



Research Synthesis after the Reproducibility and Open Science Reforms: How do Systematic Reviews and Meta-Analysis Practices in Education Need to Adapt?

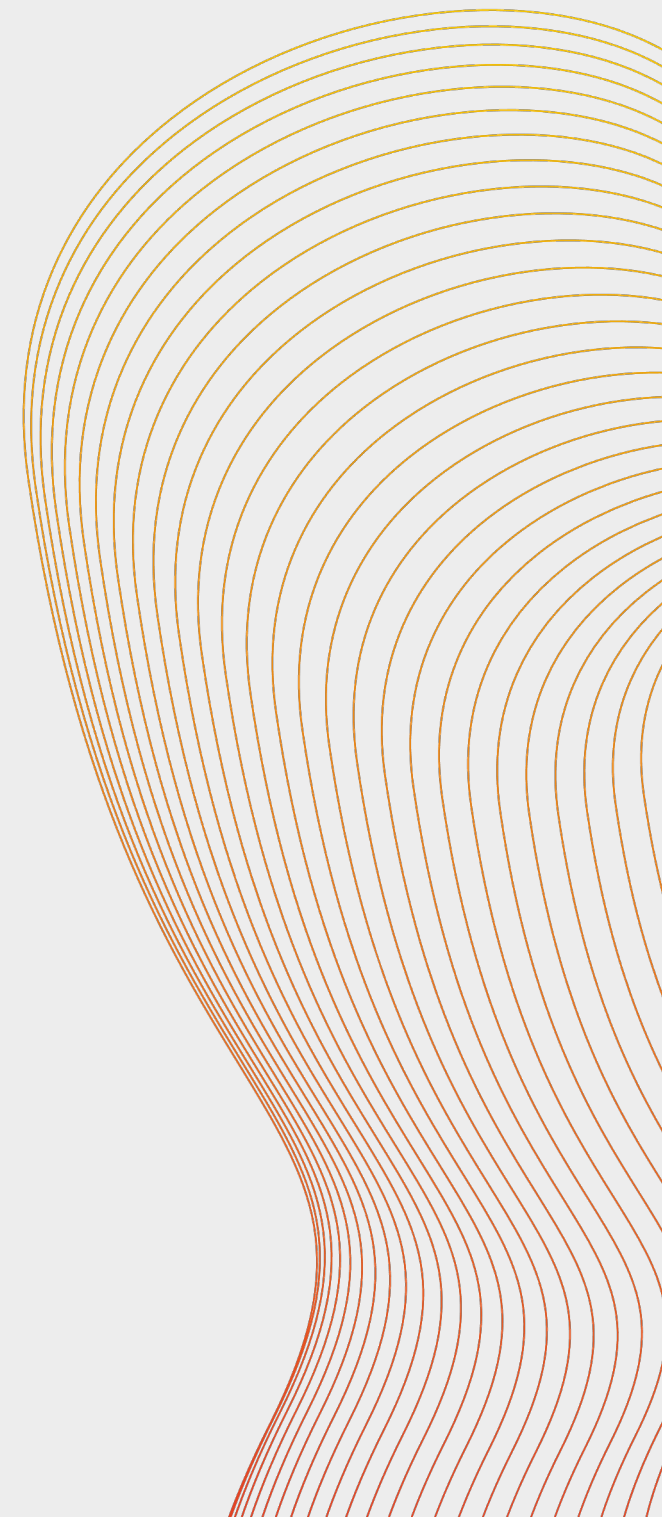
Diego G. Campos

Centre for Educational Measurement (CEMO), Faculty of Educational Sciences, University of Oslo, Norway

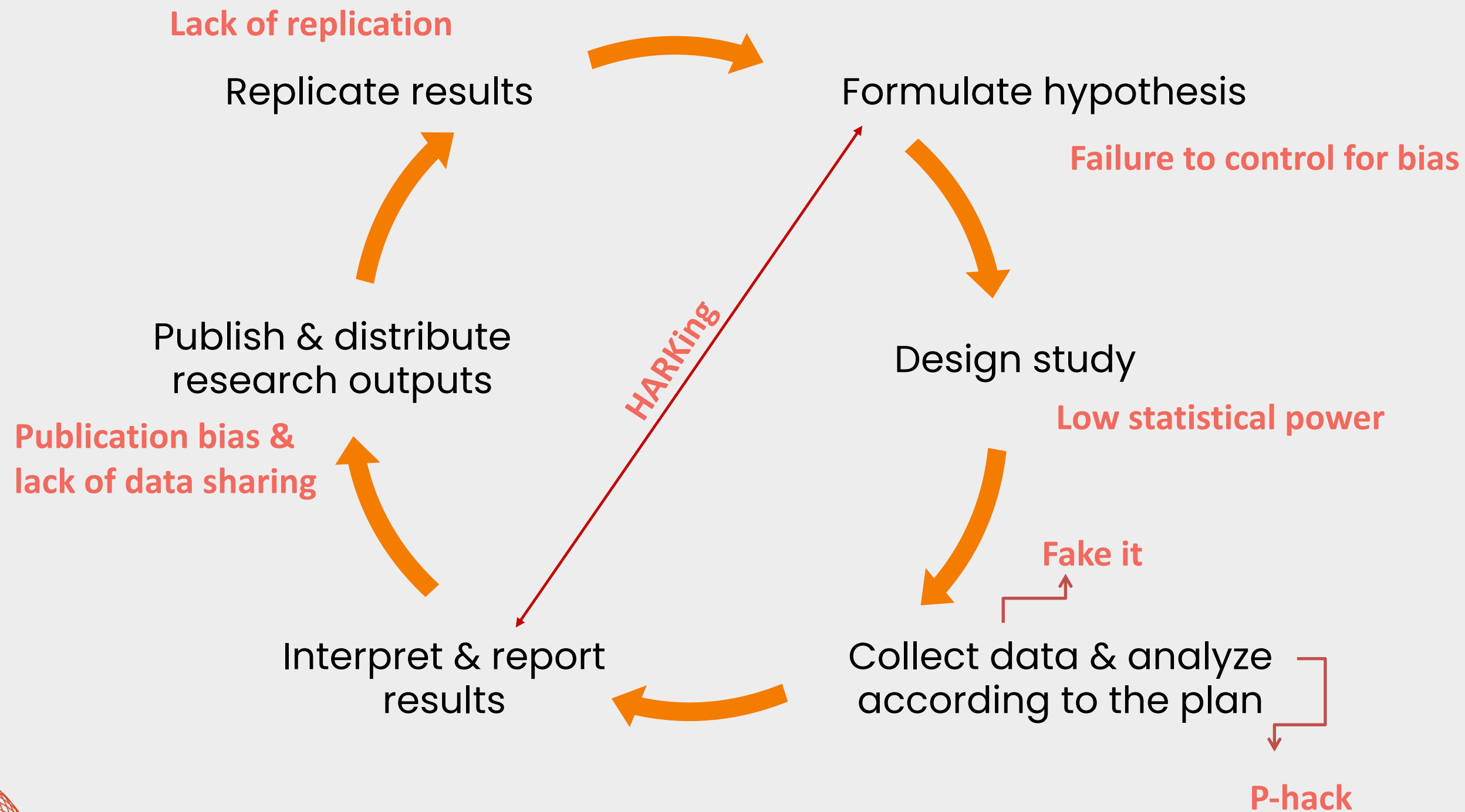


Research Synthesis after the Reproducibility and Open Science Reforms: How do Systematic Reviews and Meta-Analysis Practices in Education Need to Adapt?

