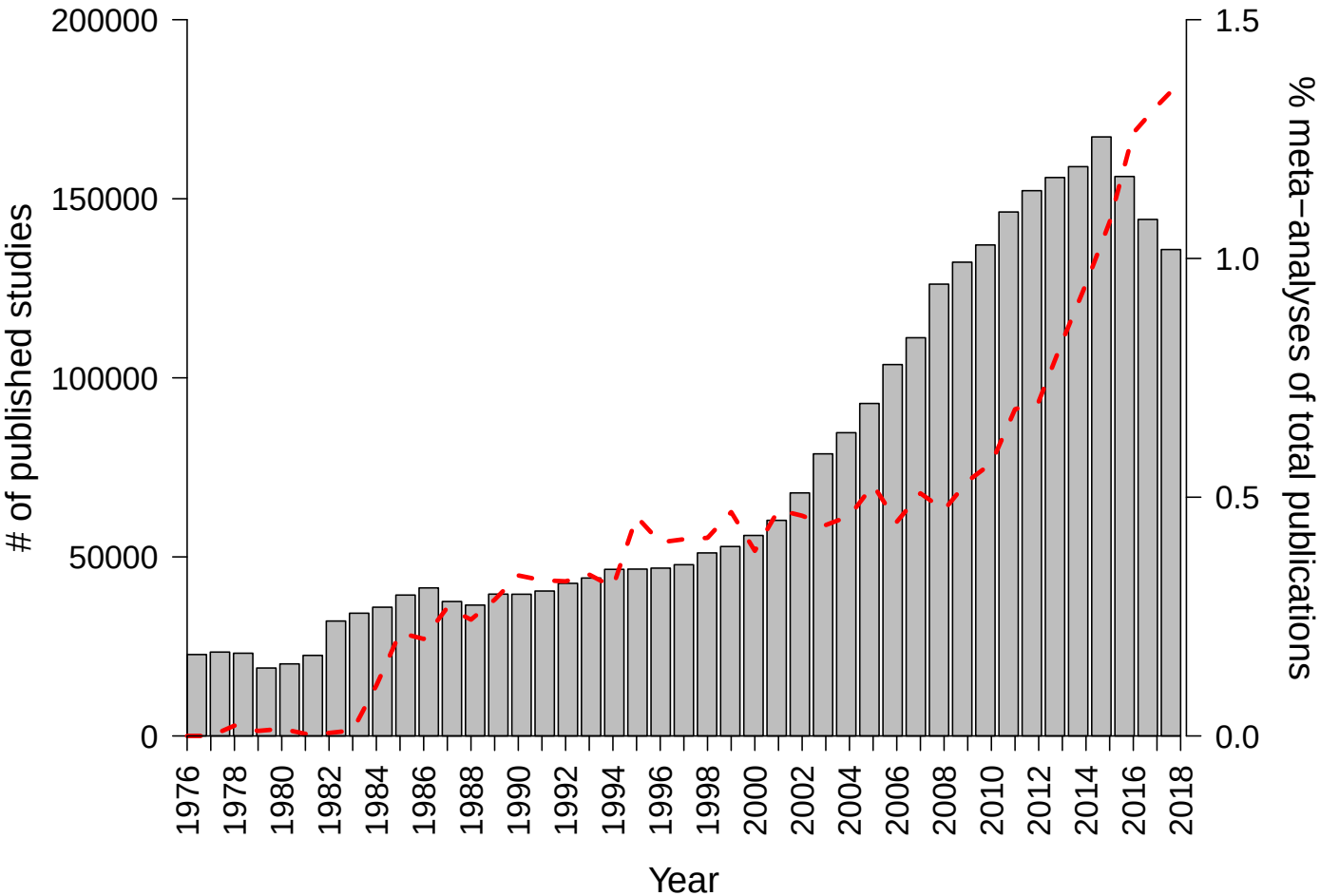
Why do we need meta-analysis?

Why do we need meta-analysis?

1. Information explosion
2. Single studies are insufficient
3. Additional (meta) information
4. Guaranteed success! However, there is a but...

1. Information explosion

- More and more studies get published → how can we keep up with reading all these studies?



1. Information explosion

Number of hits using PsycINFO:

- Sex differences and personality $\rightarrow$ 9,249 hits
- Effects of psychotherapy on depression $\rightarrow$ 8,769 hits
- Leadership and job performance $\rightarrow$ 2,238 hits
- Housing and discrimination $\rightarrow$ 773 hits
- Etc.

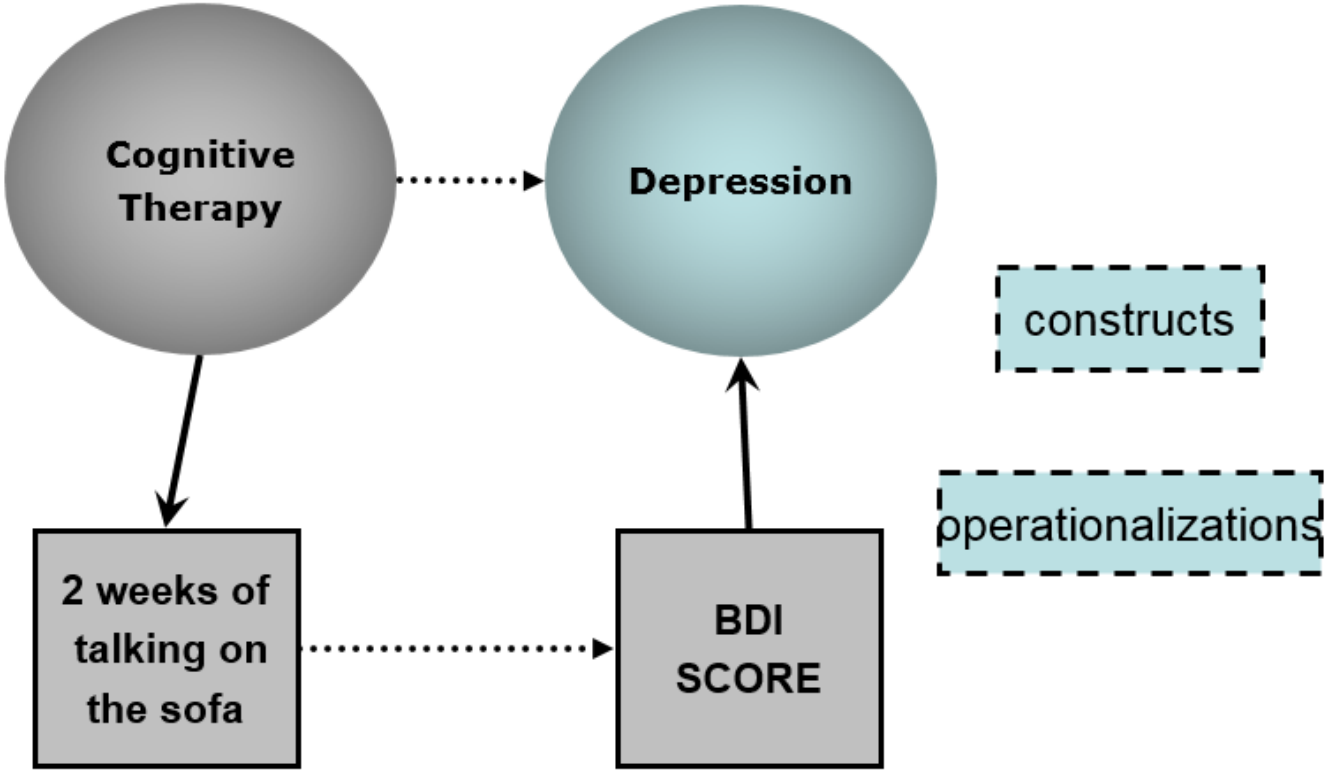
Conclusion: methods are needed to summarize research findings and that give an objective overview

2. Single studies are insufficient

- Major claims are often made in single studies that can be very context specific

2. Single studies are insufficient

Claim: "Cognitive therapy stops depression"



"Scores on the Beck Depression Inventory are lowered by two points after therapist X has talked to 30 depressed Psy-students daily for two hours over a period of two weeks"

2. Single studies are insufficient

- Major claims are often made in single studies that can be very context specific

- What about generalizability?

