



Building and Synthesizing Evidence: Replication

Dr. Sean Grant

Assistant Professor in Social & Behavioral Sciences

Richard M. Fairbanks School of Public Health

Indiana University, Indianapolis (IUPUI)

Acknowledgements

- **TRUST Collaborators**
 - Evan Mayo-Wilson, Indiana University School of Public Health—Bloomington
 - Lauren Supplee, Child Trends
- **TRUST Funding**
 - Arnold Ventures



This presentation addresses the following questions:

- How do open science practices help us to build a credible social policy evidence base?
 - What role do replication studies play?
 - How can researchers conduct replicable studies?
 - How can we shift the evidence ecosystem toward transparent, rigorous replications?



Distinguishing “Reproducibility” from “Replicability”

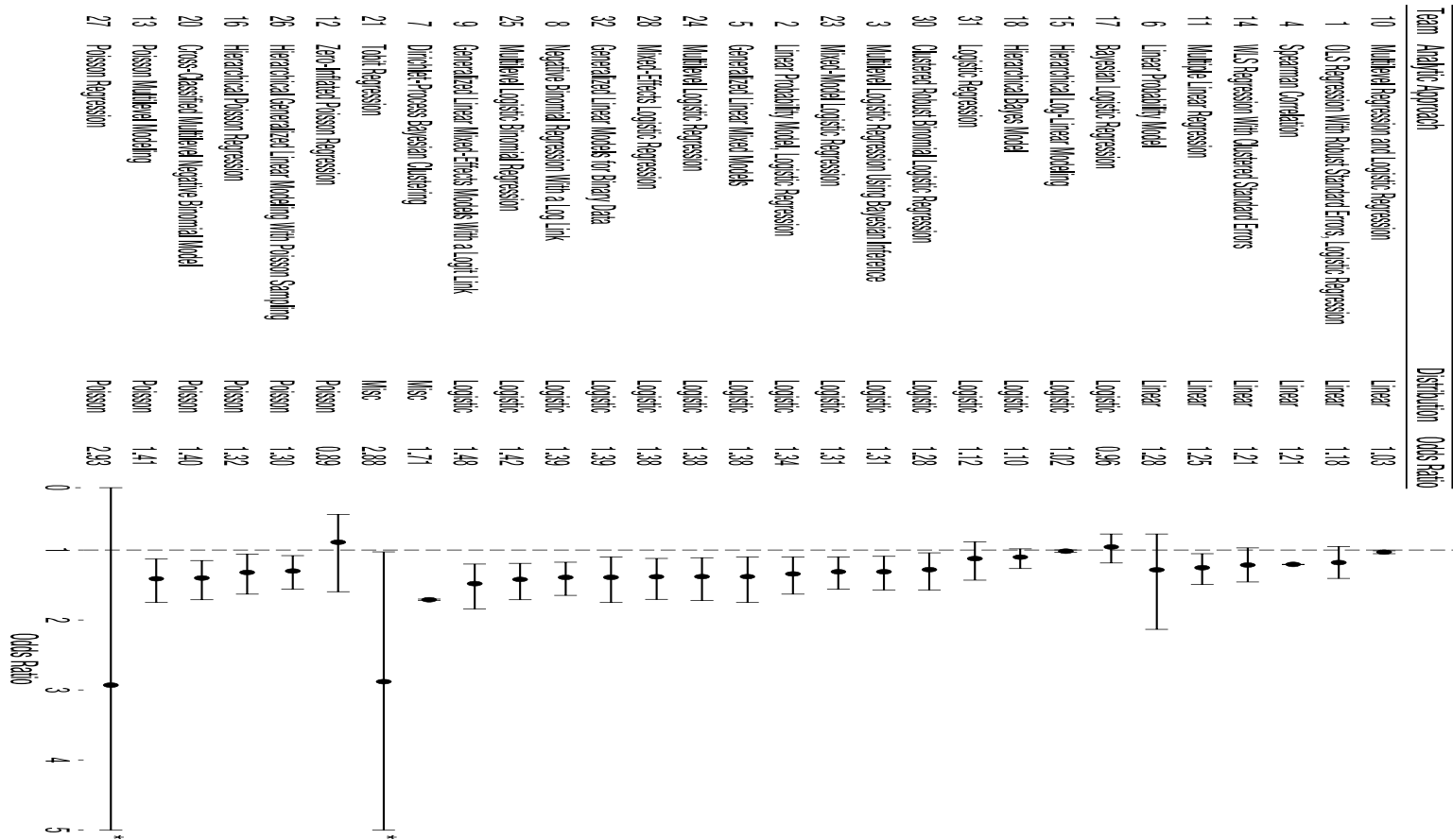
Reproducibility

Same data,
Same procedures,
Same results

Replicability

New data,
Same procedures,
Same results

The Social and Behavioral Sciences Have a Reproducibility Problem



The Social and Behavioral Sciences Have a Replication Problem

Study	Discipline	Objective	Replication Findings
Camerer et al. (2016)	Economics	Attempt to replicate 18 studies from AER and QJE in 2011-2014	<ul style="list-style-type: none"> - 61% significant effect, same direction as original - Replicated effect size 66% of original on average
Camerer et al. (2018)	Social Sciences	Attempt to replicate 21 experimental studies in Nature and Science	<ul style="list-style-type: none"> - 62% significant effect, same direction as original - Replicated effect size 50% of original on average