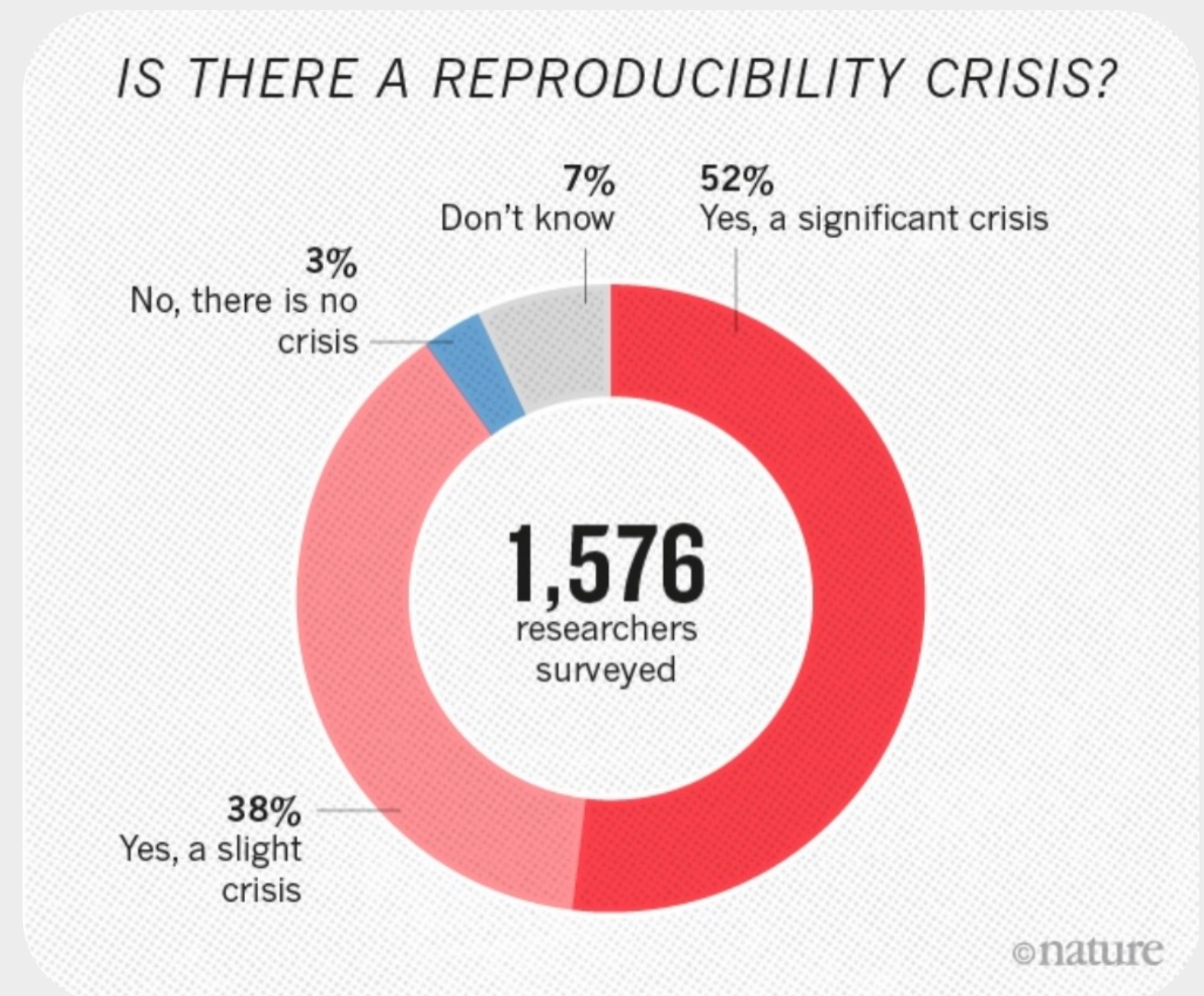
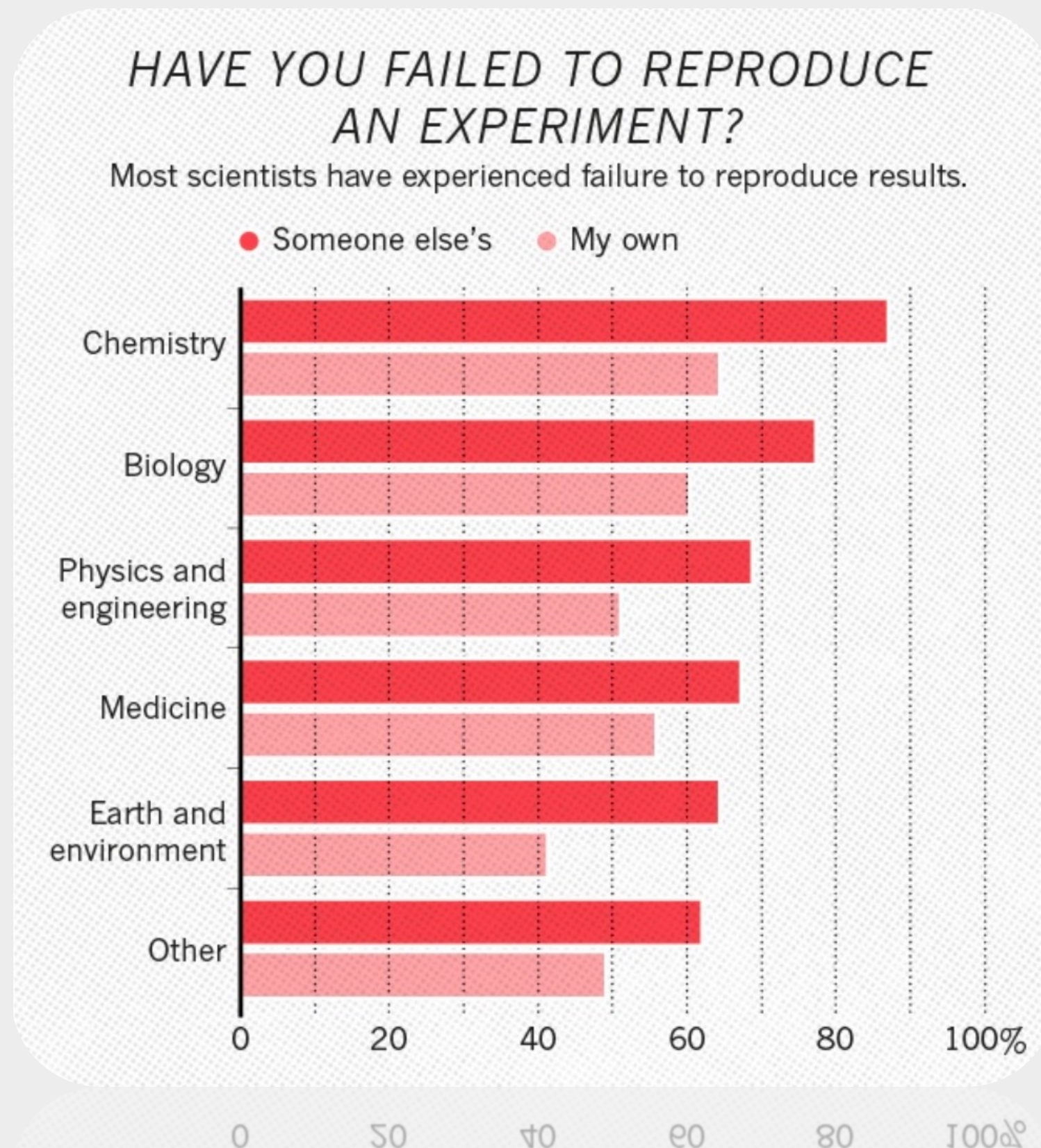
- Reproducibility and Open Science
- Are research synthesis reproducible?
- Adapting research synthesis to open science
- Challenges and future directions
- Conclusions



Questionable Research Practices



Reproducibility and Open Science



Baker, M. (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533, 452–454 (2016).
<https://doi.org/10.1038/533452a>

Reproducibility and Open Science

Reproducibility refers to testing the reliability of a prior finding using the same data and analysis strategy

- Process reproducibility
- Outcome reproducibility

Open science promotes the sharing of research data, materials, and methods to foster collaboration, scrutiny, and advancement of knowledge