This study is about:

- a particular experiment
- in a particular lab setting
- using a particular test
- in a particular student sample

We want to draw conclusions on:

- a general effect
- in a general population

3. Additional (meta) information

Characteristics of a meta-analysis:

1. Synthesizing results from multiple studies
2. Estimating the average effect size across included studies
3. Estimating differences in studies' true effect size (variance)
4. Examining whether study-level characteristics matter $\rightarrow$ moderators
5. Sensitivity analyses (study publication bias)

4. Guaranteed success!

- You will always have a finding that is worthwhile!
- More citations than a primary study
- Meta-analyses are often used for policy making
- You get a good overview of the literature yourself as well

4. Guaranteed success! But...

Performing a meta-analysis is a **lot of work**!

search the literature

scan 2,000 papers for relevance

select suitable studies

code studies

compute effect size

do the analysis

write the paper

4. Guaranteed success! But...

Not there yet, really a **lot of work**!

submit the paper and get the reviews

redo the entire literature search

scan another 1,200 papers for relevance

select suitable studies again

recompute all effect sizes

redo all analyses

revise the paper

resubmit it

fingers crossed

Stages of a meta-analysis

Stages of a meta-analysis

1. Formulating a problem/research question
2. Literature search
3. Quality evaluation and extracting information
4. Data preparation (converting effect sizes)
5. Combining effect sizes and interpretation
6. Sensitivity analysis
7. Presentation of results

Meta-analysis models

Meta-analysis models

- Example:
  - Researchers wonder what the job satisfaction of employees is during the COVID-19 pandemic
  - Twenty studies are deemed eligible for inclusion
  - Studies differ on characteristics (e.g., number of weeks participants worked from home, country where a study was conducted)

Meta-analysis models

- Example:
  - Researchers wonder what the job satisfaction of employees is during the COVID-19 pandemic
  - Twenty studies are deemed eligible for inclusion
  - Studies differ on characteristics (e.g., number of weeks participants worked from home, country where a study was conducted)

Question: What meta-analysis model do we need?

- Variability in observed effect size estimates (y_i) is caused by:
  1. Sampling variability
  2. Variability among the true effect sizes
  3. Systematic differences among the true effect sizes

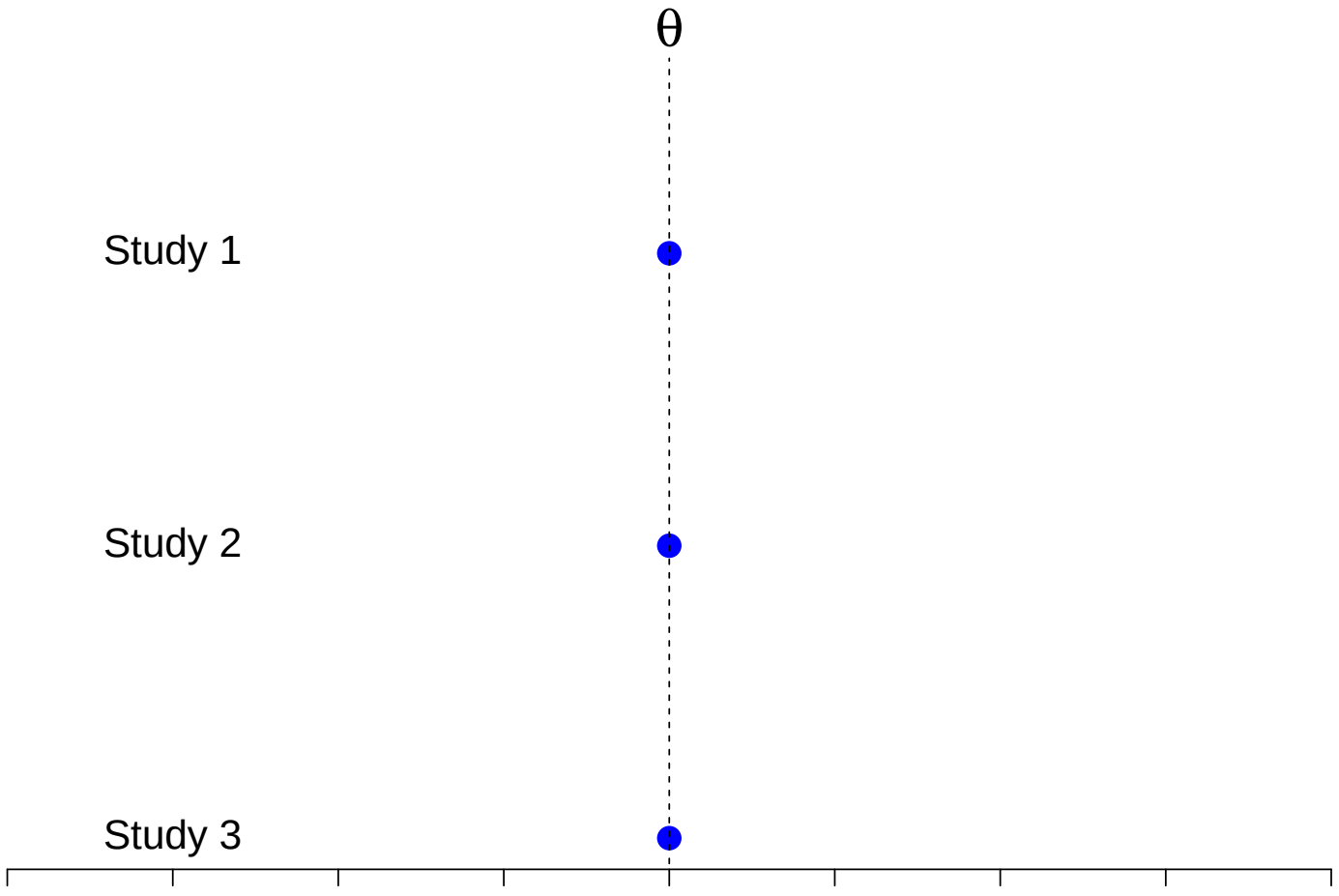
Meta-analysis models

- Example:
  - Researchers wonder what the job satisfaction of employees is during the COVID-19 pandemic
  - Twenty studies are deemed eligible for inclusion
  - Studies differ on characteristics (e.g., number of weeks participants worked from home, country where a study was conducted)