Chang and Li (2015)	Economics	Attempt to reproduce findings from 67 papers using original data and code	<ul style="list-style-type: none"> - 33% replication of key qualitative result - 49% replication with original author assistance
Klein et al. (2014)	Psychology	Attempt to replicate 13 psychological effects using 36 independent samples	<ul style="list-style-type: none"> - 10 effects replicated consistently - Effects did not differ by setting or country
Open Science Collaboration (2015)	Psychology	Attempt to replicate 100 studies from three high-ranking journals	<ul style="list-style-type: none"> - 36% were statistically significant - 47% had 95% CI containing original effect size



Social Determinants of These Problems: The Scientific Ecosystem




Social Determinants of These Problems: The Scientific Ecosystem



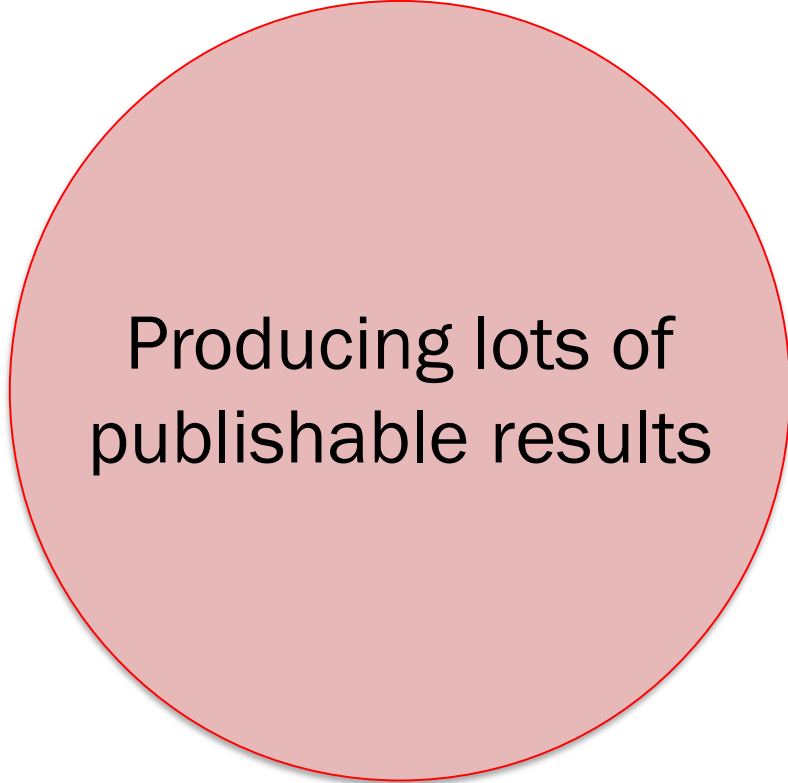
Scientists (People) Respond to Their Environment

What's best for *research*



High quality
research,
regardless of
outcome

What's best for *researchers*



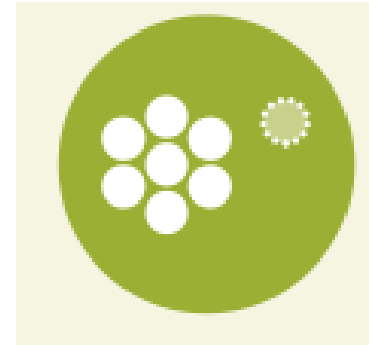
Producing lots of
publishable results



Behavioral Determinants: Questionable Research Practices

Underspecified Protocols

Methods and analytic plan are not shared with other scientists in sufficient detail.

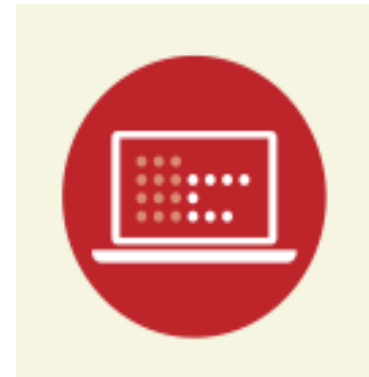


Reporting Bias

When scientists or journals decide not to publish analyses, outcomes, or entire studies (e.g., results are not statistically significant).