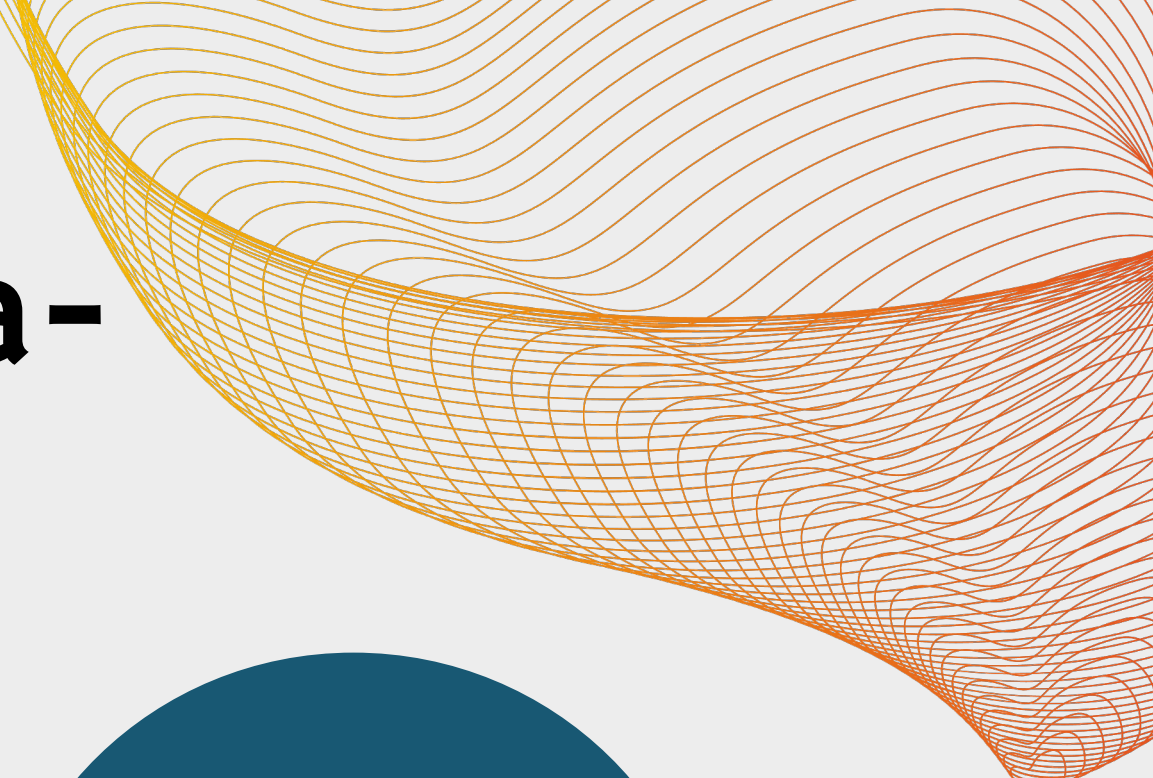




**What is the status of
reproducibility and open
science in systematic reviews?**



Systematic Reviews and Meta-Analysis



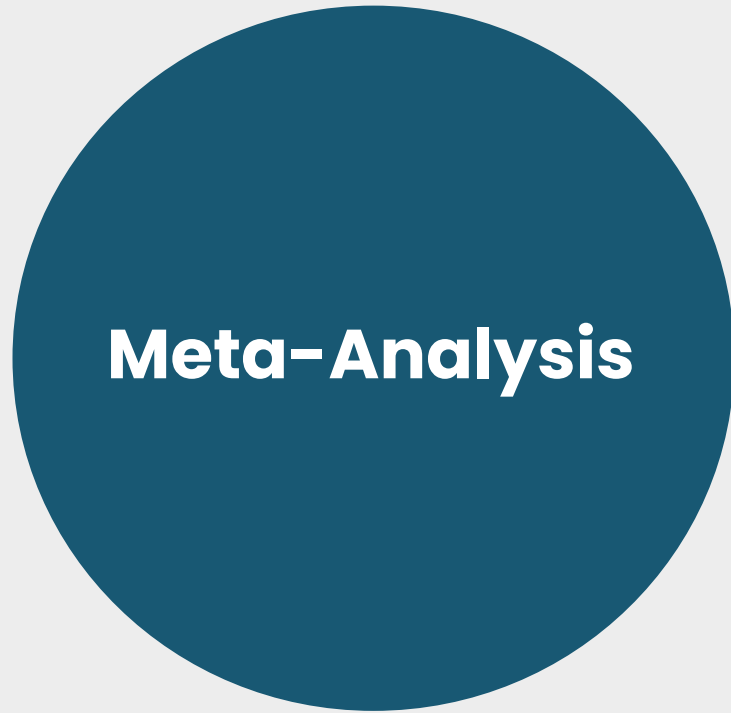
Research Synthesis

Summarize past empirical research that addresses related hypotheses for drawing generalizations



Systematic Review

A set of systematic and replicable search methods used to identify a body of relevant evidence



Meta-Analysis

A set of statistical models used to combine quantitative information from multiple sources of evidence



Systematic Reviews and Meta-Analysis



Inform evidence-based practice, policy guidance, and future research directions

Homework

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, 76(1), 1–62. <https://doi.org/10.3102/00346543076001001>

Class size

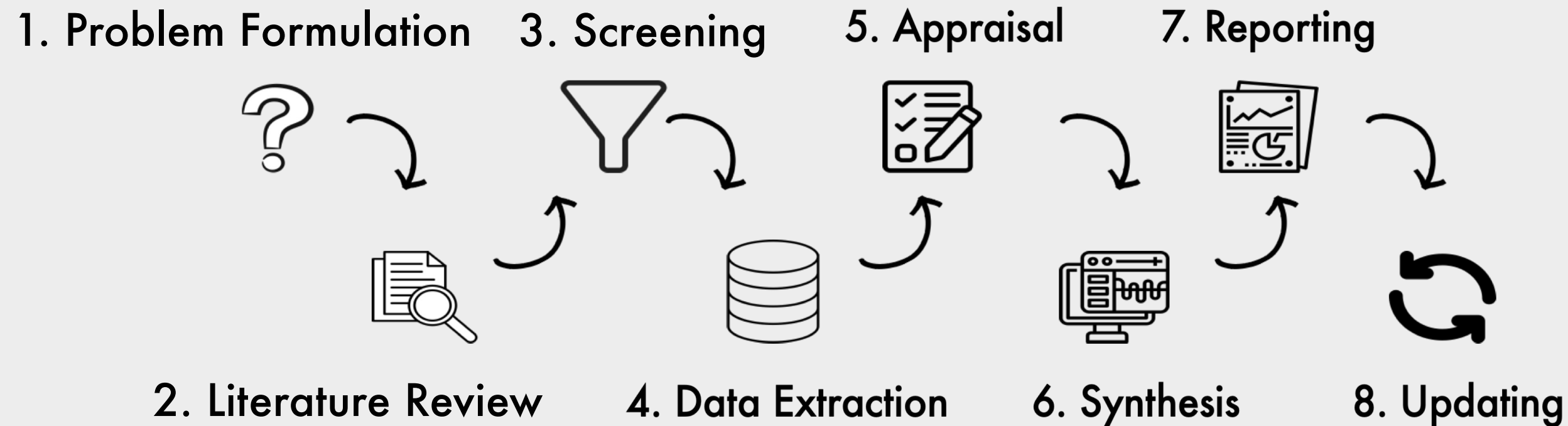
Filges, T., Sonne-Schmidt, C. S., & Nielsen, B. C. V. (2018). Small class sizes for improving student achievement in primary and secondary schools: A systematic review. *Campbell Systematic Reviews*, 14(1), 1–107. <https://doi.org/10.4073/csr.2018.10>

A low-angle photograph looking up at a tree with many small, light-colored blossoms, possibly cherry blossoms, against a bright, overcast sky. The branches are dark and intricate, creating a complex web of lines. The blossoms are small and numerous, creating a soft, textured background.

The Garden of Forking Paths

Photo by [Robin Huang](#)

Are Systematic Reviews and Meta-Analyses Reproducible?



Are Systematic Reviews and Meta-Analyses Reproducible?

Literature Review

Search Procedures



Elson, M., & Ferguson, C. J. (2014). Twenty-five years of research on violence in digital games and aggression: Empirical evidence, perspectives, and a debate gone astray. *European Psychologist*, 19(1), 33–46. <https://doi.org/10.1027/1016-9040/a000147>

Bushman, B. J., & Huesmann, L. R. (2014). Twenty-five years of research on violence in digital games and aggression revisited: A reply to Elson and Ferguson (2013). *European Psychologist*, 19(1), 47–55. <https://doi.org/10.1027/1016-9040/a000164>

Are Systematic Reviews and Meta-Analyses Reproducible?