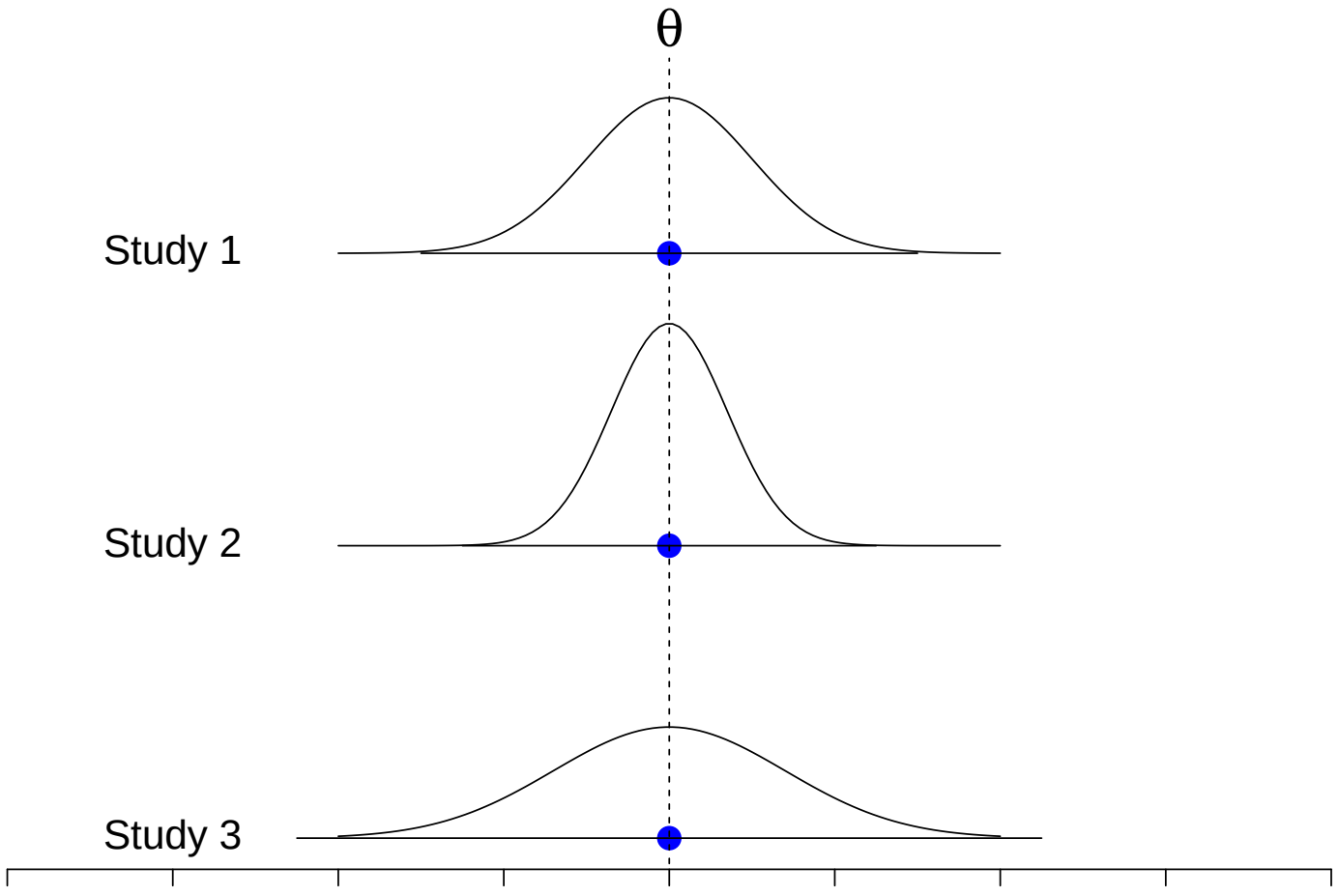
Question: What meta-analysis model do we need?

- Variability in observed effect size estimates (y_i) is caused by:
  1. Sampling variability
  2. Variability among the true effect sizes
  3. Systematic differences among the true effect sizes

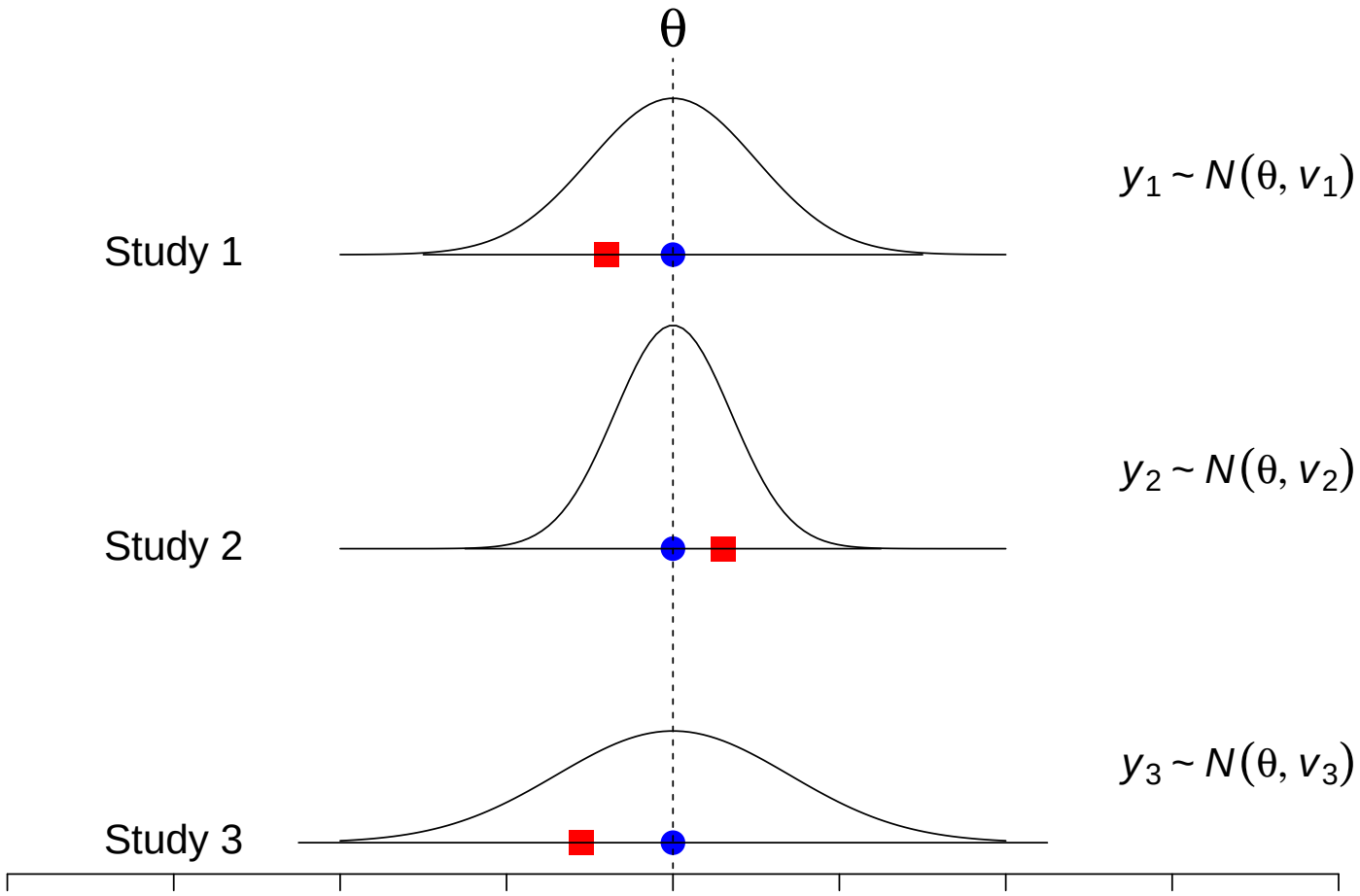
Meta-analysis models: Equal-effects model