Human Error

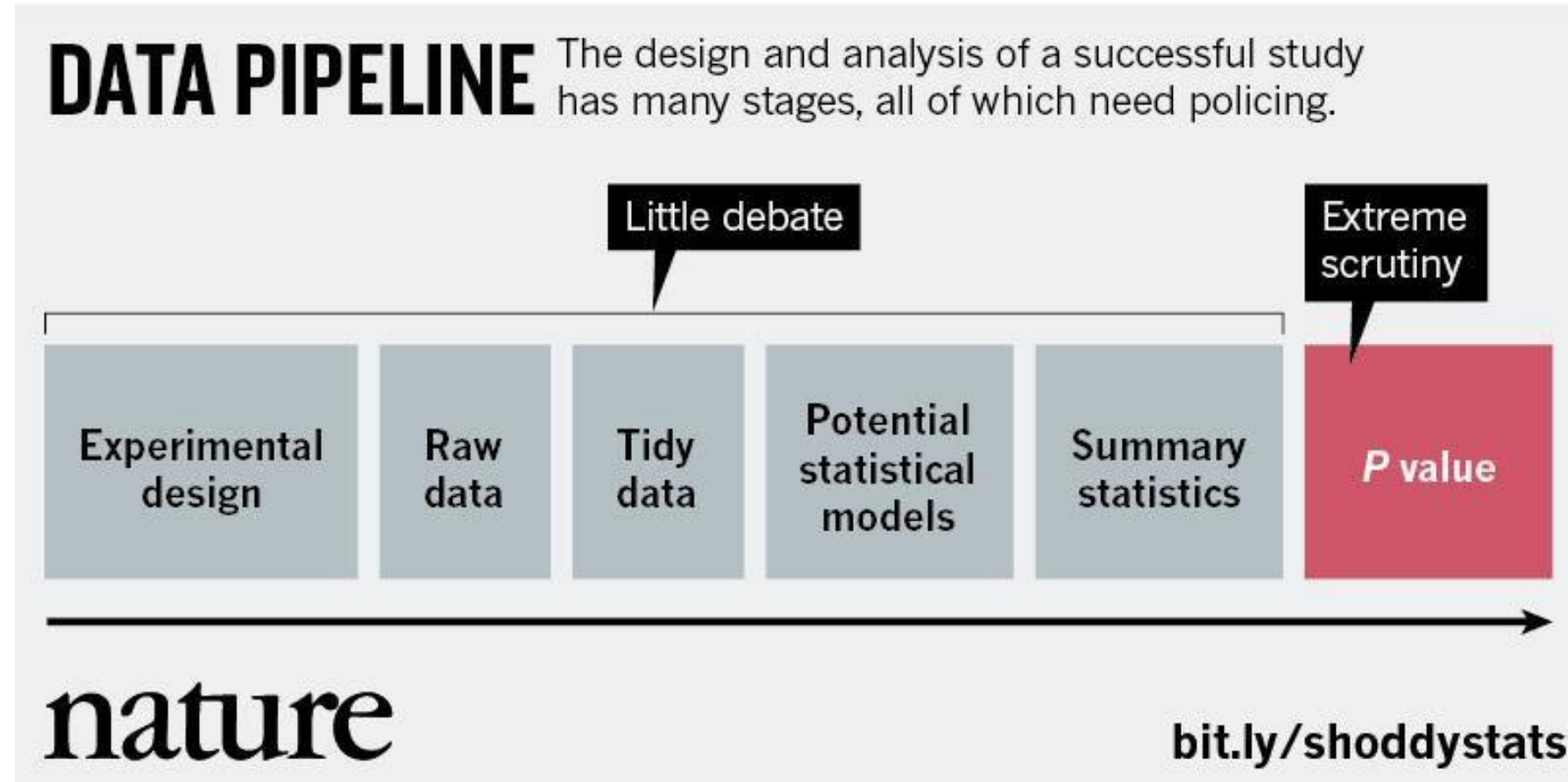
Technical errors may exist within a study, (e.g., computational errors, copy/paste mistakes).



Data dredging (p-hacking)

Repeatedly searching a dataset or trying alternative analyses until a (significant) result is found.