Screening

Inclusion/Exclusion criteria



Anderson CA, Shibuya A, Ihori N, Swing EL, Bushman BJ, Sakamoto A, Rothstein HR, Saleem M.. Violent video game effects on aggression, empathy, and prosocial behaviour in Eastern and Western countries: A meta-analytic review. Psychol Bull. 2010;136:151–73. doi:10.1037/a0018251.24

Are Systematic Reviews and Meta-Analyses Reproducible?

Data Extraction

Coding mistakes



Gøtzsche, P. C., Hróbjartsson, A., Maric, K., & Tendal, B. (2007). Data extraction errors in meta-analyses that use standardized mean differences. *JAMA*, 298(4), 430–437. <https://doi.org/10.1001/jama.298.4.430>

Are Systematic Reviews and Meta-Analyses Reproducible?

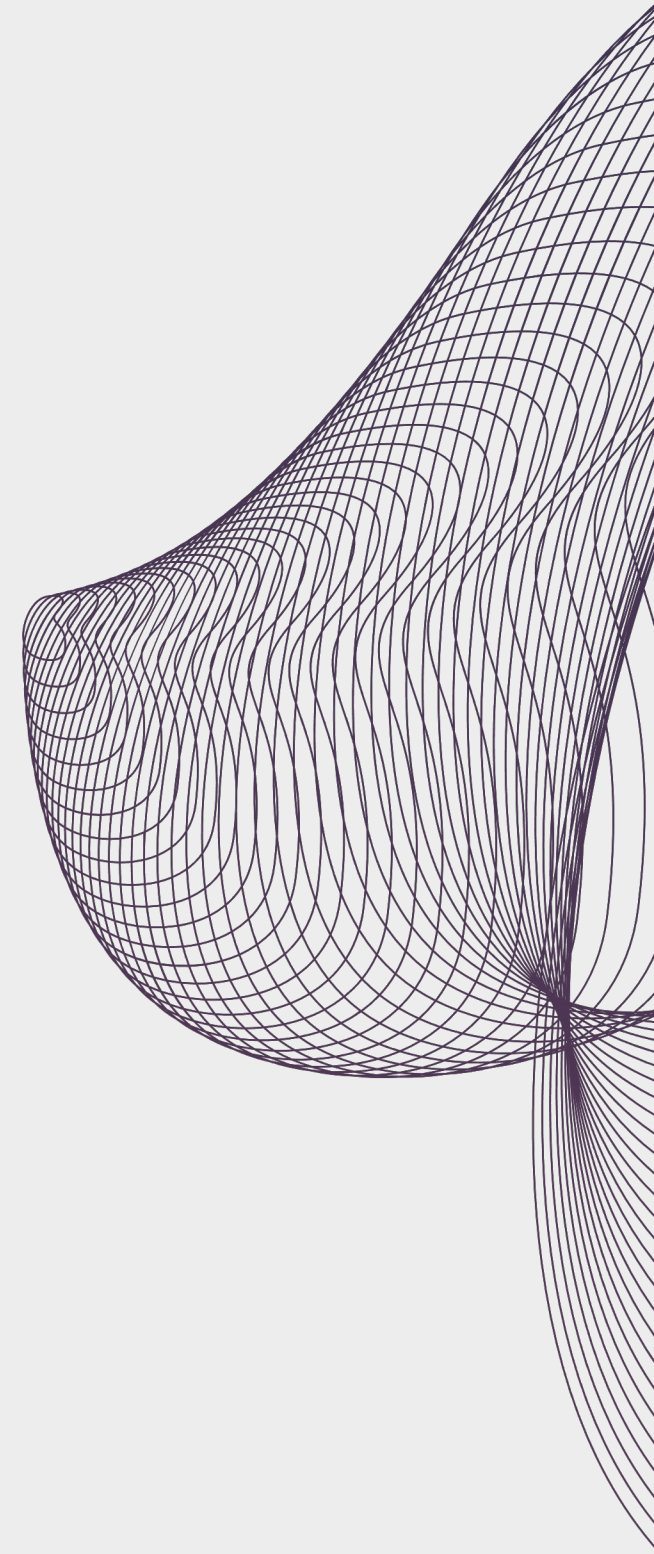
Synthesis

Analyses plan



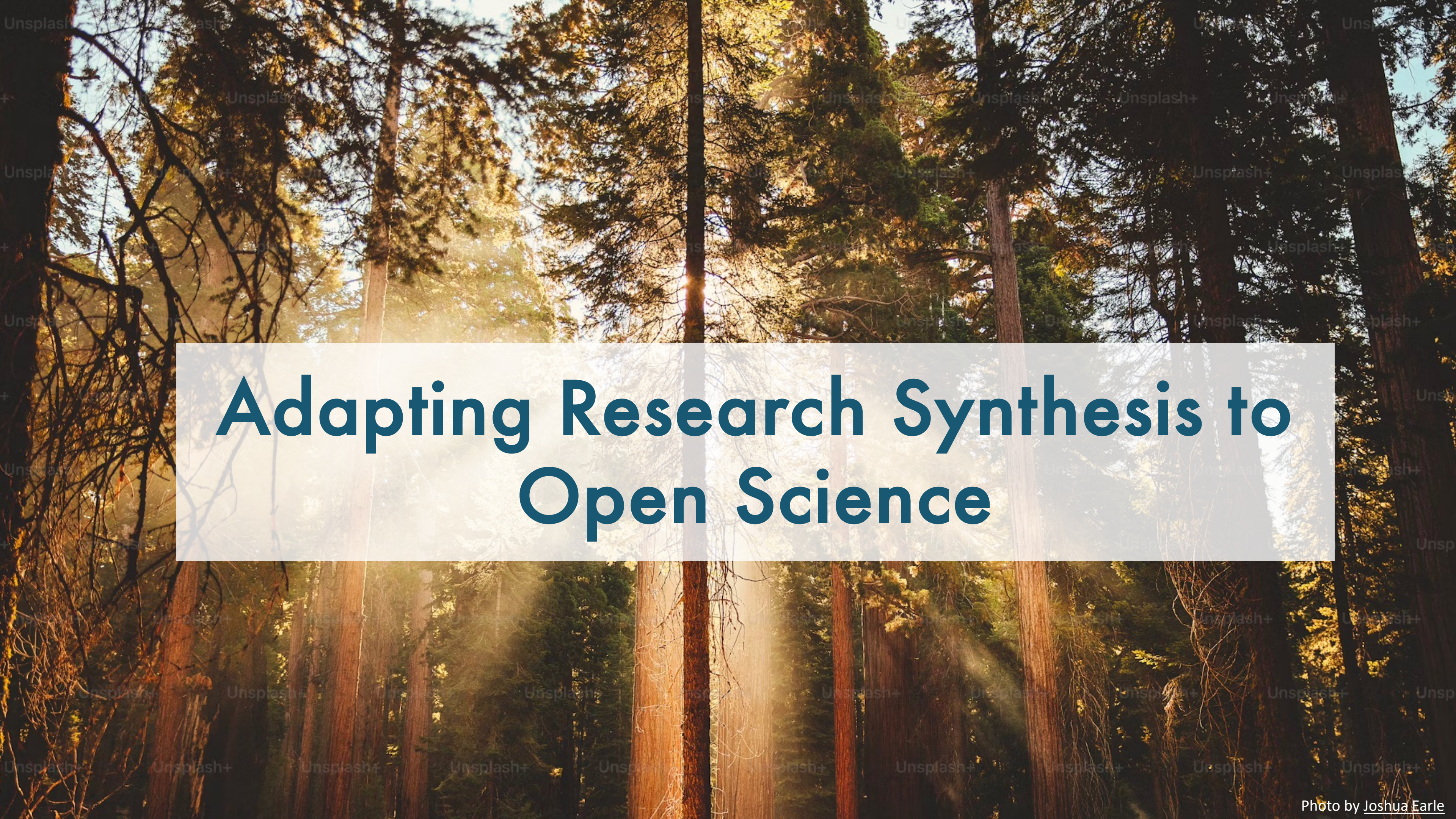
Pietschnig, J., Gerdesmann, D., Zeiler, M., & Voracek, M. (2022). Of differing methods, disputed estimates and discordant interpretations: the meta-analytical multiverse of brain volume and IQ associations. *R. Soc. Open Sci.*9211621

$r = 0.10$ to 0.37



“It may be that meta-analyses are not capable of answering big questions due to issues of methodological flexibility which introduce exactly the type of subjectivity they were designed to avoid.”