Meta-analysis models: Equal-effects model



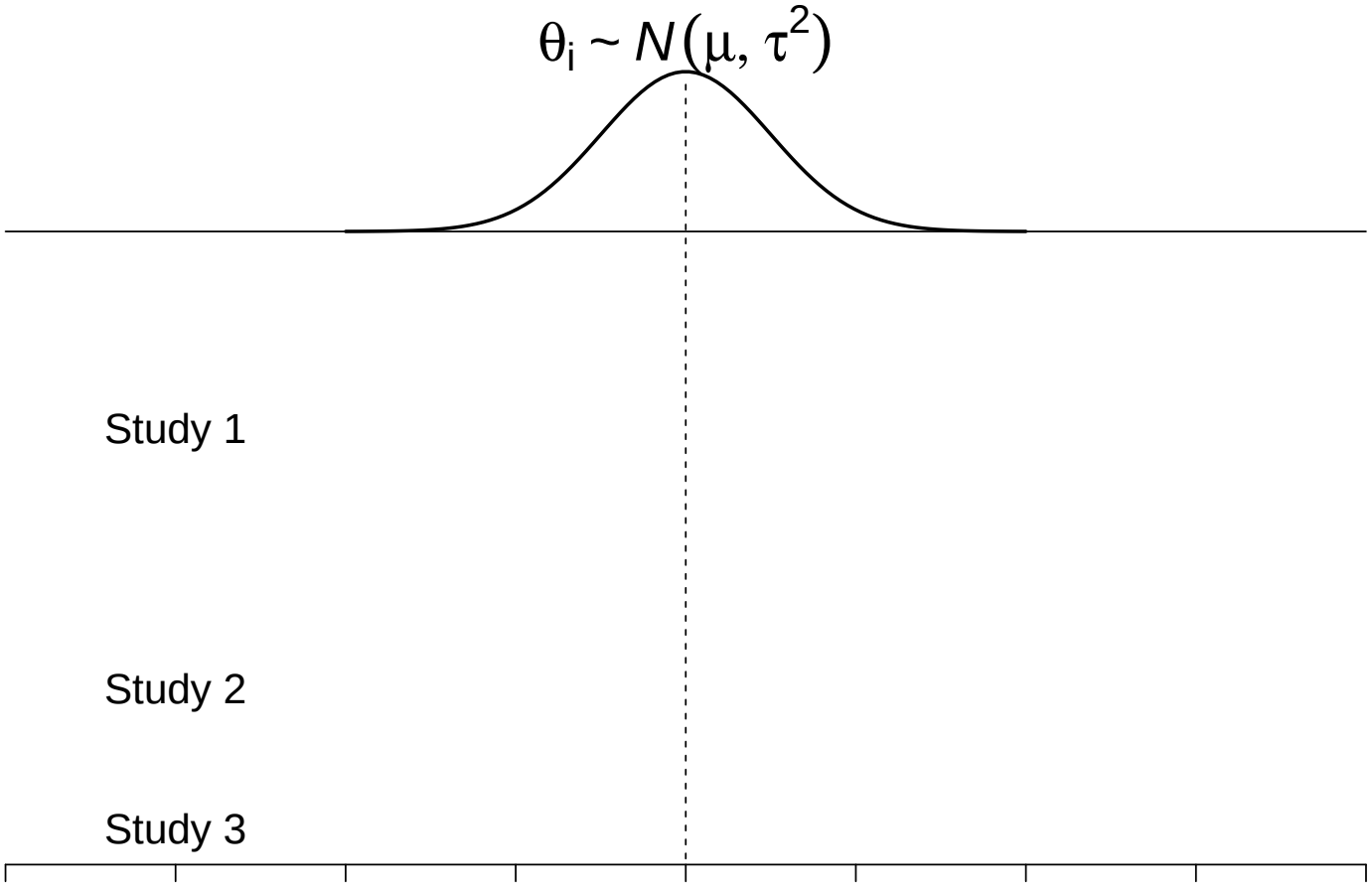
Meta-analysis models: Equal-effects model



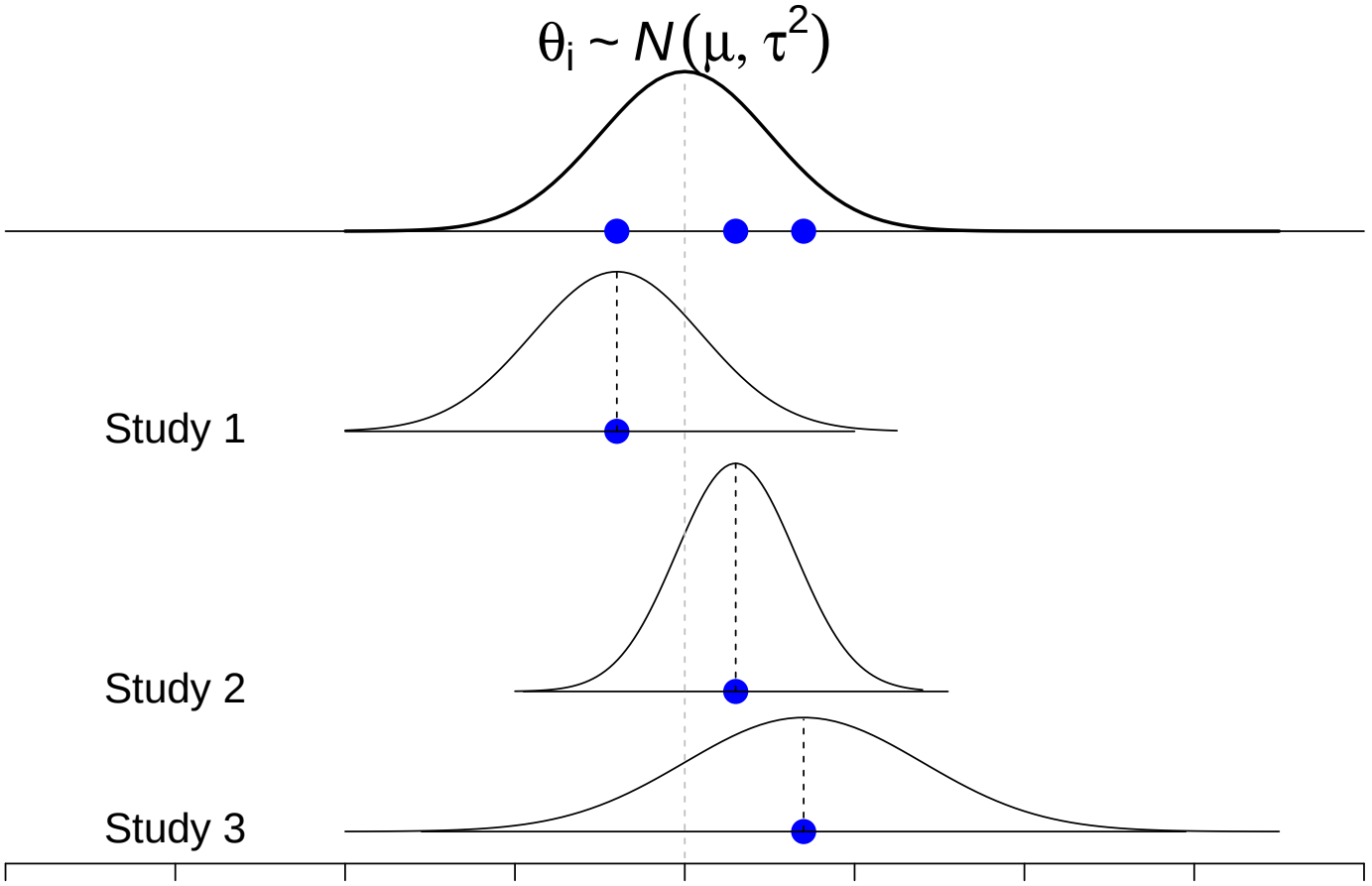
Meta-analysis models: Random-effects model

- Variability in observed effect size estimates (y_i) is caused by:
  1. Sampling variability
  2. Variability among the true effect sizes
  3. Systematic differences among the true effect sizes