What Enables the QRPs Causing Our Problems? Closed Research Workflows



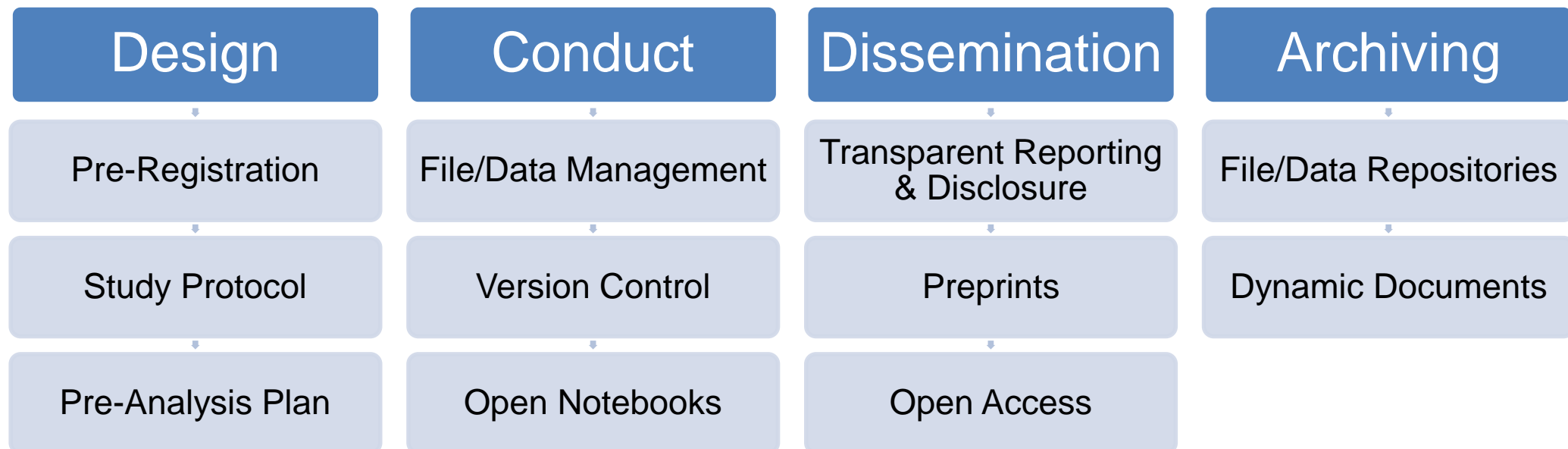
Open Science as a Solution

- Transparent workflows with replication documentation
- Conduct “robustness check” prior to replication
- Adopt “Registered (Replication) Reports” model for funding and publishing
- Include open science practices in standards of evidence on “What Works” (i.e., designating “evidence-based” practices)

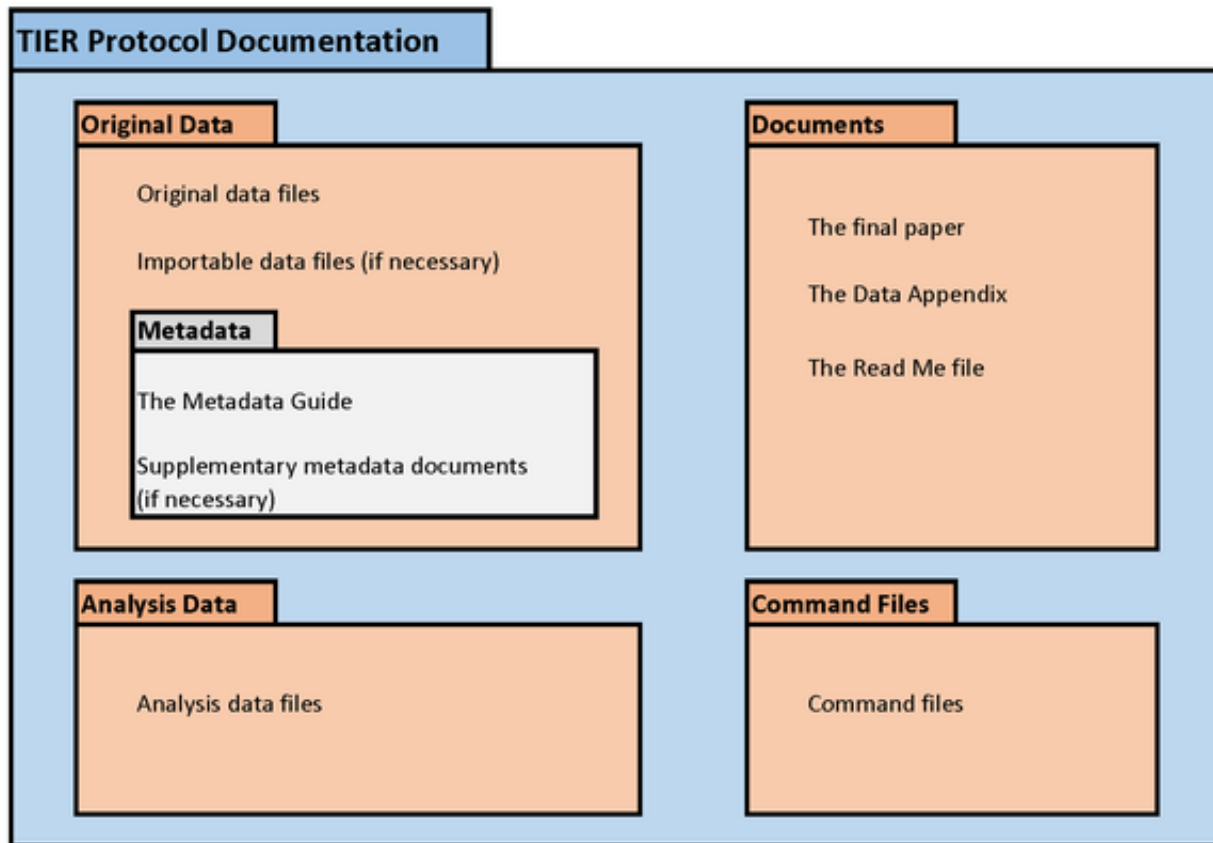


Open Science as a Solution: Transparent Workflows

Organized Workflow and File Management (Open Science Framework)



Open Science as a Solution: Replication Documentation (TIER/DRESS Protocols)



- List of results documentation should reproduce
- Information on software and original data
- Processed data, command, and README files
- Instructions for replicating

Open Science as a Solution: “Robustness Check” Prior to Replication

- Check the internal consistency of the statistical results (Statcheck)
- Reanalyze the data using the original analytical strategy (Computational Reproducibility)
- Check if the result is robust to alternative analytical choices (sensitivity, multiverse analyses)
- *Then consider a replication study in a new sample