(Fergusson, 2014, p.1)



Adapting Research Synthesis to Open Science

Photo by [Joshua Earle](#)

Improving Reproducibility in Systematic Reviews and Meta-Analysis

MARS



APA Style JARS
Journal Article Reporting Standards

JARS–Quant | Table 9
Quantitative Meta-Analysis Article Reporting Standards
Information Recommended for Inclusion in Manuscripts Reporting Quantitative Meta-Analyses

Title and Title Page	
Title	State the research question and the type of research synthesis (e.g., narrative synthesis, meta-analysis).
Author Note	
- List all sources of monetary and in-kind funding support; state the role of funders in conducting the synthesis and deciding to publish the results, if any. - Describe possible conflicts of interest, including financial and other nonfinancial interests. - Give the place where the synthesis is registered and its registry number, if registered. - Provide the name, affiliation, and e-mail address of the corresponding author.	
Abstract	
Objectives	State the research problems, questions, or hypotheses under investigation.
Eligibility Criteria	
Describe the characteristics for inclusion of studies, including independent variables (treatments, interventions), dependent variables (outcomes, criteria), and eligible study designs.	
Methods of Synthesis	
- Describe the methods for synthesizing study results, including - statistical and other methods used to summarize and to compare studies - specific methods used to integrate studies if a meta-analysis was conducted (e.g., effect-size metric, averaging method, the model used in homogeneity analysis)	
Results	
- State the results of the synthesis, including - number of included studies and participants, and their important characteristics - results for the primary outcome(s) and moderator analyses - effect size(s) and confidence interval(s) associated with each analysis if a meta-analysis was conducted	
Conclusions	
Describe strengths and limitations of the evidence, including evidence of inconsistency, imprecision, risk of bias in the included studies, and risk of reporting biases.	

© 2020 American Psychological Association

www.apastyle.org/jars

JARS–Quant | Table 9 | Page 1 of 3

PRISMA



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
Reporting bias assessment	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	

Improving Reproducibility in Systematic Reviews and Meta-Analysis



A Reproducible Meta-Analysis of the Effect of Violent Video Game Play on Aggressive Behaviour

**Information
sources**

**Search
strategy**

**Eligibility
criteria**

Improving Reproducibility in Systematic Reviews and Meta-Analysis



Data extraction

Coding	Reference	NMale	NFemale	Experimental	SMD
0	Blackwell et al. (2007)	44	50	1	0,3
1	Blackwell et al. (2007)	44	50	1	0,4
2	Blackwell et al. (2007)	44	50	1	0,2

Analyses

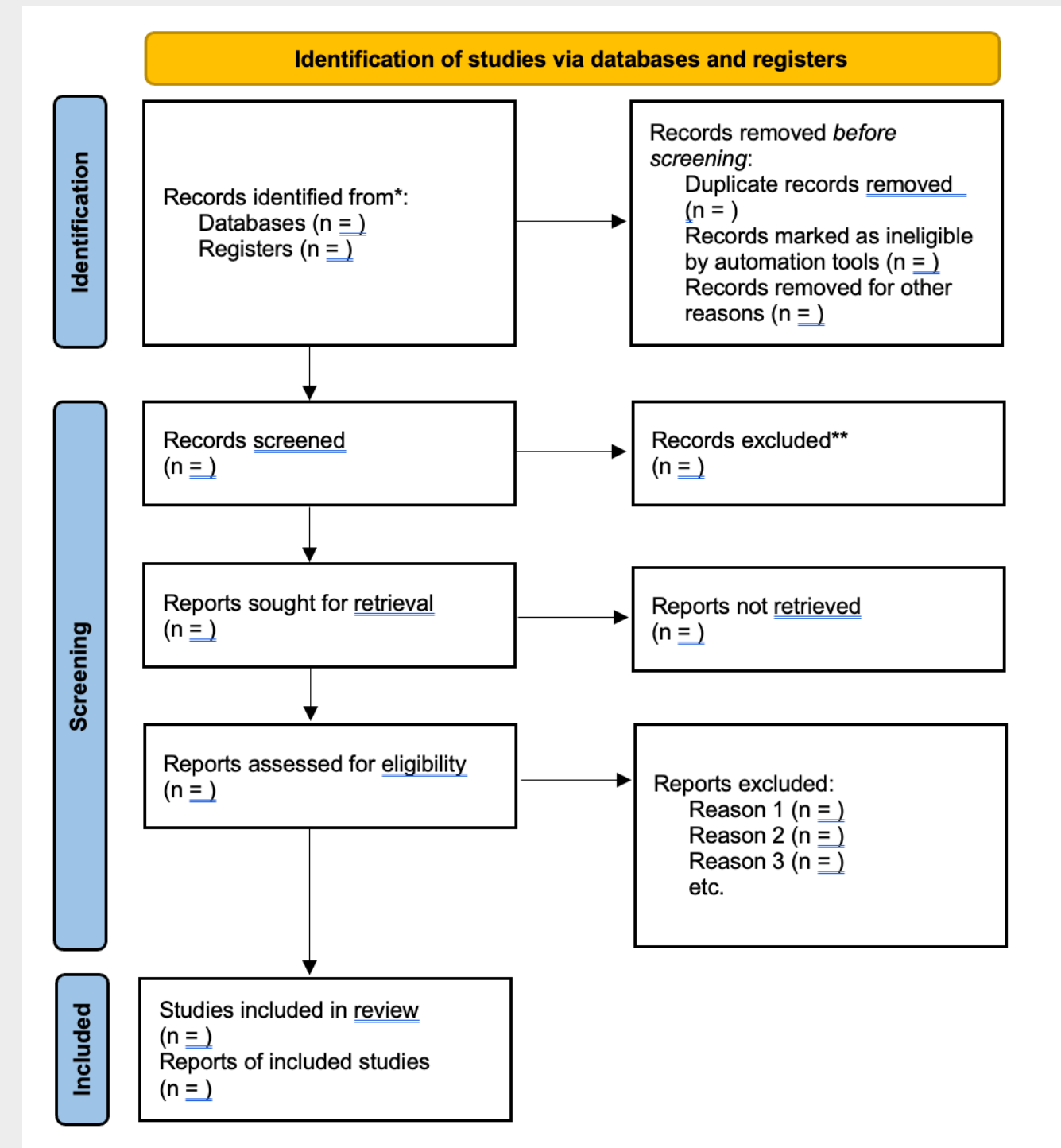
Fixed-effects

Random-effects

Effect size dependencies

Improving Reproducibility in Systematic Reviews and Meta-Analysis

Improving the reporting standards in systematic reviews and meta-analyses can significantly enhance the reproducibility of research findings



Open Science Reforms

Protocols and Pre-Registration

The **study eligibility, screening, and coding procedures** as well as the **analytic model** should be specified before conducting the systematic review and meta-analysis

RESEARCH ARTICLE

Open Access

Analysis of decisions made in meta-analyses of depression screening and the risk of confirmation bias: A case study

Felicity A Goodyear-Smith^{1*}, Mieke L van Driel^{2,3,4}, Bruce Arroll¹ and Chris Del Mar⁴

- Reduces analytical flexibility
- Get early feedback



Open Materials

Making components of the research methodology needed to reproduce the reported procedure and analysis publicly available

- **Search Syntax**
- **Coding schemes**
- **Analyses scripts**

RESEARCH ARTICLE

Open Access

Analysis of decisions made in meta-analyses of depression screening and the risk of confirmation bias: A case study

Felicity A Goodyear-Smith^{1*}, Mieke L van Driel^{2,3,4}, Bruce Arroll¹ and Chris Del Mar⁴



Open Data

- Erroneous calculations of effect sizes
- Inconsistent calculation of effect sizes
- Incorrect inclusion of effect sizes
- Enable researchers to detect outcome reproducibility failures
- Enable re-use of data to update a review or answer other research questions
- Enables cumulative meta-analysis