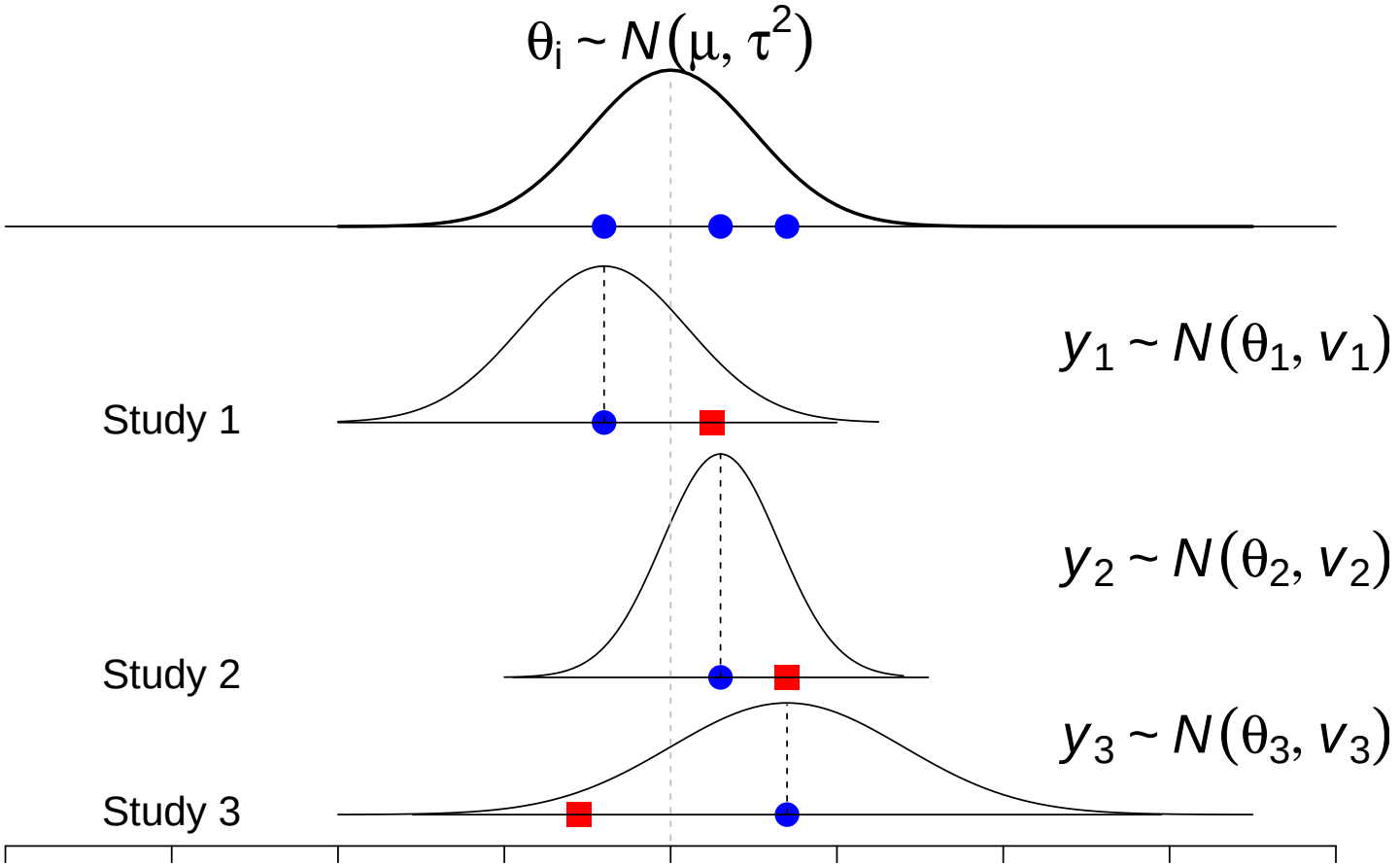
Meta-analysis models: Random-effects model



Meta-analysis models: Random-effects model



Meta-analysis models: Random-effects model



Heterogeneity in true effect size

- Variability in observed effect size estimates (y_i) is caused by:

1. Sampling variability
2. Variability among the true effect sizes
3. Systematic differences among the true effect sizes

- Research questions to study systematic differences in 3.
  - Are there differences in job satisfaction across countries?
  - What is the effect of working from home on job satisfaction?

- Moderator analysis can be used to study whether true effect sizes are related to moderators

Issues with meta-analysis

Issues with meta-analysis: Criticism

- Meta-analysis is *an exercise of mega-silliness* (Eysenck, 1978)

- Meta-analysis is *statistical alchemy for the 21st century* (Feinstein, 1995)

Main criticisms:

- Mixing apples and oranges

Issues with meta-analysis: Criticism

- Meta-analysis is *an exercise of mega-silliness* (Eysenck, 1978)

- Meta-analysis is *statistical alchemy for the 21st century* (Feinstein, 1995)

Main criticisms:

- Mixing apples and oranges

Issues with meta-analysis: Criticism

- Meta-analysis is *an exercise of mega-silliness* (Eysenck, 1978)

- Meta-analysis is *statistical alchemy for the 21st century* (Feinstein, 1995)

Main criticisms:

- Mixing apples and oranges
- Garbage in, garbage out
- Publication bias invalidates meta-analysis $\rightarrow$ yes, but there are methods to correct for this

Issues with meta-analysis: Do I have enough studies?

- You can already apply meta-analysis methods for two studies
  - Equal-effect model $\rightarrow$ more precise effect size estimate
  - Random-effects model $\rightarrow$ be cautious, because of imprecise estimated τ^2

- However, statistical power for testing for no effect and homogeneity will be low

Publication bias

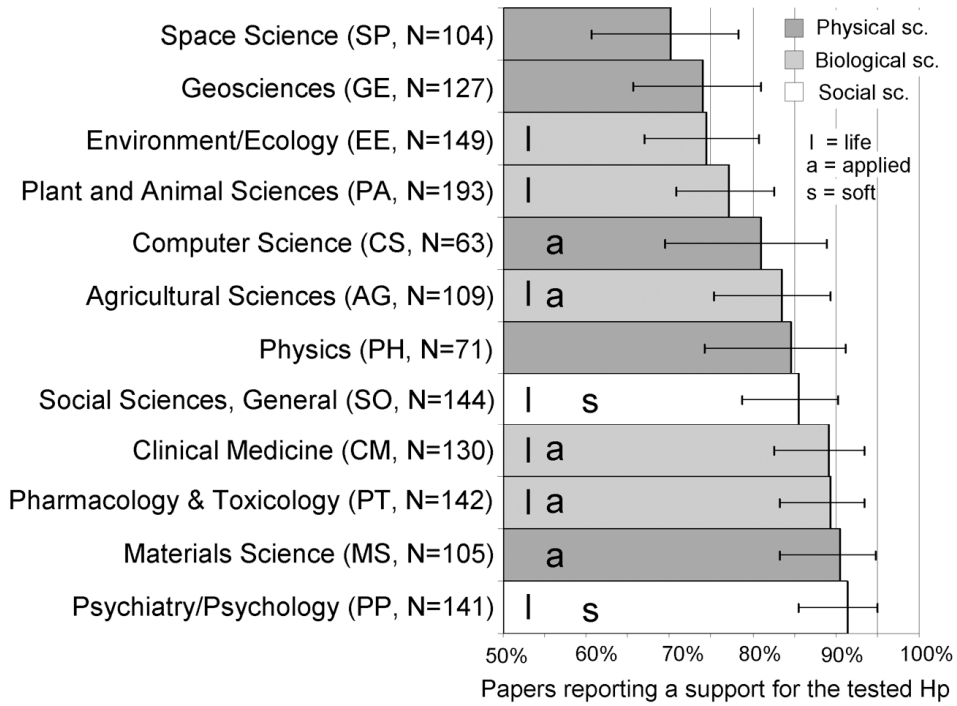
Publication bias

- Publication bias is "the selective publication of studies with a significant outcome"

- ≈90% of main hypotheses are significant in psychology

- But this is not in line with average statistical power (about 20-50%)

- Consequences:
  - Overestimation
  - False impression



Adapted from Fanelli (2010)

Replications and meta-analysis: The problem

- Example of a common problem (independent samples t-test):

	Cohen's d	t-statistic
Original	0.5	$t(78) = 2.24$, $p = .028$
Replication	0.23	$t(170) = 1.5$, $p = .135$