Open Science as a Replicability Solution: Registered (Replication) Reports



Authors and Reviewers Systematically Compare Replication Study to the Original Study

Replication Continuum

Design Facet	Direct Replication			Conceptual Replication	
	Exact Replication (All facets under researcher control are the same)	Very Close Replication (Procedure or physical setting is different)	Close Replication (IV or DV stimuli are different)	Far Replication (IV or DV operationalization or population is different)	Very Far Replication (IV or DV constructs are different)
Effect, Hypothesis	Same	Same	Same	Same	Same
IV Construct	Same	Same	Same	Same	Different
DV Construct	Same	Same	Same	Same	Different
IV Operationalization	Same	Same	Same	Different	
DV Operationalization	Same	Same	Same	Different	
Population (e.g., age)	Same	Same	Same	Different	
IV Stimuli	Same	Same	Different		
DV Stimuli	Same	Same	Different		
Procedural Details	Same	Different			
Physical Setting	Same	Different			
Contextual Variables	Different				
⋮	⋮				



Funders, Journals, and Clearinghouses Coordinate on a Registered Report “What Works” Scheme

- **Funders** sponsor work that will be published, transparent, and reproducible
- **Journals** publish high-quality, impactful work
- **Institutions** coordinate research compliance (e.g., IRBs) with funders and journals
- **Clearinghouses** obtain evidence that they can trust












Clearinghouses Describe the Transparency of Research Studies Evaluations Using Badges


Infidelity Distress Effect

Original instantiation of effect

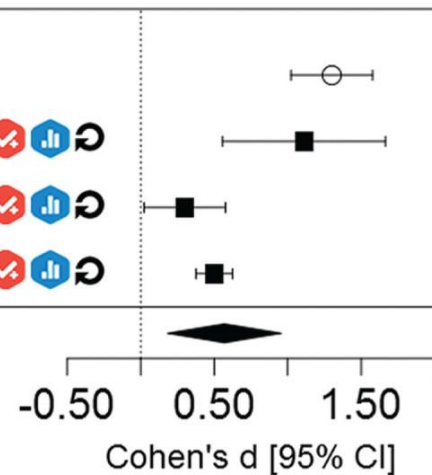
Buss et al. (1999) Study 2

IJzerman et al. (2014) Study 1     

IJzerman et al. (2014) Study 2     






IJzerman et al. (2014) Study 4     






Meta-analytic estimate of replications



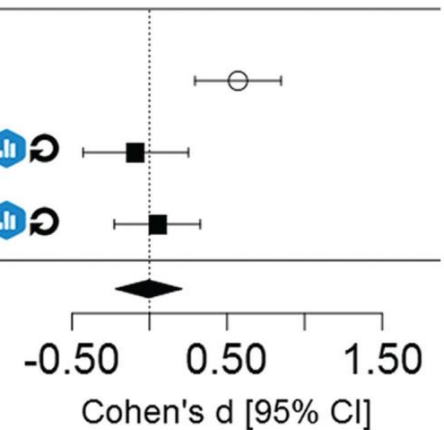
Population generalization of effect

Shackelford et al. (2004)