Search
Results

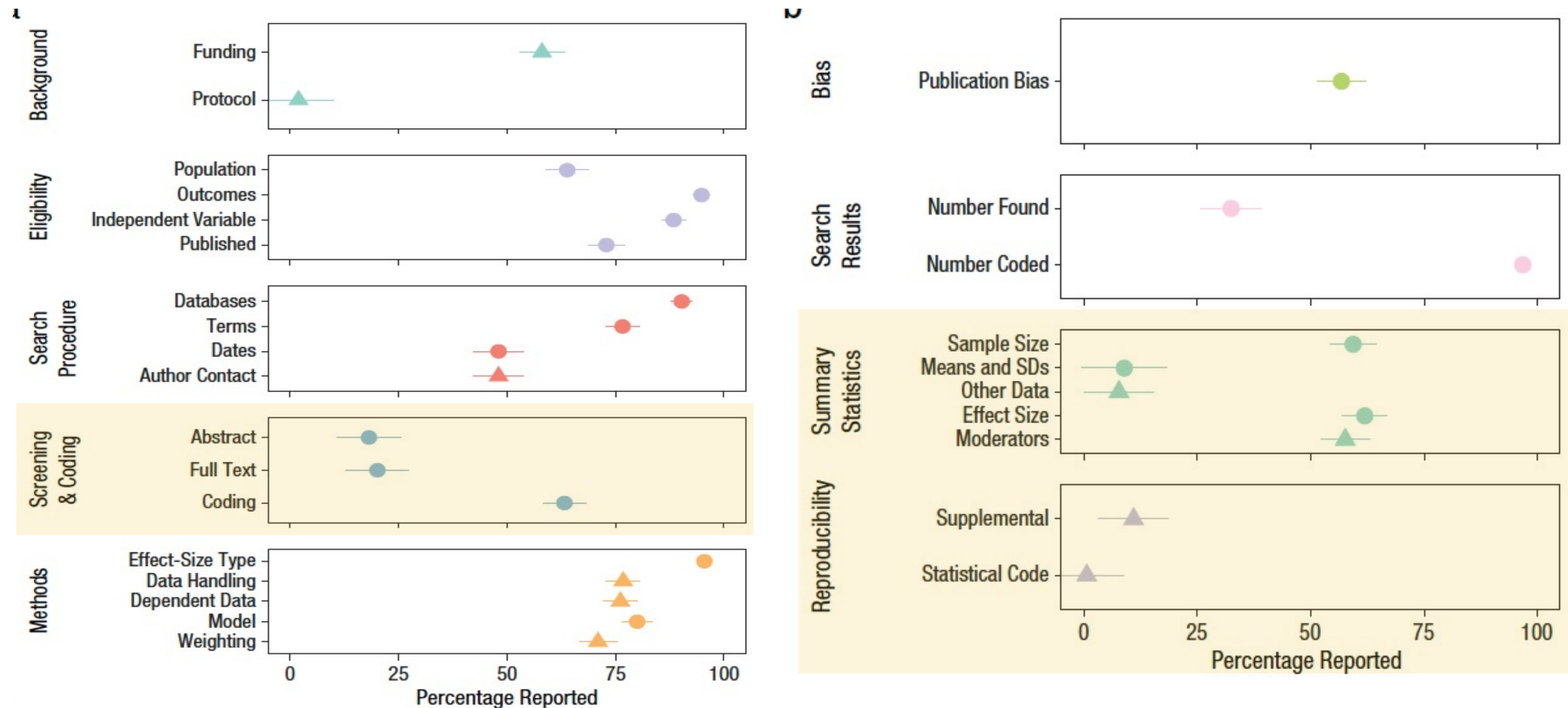
Meta-
Analytic
Data





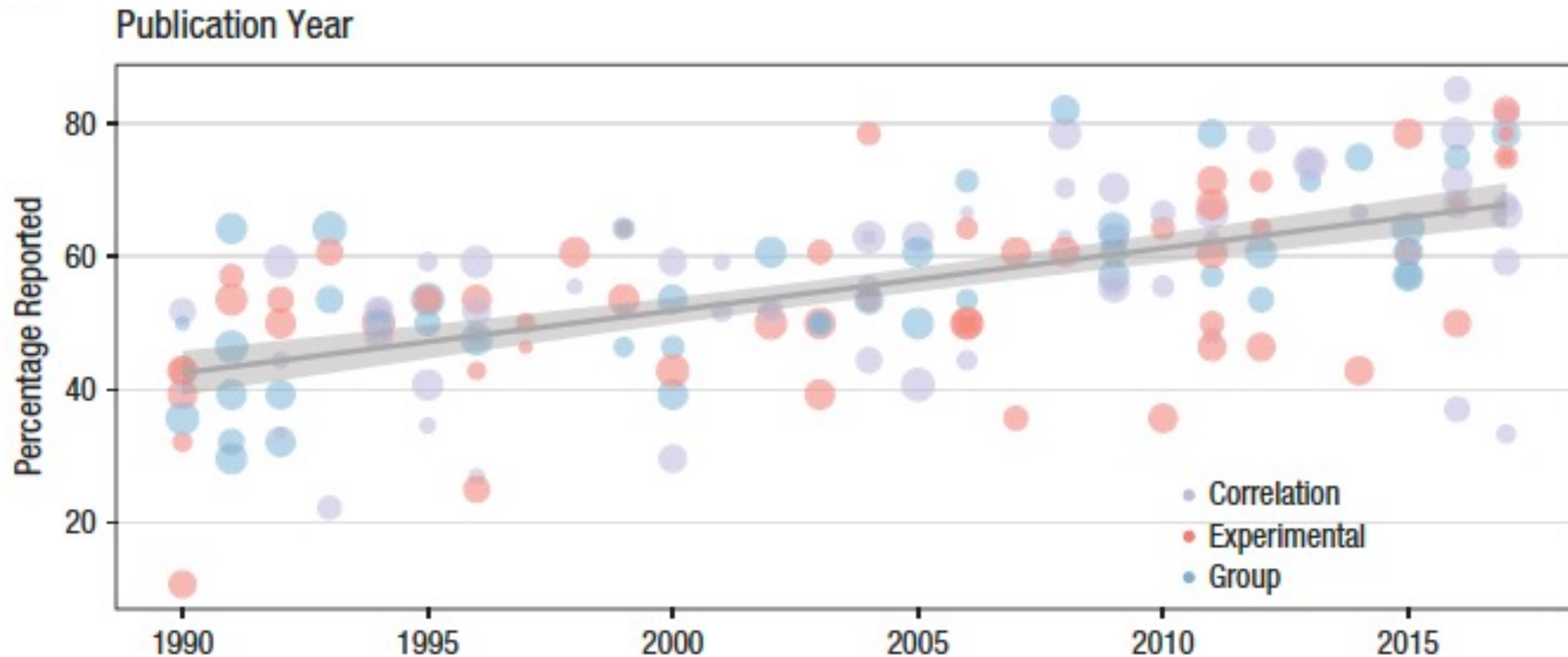
Challenges and Future Directions

Challenges and Future Directions



Polanin, J. R., Hennessy, E. A., & Tsuji, S. (2020). Transparency and Reproducibility of Meta-Analyses in Psychology: A Meta-Review. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 15(4), 1026–1041. <https://doi.org/10.1177/1745691620906416>