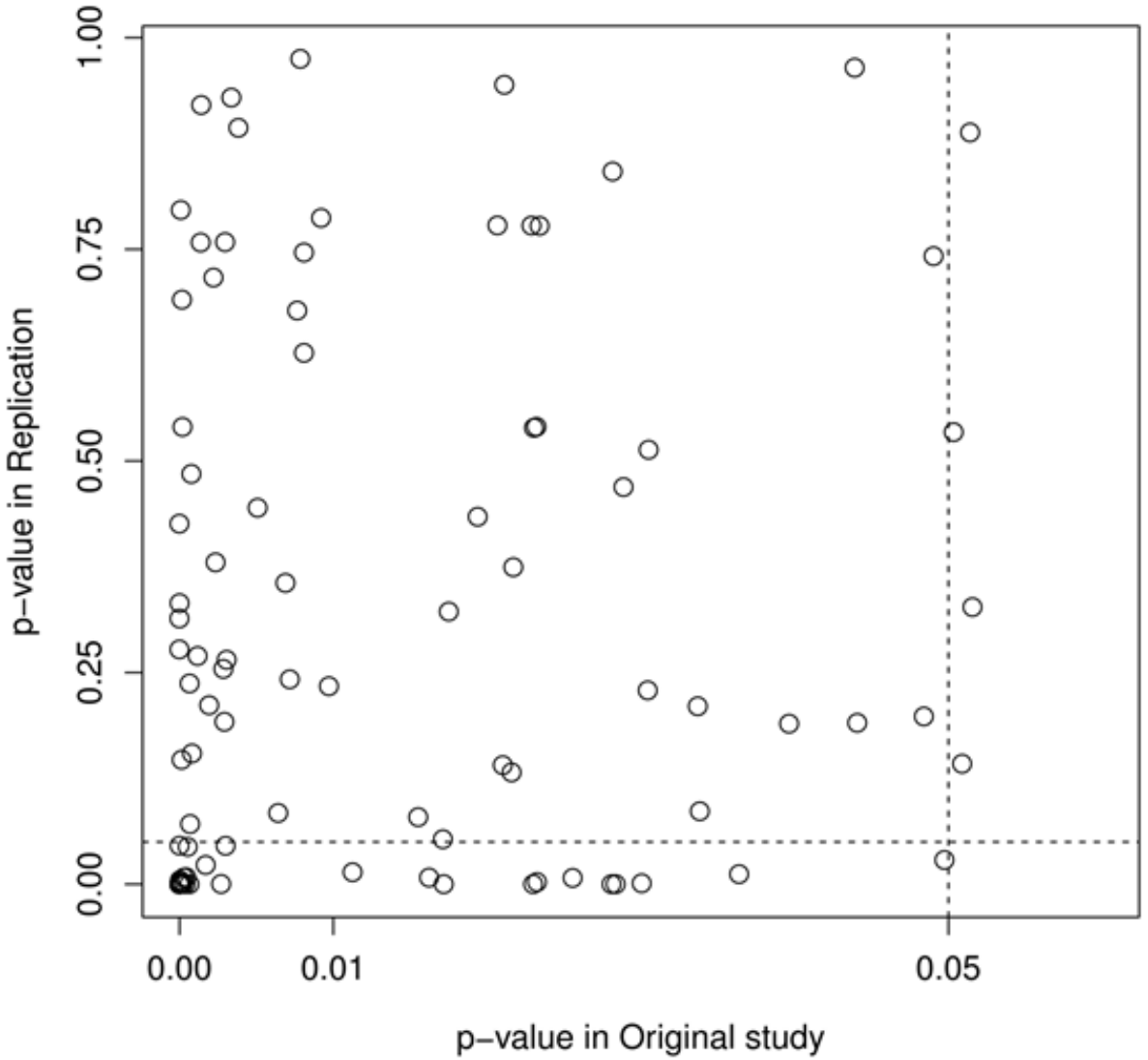
What to conclude?!

Questions considered relevant:

- Does an effect exist? (0 or not)
- What is the magnitude of effect size? (best guess)

Replications and meta-analysis: The problem

- Distribution of *p*-values in Reproducibility Project: Psychology
  - Significant original and nonsignificant replication in 63.9%



Replications and meta-analysis: The problem

- Significant results are overrepresented in the literature
- Published effect sizes are therefore most probably overestimated
- Replicability projects in psychology (RPP) and economics (EE-RP) confirmed that effect sizes are overestimated:
  - RPP: $r = 0.403$ vs. 0.197
  - EE-RP: $r = 0.506$ vs. 0.303
- **Conclusion:** We should take statistical significance of original study into account

Snapshot method

Snapshot method

- **Snapshot** Bayesian Hybrid Meta-Analysis Method
  - Assume four effect sizes (zero, small, medium, large) $\rightarrow$ *snapshots*
- Snapshot **Bayesian** Hybrid Meta-Analysis Method
  - Compute posterior probability of these four effects $\rightarrow$ *Bayesian*
- Snapshot Bayesian **Hybrid** Meta-Analysis Method
  - Take statistical significance of original study into account $\rightarrow$ *hybrid*
- Snapshot Bayesian Hybrid **Meta-Analysis** Method
  - Combine original study with replication $\rightarrow$ *meta-analysis*

Snapshot method: Basic idea

- Density of the replication is "normal" pdf because no selection:

$$f_r = f(y = y_r; \theta)$$

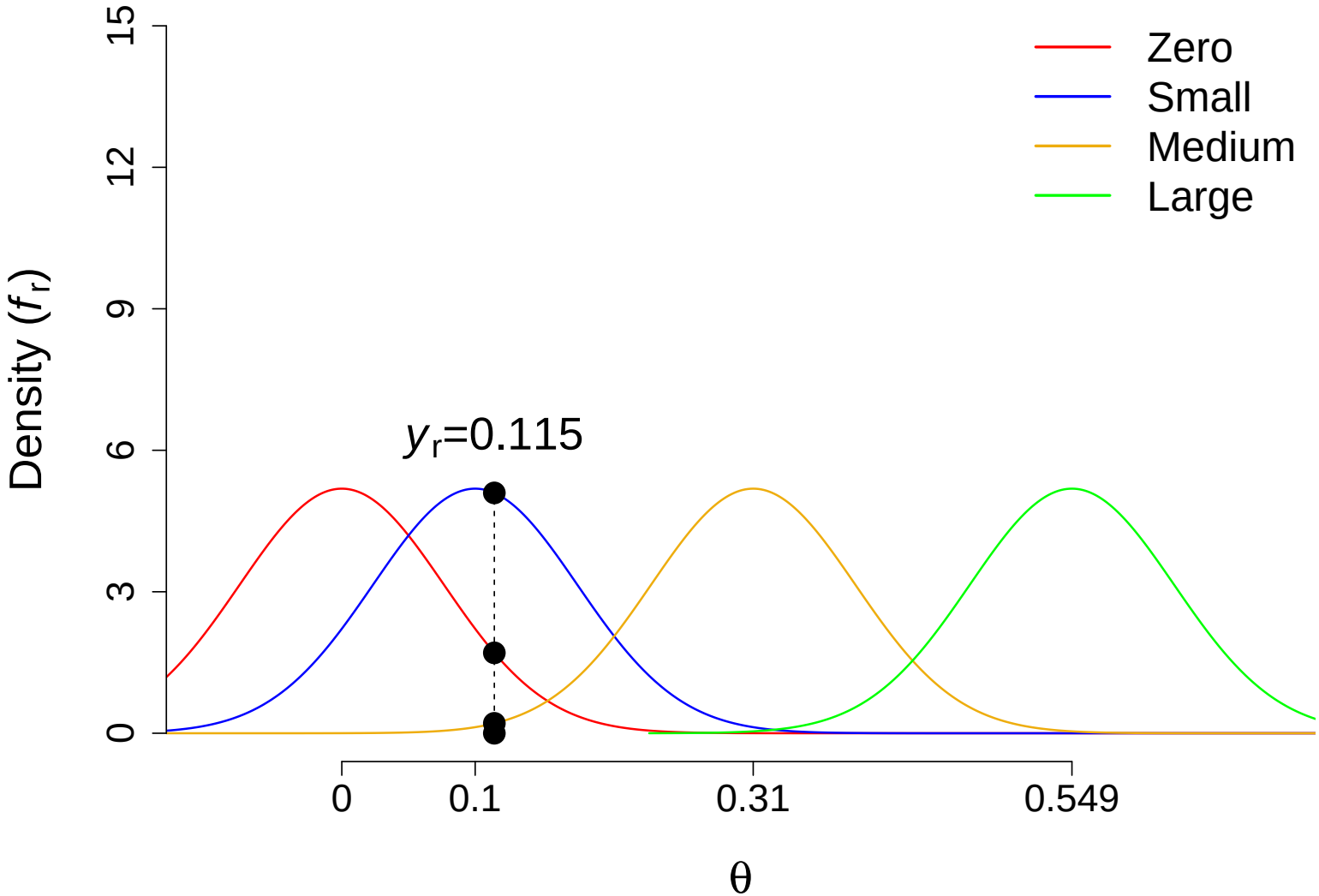
- Density of the original study is pdf *conditional on effect size being statistically significant*:

$$f_o = \frac{f(y = y_o; \theta)}{P(y \geq y_{cv}; \theta)}$$

- Assumptions:
  - Original study is statistically significant
  - Both studies estimate the same effect (fixed-effect)