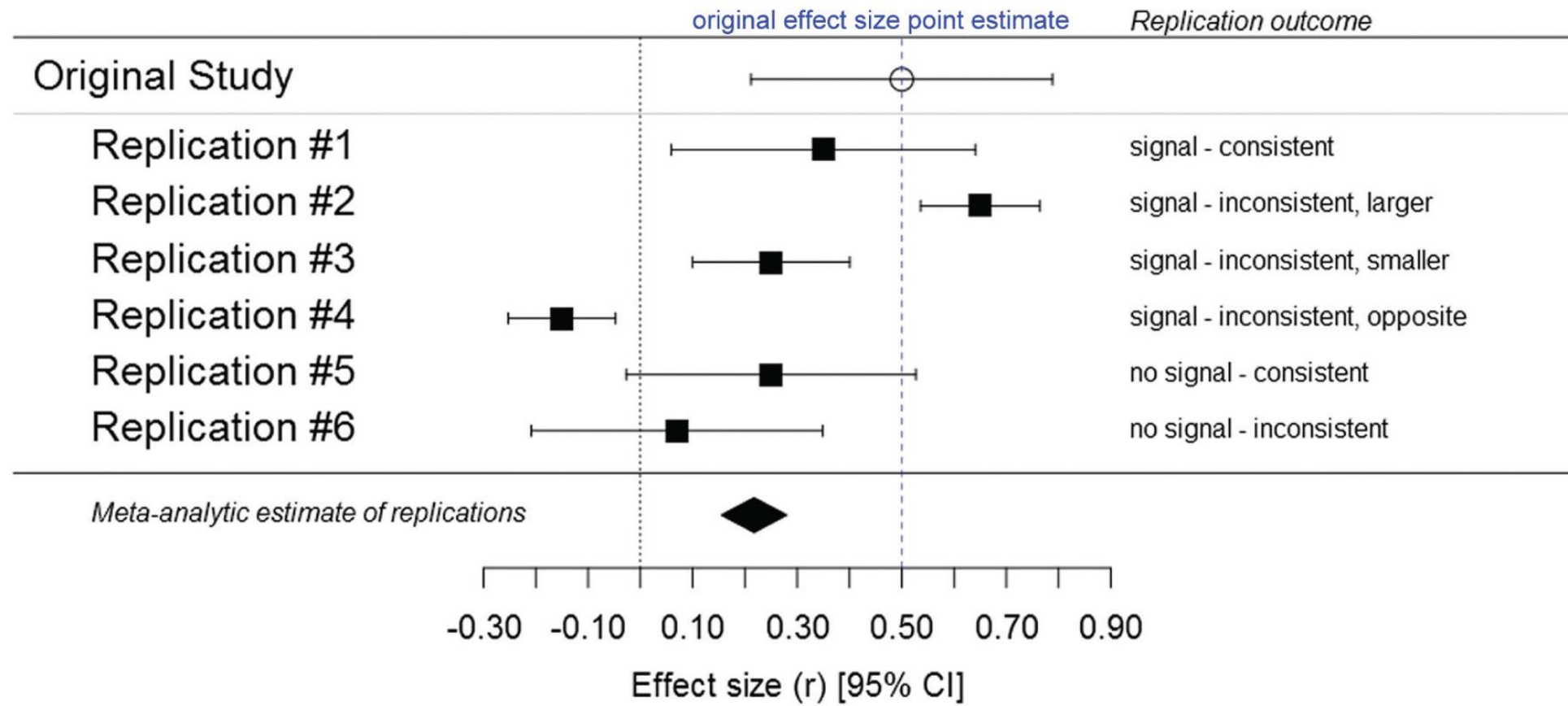
IJzerman et al. (2014) Study 3     

IJzerman et al. (2014) Study 4     

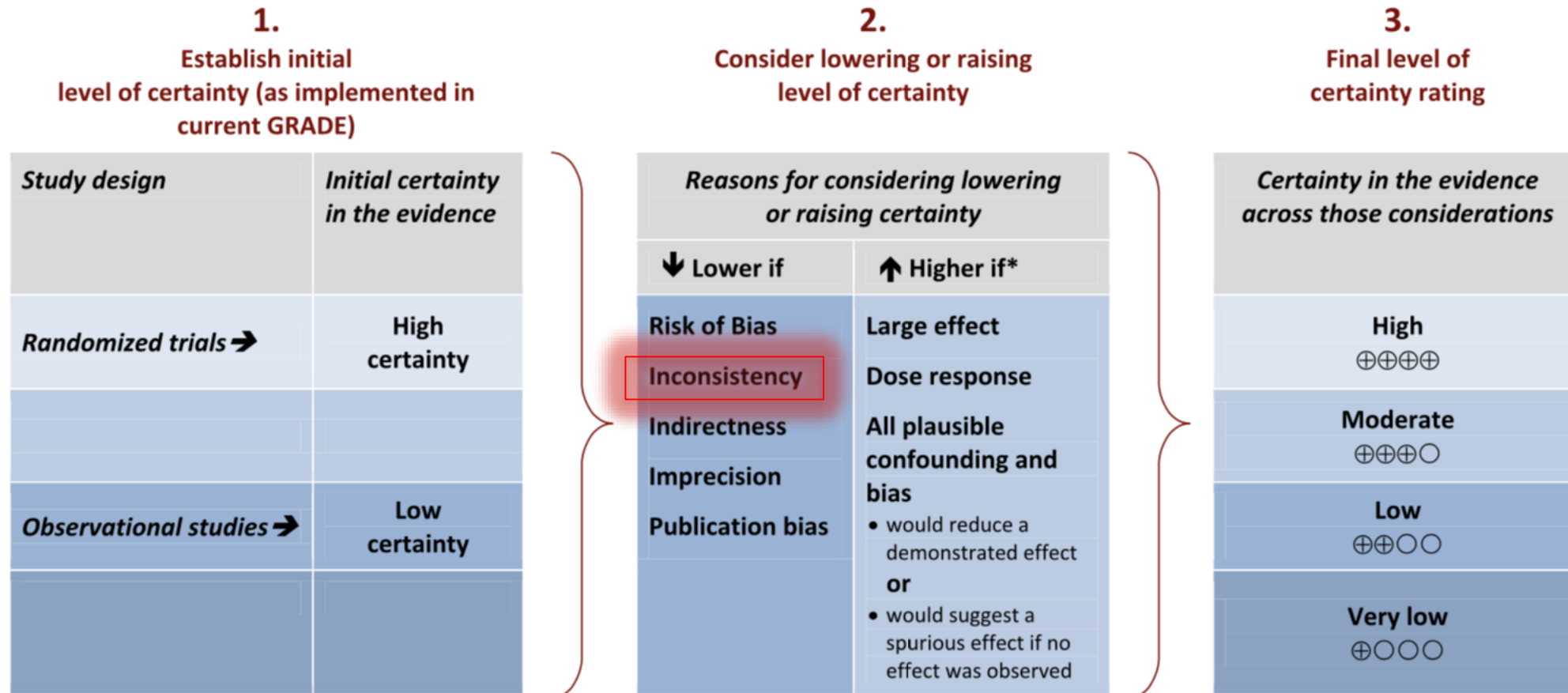
Meta-analytic estimate of replications



Clearinghouses Assess the Signal, Consistency, and Size of Effects across Replications