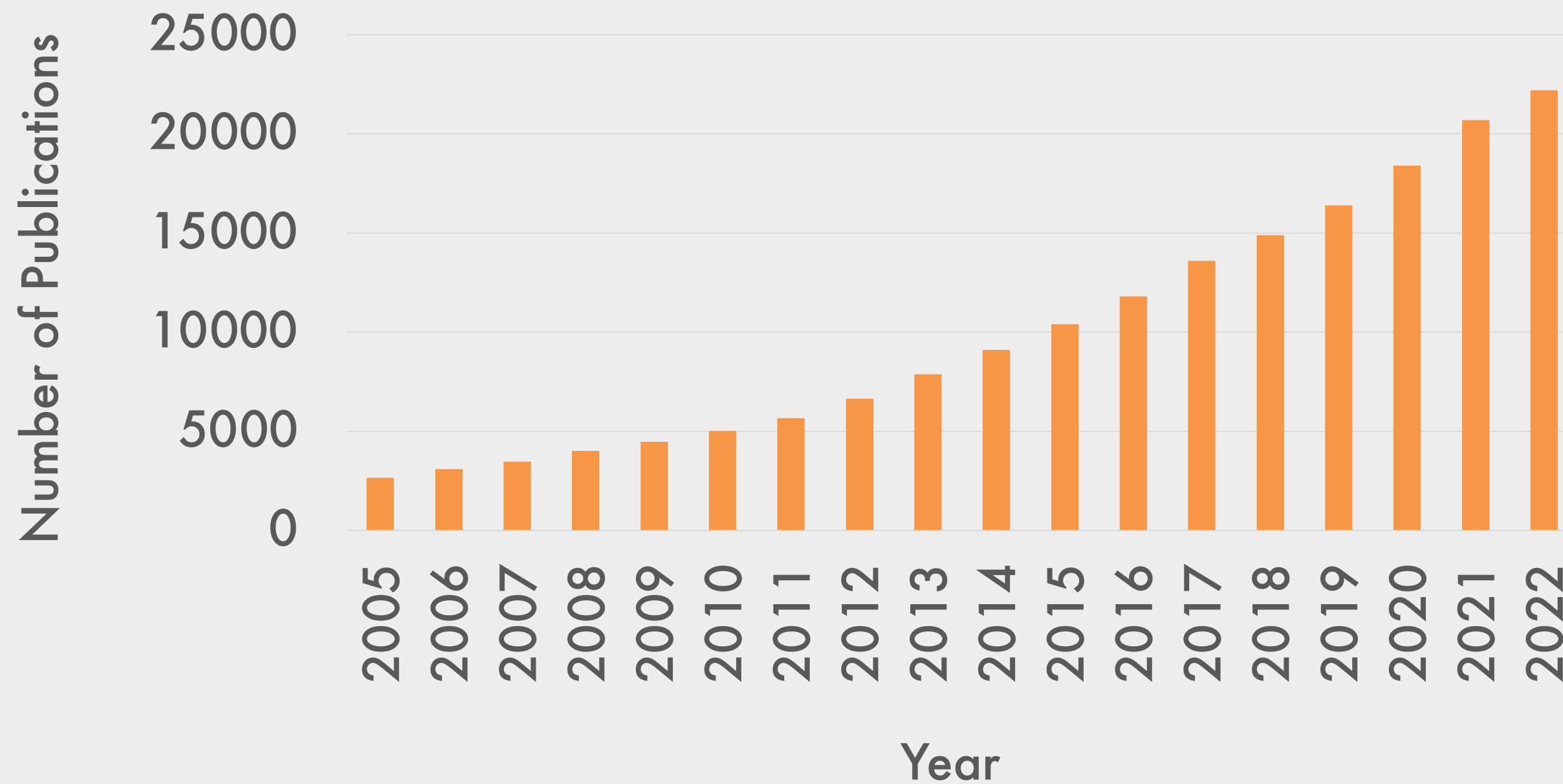
Challenges and Future Directions



Polanin, J. R., Hennessy, E. A., & Tsuji, S. (2020). Transparency and Reproducibility of Meta-Analyses in Psychology: A Meta-Review. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 15(4), 1026–1041. <https://doi.org/10.1177/1745691620906416>

Challenge #1

Growing amount of literature



Challenge #2

Individual Participant Data Meta-Analysis

- Researcher degrees of freedom
- Selection or availability bias
- Reporting deficiencies



Challenge #3

Synthesis of Complex Effect Sizes

- Mediation effects
- Reliability coefficients
- Factor loadings

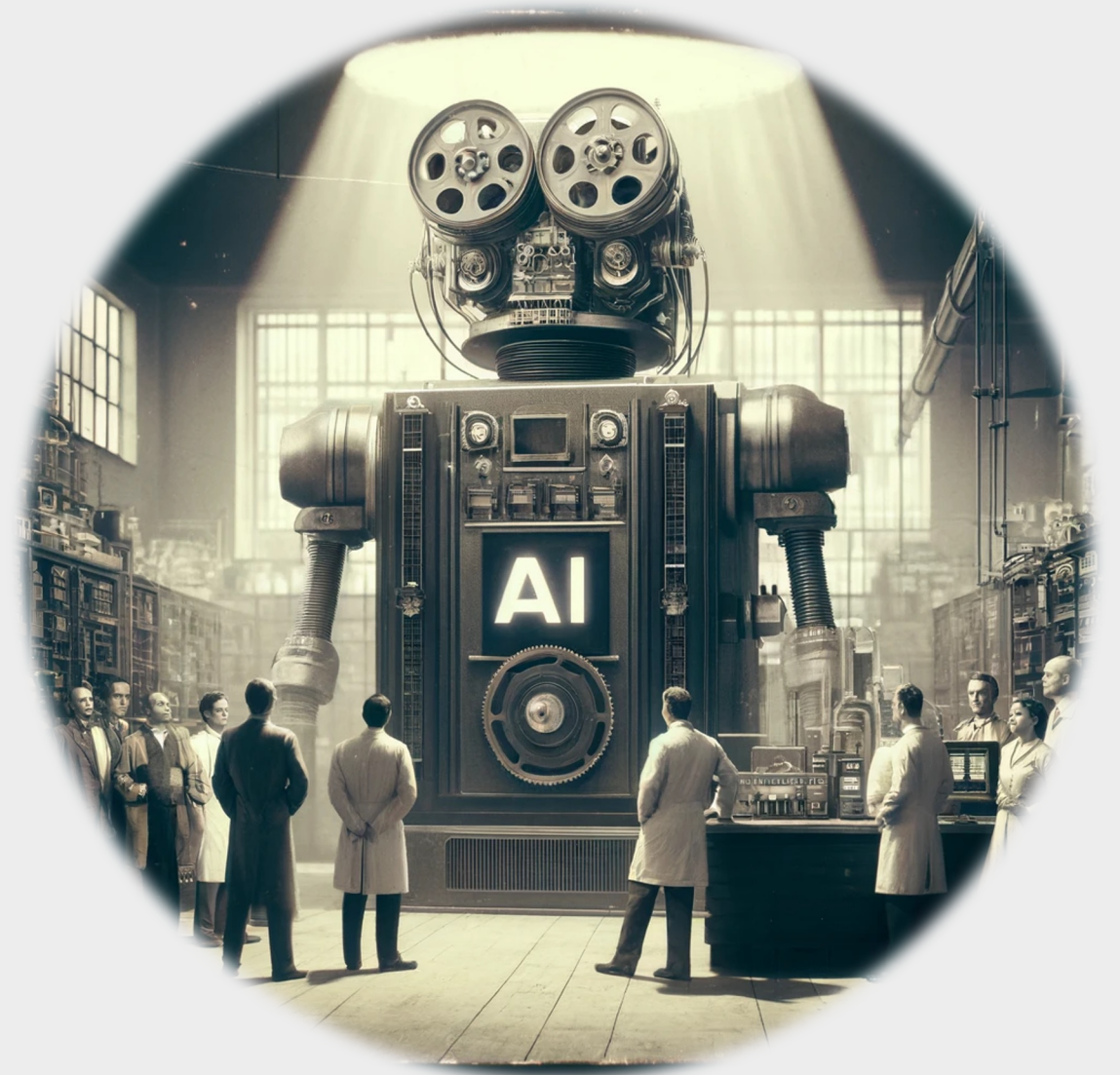


Campos, D. G., Cheung, M. W., & Scherer, R. (2023). A primer on synthesizing individual participant data obtained from complex sampling surveys: A two-stage IPD meta-analysis approach. *Psychological methods*, 10.1037/met0000539. Advance online publication. <https://doi.org/10.1037/met0000539>

Challenge #4

AI Systematic Reviews

- Black box algorithms
- Data biases
- Reproducibility of AI workflows



An underwater photograph of a sea turtle swimming over a coral reef. The turtle is positioned in the center-left of the frame, moving towards the right. Its head is above the water surface, and its front flippers are visible. The coral reef below is covered in green algae and small coral structures. The water is clear and blue. A semi-transparent white rectangular box is overlaid on the center of the image, containing the word "Conclusions" in a bold, yellow, sans-serif font.