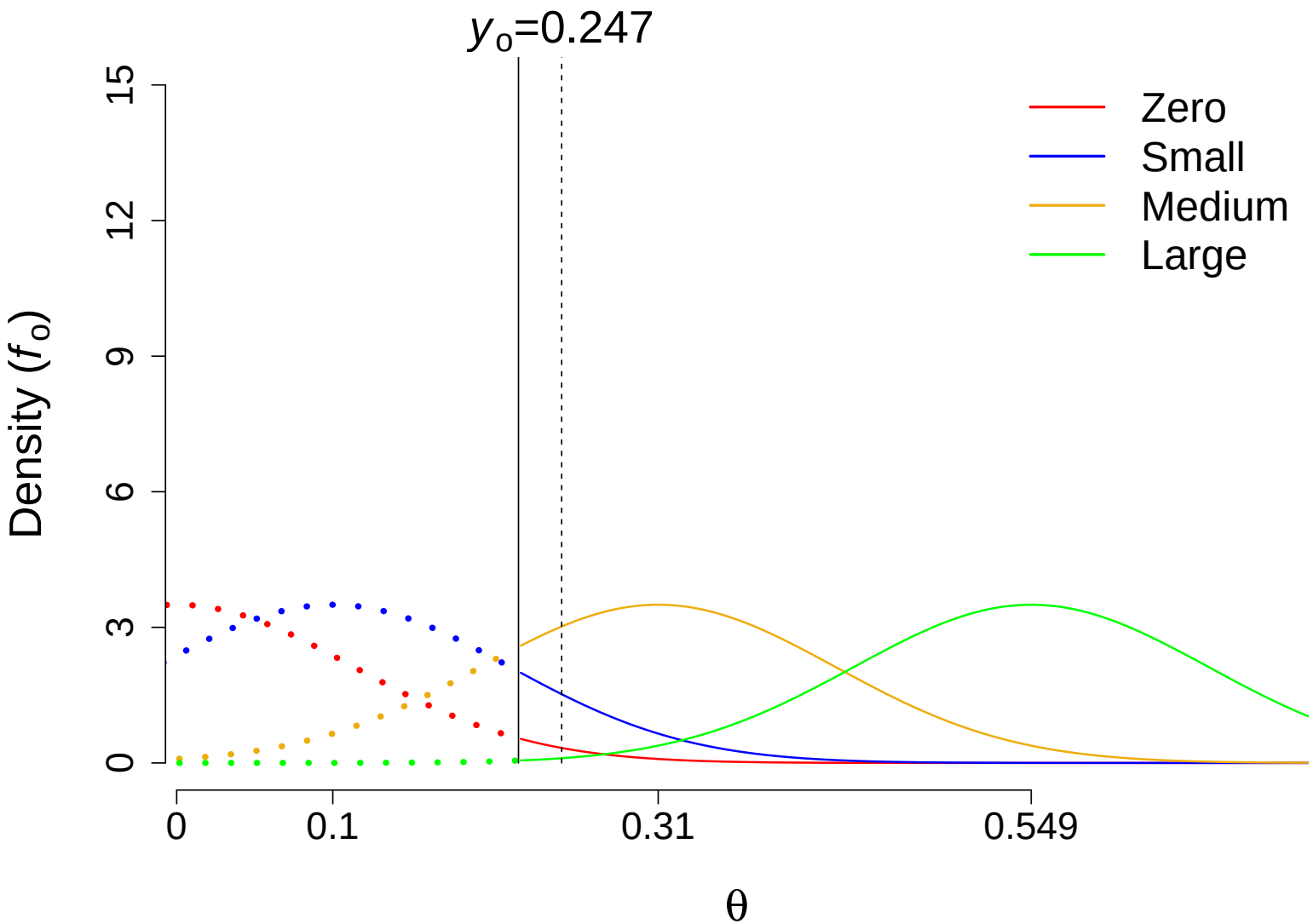
Snapshot method: Basic idea

- Densities replication: $d = 0.23$, $t(170) = 1.5$, $p = 0.135$



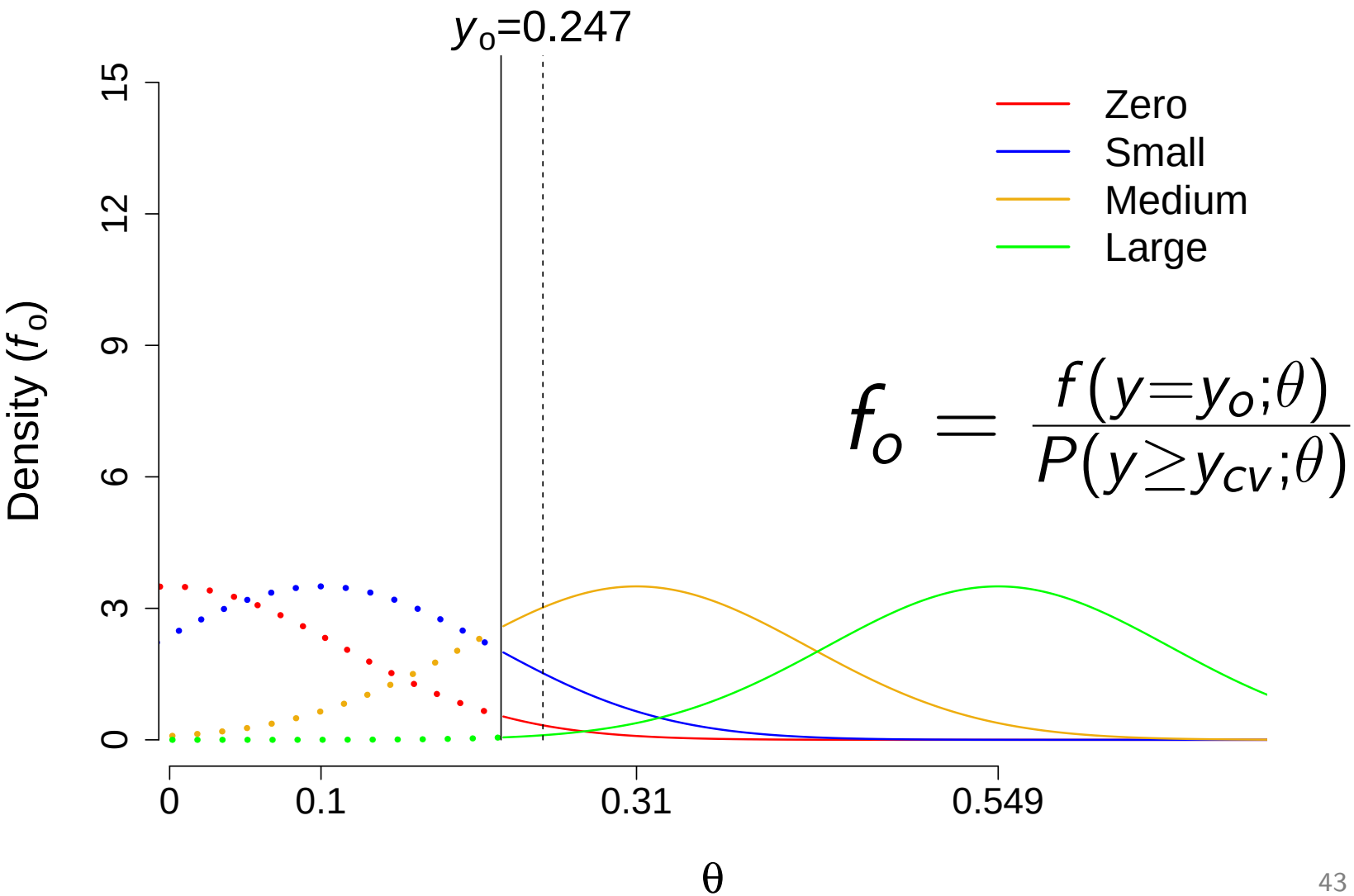
Snapshot method: Basic idea

- Densities original study (naïve): $d = 0.5$, $t(78) = 2.24$, $p = 0.028$



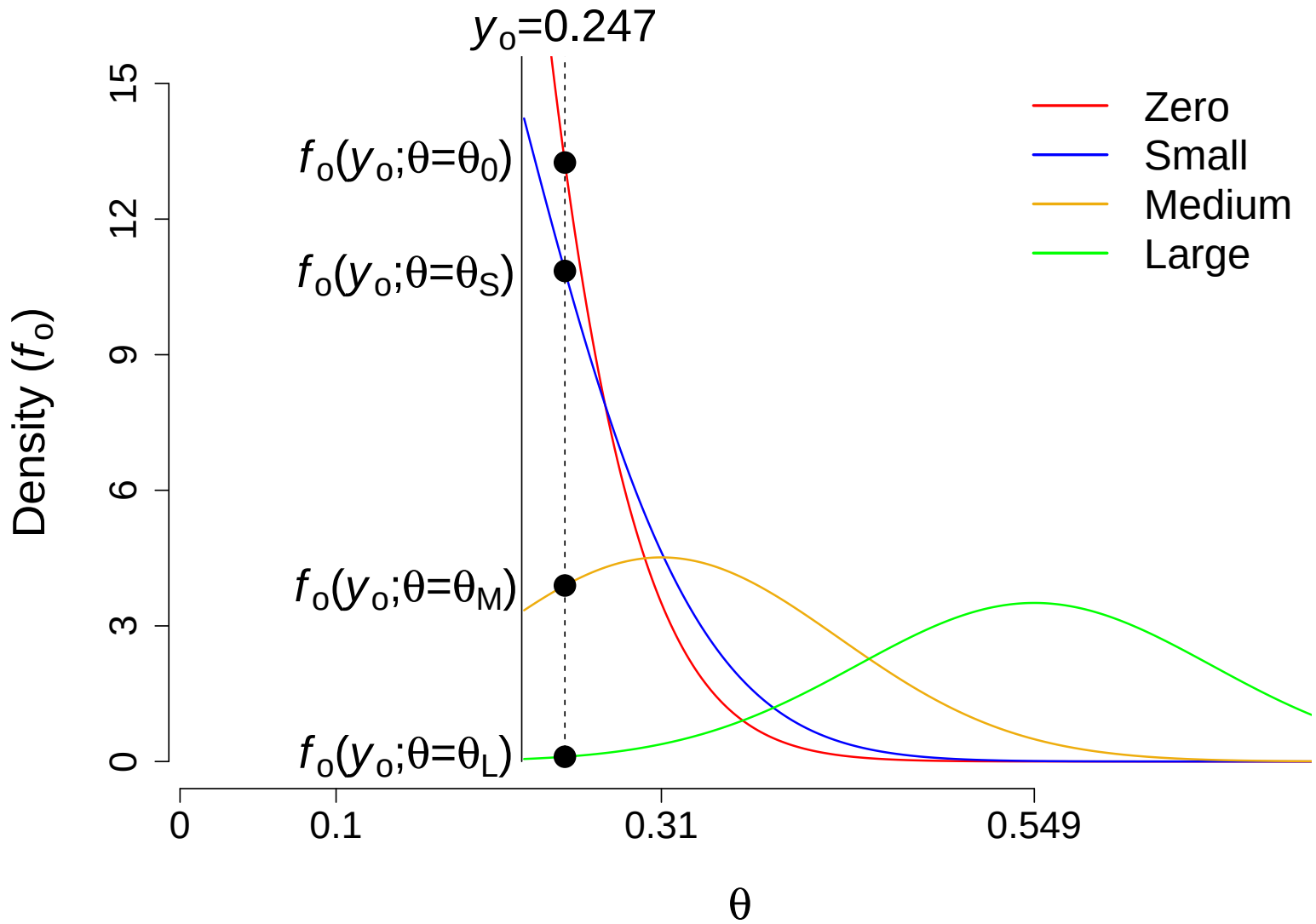
Snapshot method: Basic idea

- Densities original study (naïve): $d = 0.5$, $t(78) = 2.24$, $p = 0.028$



Snapshot method: Basic idea

- Densities original study: $d = 0.5$, $t(78) = 2.24$, $p = 0.028$



Snapshot method: Basic idea

- Combined likelihood:

$$L(\theta) = f_o(\theta) \times f_r(\theta)$$

- Posterior probabilities assuming a uniform prior for each snapshot are computed with:

$$\pi_x = \frac{L(\theta = x)}{L(\theta = \theta_0) + L(\theta = \theta_S) + L(\theta = \theta_M) + L(\theta = \theta_L)}$$

- Advantages of the method:
  - Easy and insightful
  - Easy (re)computation posterior for other (than uniform) prior:

$$\pi_x^* = \frac{p_x \pi_x}{p_0 \pi_0 + p_S \pi_S + p_M \pi_M + p_L \pi_L}$$

Snapshot method: Example

- Example of a common problem (independent samples t-test):

	Cohen's d	t-statistic
Original	0.5	$t(78) = 2.24$, $p = .028$
Replication	0.23	$t(170) = 1.5$, $p = .135$

- Applying snapshot method:

	Zero	Small	Medium	Large
Naïve	0.063	0.866	0.071	0
Snapshot				

Snapshot method: Example

- Example of a common problem (independent samples t-test):

	Cohen's d	t-statistic
Original	0.5	$t(78) = 2.24$, $p = .028$
Replication	0.23	$t(170) = 1.5$, $p = .135$

- Applying snapshot method:

	Zero	Small	Medium	Large
Naïve	0.063	0.866	0.071	0