Replication Studies Build Our Confidence in the Cumulative Evidence Base on “What Works”



Why This Room Should Care: Reproducibility and Replicability are Essential to Social Policy Research and Evaluation

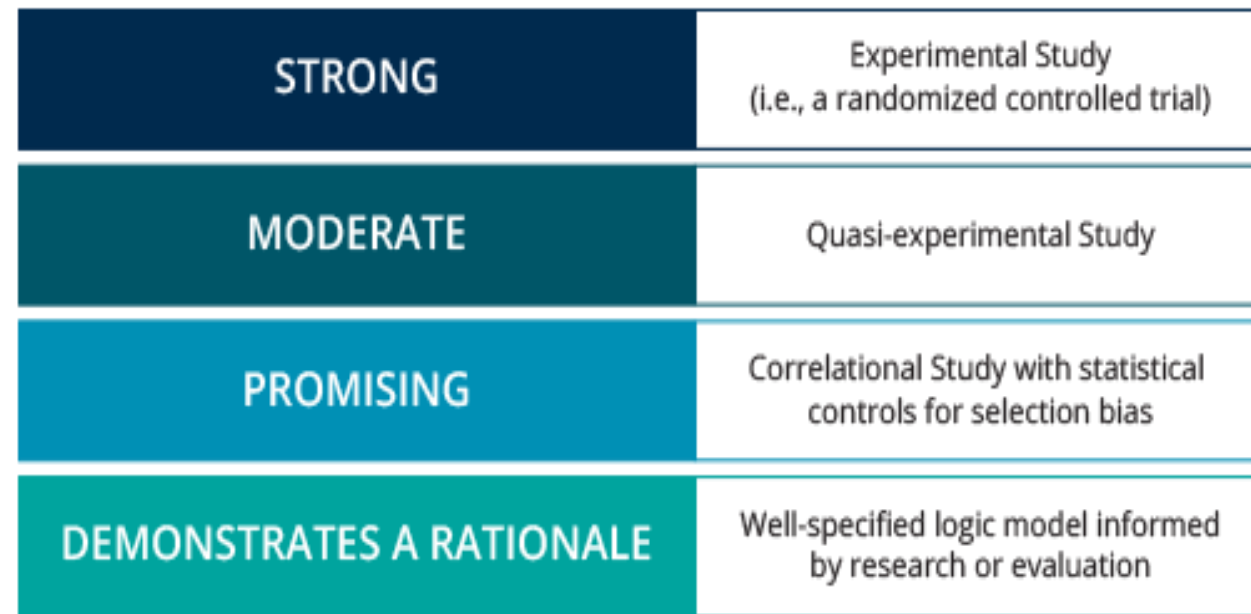
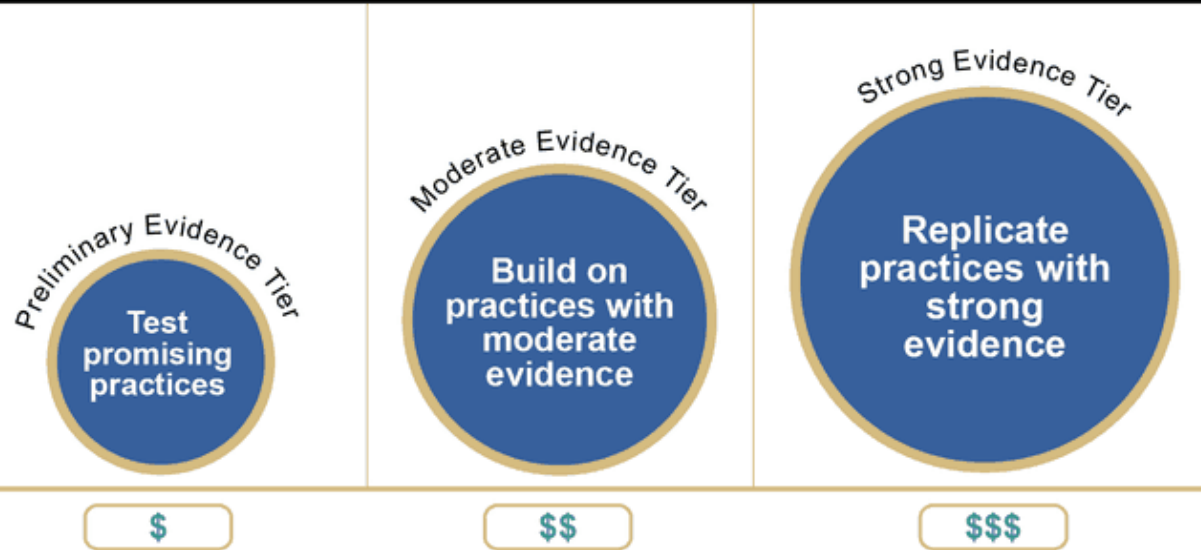
- Evidence-based policy and practice: research and analysis inform policy and practice decision-making
 - Intervention studies with reproducible finding and replicable effects
- Large focus on rigorous causal inference
 - High-quality policy analyses, research syntheses, randomized trials, and quasi-experiments



They Directly Influence Our Labels of “Evidence-Based” and “What Works”

Tiered Evidence Grant Model

(The size of grant awards is linked to the strength of the evidence provided in the application.)