Conclusions



A Brief Recap

- Systematic reviews and meta-analyses are central to informing practice, policy, and research
 - During the reviewing process, researchers make several decisions that can result in conflicting outcomes
 - Adopting reproducibility and open science practices is central to ensuring the credibility of systematic reviews and meta-analyses
 - Emerging methodological approaches and technological advancements are posing a challenge to the replicability of systematic reviews and meta-analyses
- 



Practical Recommendations



- **Adhere to reporting standards:** Follow PRISMA or MARS standards
- **Pre-register:** Pre-register the research synthesis protocol
- **Facilitate quality control:** Specify analyses to estimate effect sizes and who extracted and coded the data
- **Facilitate reproducibility:** Provide search syntax, coding schemes, reproducible scripts, and meta-analytic data

"It is arguable that systematic review and meta-analysis exist as research techniques because of the transparency of primary research.



Without the transparency of research design, sample and measurement characteristics, and summary statistics, the meta-analyst would have little information to synthesize."

(Polanin, et al., 2020, p.1027)

References

- Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H. R., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. *Psychological bulletin*, 136(2), 151–173. <https://doi.org/10.1037/a0018251>
- Bushman, B. J., & Huesmann, L. R. (2014). Twenty-five years of research on violence in digital games and aggression revisited: A reply to Elson and Ferguson (2013). *European Psychologist*, 19(1), 47–55. <https://doi.org/10.1027/1016-9040/a000164>
- Campos, D. G., Cheung, M. W., & Scherer, R. (2023). A primer on synthesizing individual participant data obtained from complex sampling surveys: A two-stage IPD meta-analysis approach. *Psychological methods*, 10.1037/met0000539. Advance online publication. <https://doi.org/10.1037/met0000539>
- Elson, M., & Ferguson, C. J. (2014). Twenty-five years of research on violence in digital games and aggression: Empirical evidence, perspectives, and a debate gone astray. *European Psychologist*, 19(1), 33–46. <https://doi.org/10.1027/1016-9040/a000147>
- Ferguson, C. J. (2014). Comment: Why Meta-Analyses Rarely Resolve Ideological Debates. *Emotion Review*, 6(3), 251–252. <https://doi.org/10.1177/1754073914523046>
- Gøtzsche, P. C., Hróbjartsson, A., Maric, K., & Tendal, B. (2007). Data extraction errors in meta-analyses that use standardized mean differences. *JAMA*, 298(4), 430–437. <https://doi.org/10.1001/jama.298.4.430>
- Lakens, D., Hilgard, J. & Staaks, J. (2016). On the reproducibility of meta-analyses: six practical recommendations. *BMC Psychol* 4, 24. <https://doi.org/10.1186/s40359-016-0126-3>
- López-Nicolás, R., López-López, J. A., Rubio-Aparicio, M., & Sánchez-Meca, J. (2022). A meta-review of transparency and reproducibility-related reporting practices in published meta-analyses on clinical psychological interventions (2000–2020). *Behavior research methods*, 54(1), 334–349. <https://doi.org/10.3758/s13428-021-01644-z>
- Markowetz, F. (2015) Five selfish reasons to work reproducibly. *Genome Biol* 16, 274.. <https://doi.org/10.1186/s13059-015-0850-7>
- Moreau, D., & Gamble, B. (2022). Conducting a meta-analysis in the age of open science: Tools, tips, and practical recommendations. *Psychological Methods*, 27(3), 426–432. <https://doi.org/10.1037/met0000351>
- Munafò, M., Nosek, B., Bishop, D., Button, K., Chambers C. (2017). A manifesto for reproducible science. *Nat Hum Behav* 1, 0021 <https://doi.org/10.1038/s41562-016-0021>
- Nosek, B. A., & Lindsay, D. S. (2018). Preregistration becoming the norm in psychological science. *APS Observer*, 31.
- Tendal, B., Nüesch, E., Higgins, J. P., Jüni, P., & Gøtzsche, P. C. (2011). Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. *BMJ (Clinical research ed.)*, 343, d4829. <https://doi.org/10.1136/bmj.d4829>



Thank You



Diego G. Campos

@diegc_

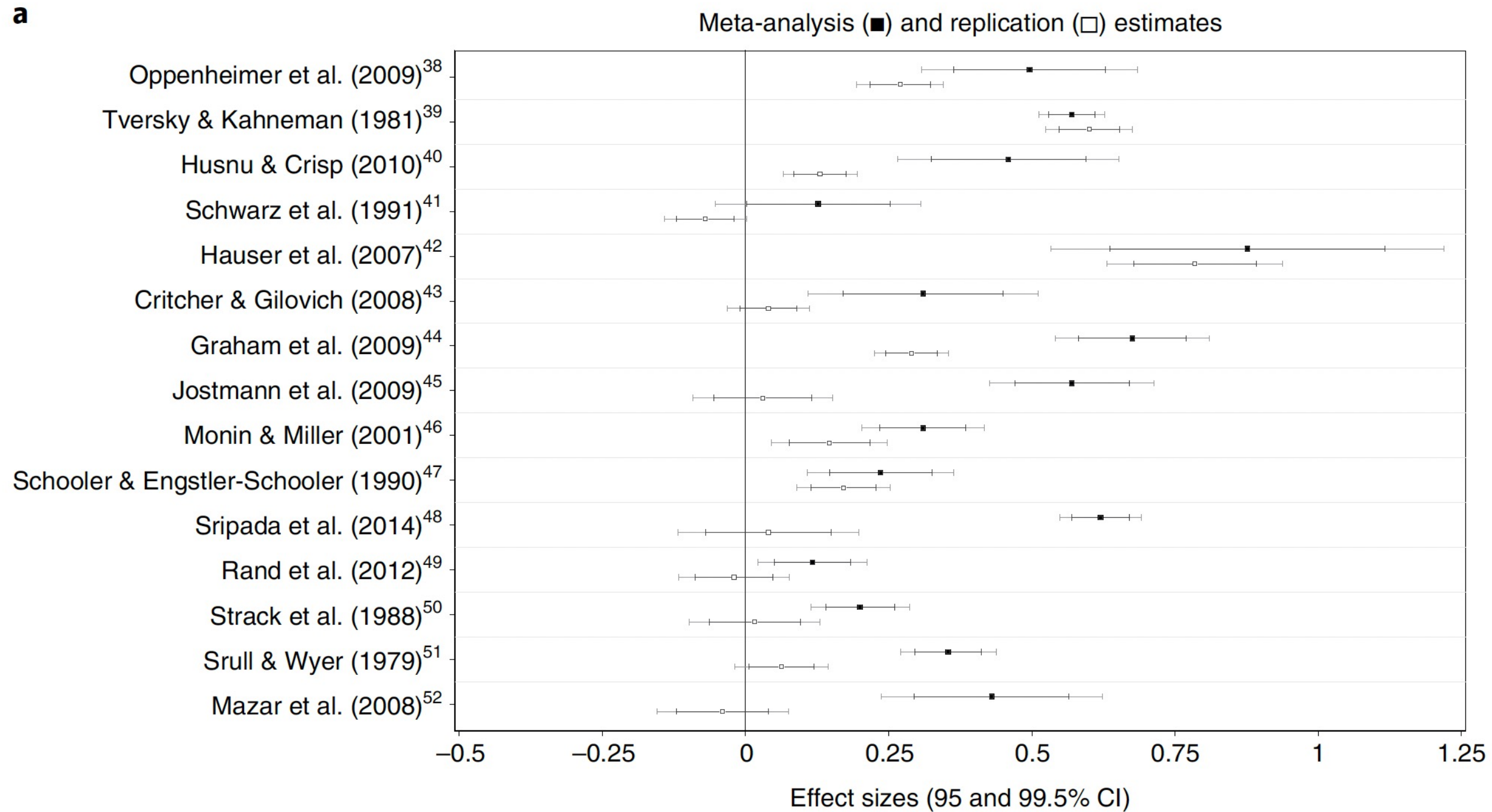
diegogc@uio.no

www.diegocampos.co



Meta-Analysis Failure?

a



Five selfish reasons to work reproducible

1. Making your **data open and transparent** helps you and your co-authors to spot errors that might undermine your results.
2. The new **automated methods** make the paper easier to write.
3. If anyone can **see how you analyzed your data**, convincing reviewers that you did the right thing is easier.
4. Openly **documenting every step** of the analysis helps you continue your work a few months later.
5. Making **everything available** shows the scientific community that you have nothing to hide and have done everything in good faith.