Snapshot	0.287	0.703	0.01	0

- Evidence of zero effect increased; best guess = small effect

Application: RPP and EE-RP

- Initiatives to study the replicability of psychological and economic research
- **RPP:** Studies from JPSP, Psychological Science, and Journal of Experimental Psychology: 67 out of 100 studies were included
- **EE-RP:** Experimental research from the American Economic Review and Quarterly Journal of Economics: 16 out of 18 studies were included
- "High-powered" replication of a key effect

Application: RPP and EE-RP

- Probability of strong evidence (π_x >.75; BF >3) using snapshot method

	Zero	Small	Medium	Large	Unknown
EE-RP	0	0.062	0.312	0.438	0.188
RPP	0.134	0.03	0.045	0.164	0.627

- **Conclusions:**
  - Studied effects larger in EE-RP than in RPP
  - Only few studies have strong evidence for zero effect in RPP (13.4%)
  - Often not enough information for determining magnitude of effect size in RPP (62.7%)

Conclusion

- Meta-analysis is a powerful tool to synthesize studies
- Meta-analysis helps researchers coping with the information explosion
- Model selection is a complicated but very important part of meta-analysis
- Methods *should* take statistical significance of original study into account

Further reading on meta-analysis

- Borenstein et al. (2009). Introduction to meta-analysis
- Cooper et al., (2019). The handbook of research synthesis and meta-analysis
- Cooper (2010). Research synthesis and meta-analysis: A step-by-step approach

Thank you for your attention

www.robbievanaert.com

www.metaresearch.nl

PDF of dissertation:

https://osf.io/preprints/metaarxiv/eqhjd/