Source: GAO analysis of agency documents. | GAO-16-818

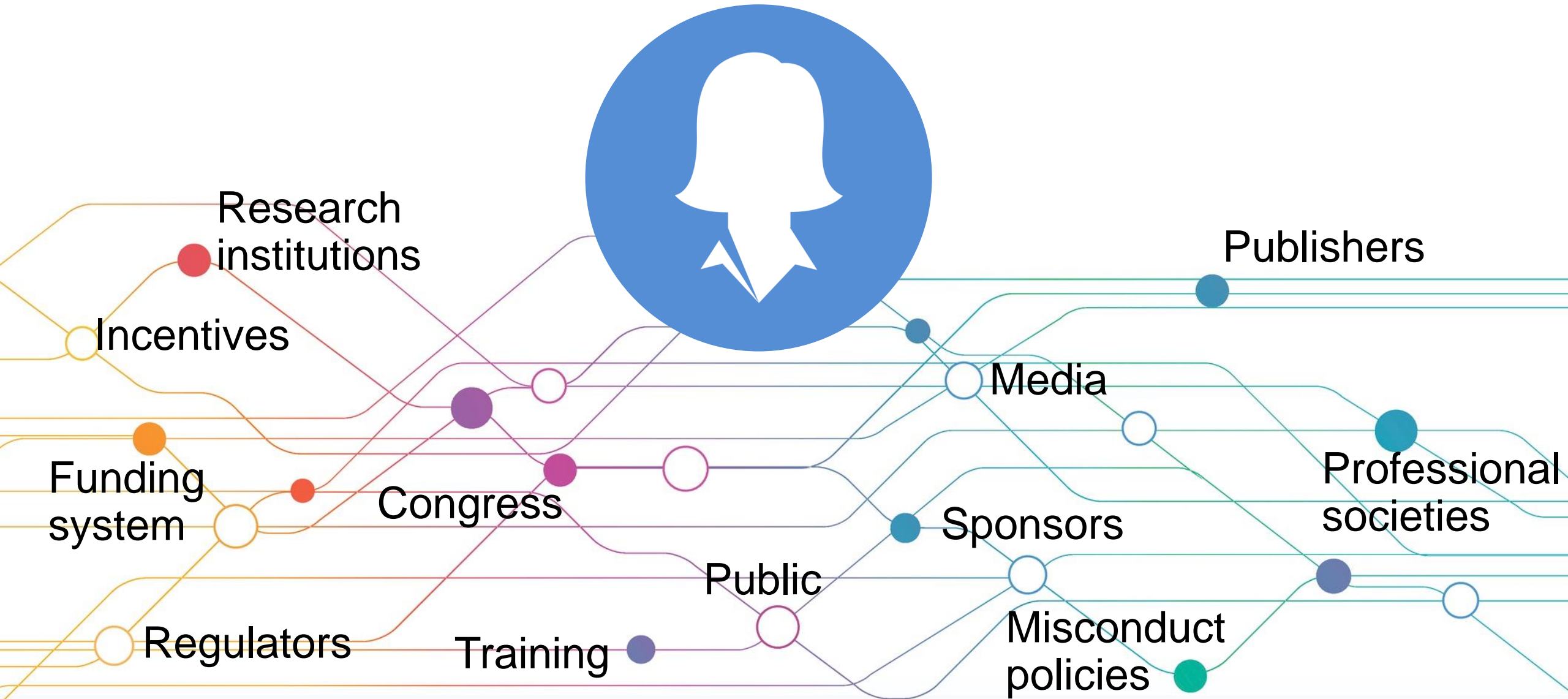


Our Evidence Ecosystem Generally Does Not Require “What Works” to be based on Open Scientific Research

Open Science Practice	Clearinghouses	Examples
Study Registration	HomVEE, WWC	Study registration was assessed and disclosed Created study registry
Protocol and Statistical Analysis Plan		
Data, Code, and Materials Sharing		
Design and Analysis Transparency	HomeVEE, WWC	Author reporting guide
Public Availability of Results	ESER, TPP, WWC	Database of results from studies in higher tiers
Replication	CS, HomVEE, TPP, WWC	Number of studies influences evidence tier Indicate whether a beneficial outcome has replicated
Investigator Conflicts of Interest	CS, HomVEE, SFER, WWC	Declare (lack of) independence of evaluator Independence of evaluator could influence study inclusion/exclusion
Citation Standards		



If ultimate goal is to change the ecosystem to produce reproducible findings and replicable effects...



... everyone has a role!

- **Researchers and analysts:** adopt transparent, reproducible workflows and conduct replications
- **Editors/peer-reviewers:** enforce transparency during article submission and review; prioritize replications
- **Sponsors/funders:** require transparent, reproducible workflows; coordinate registered (replication) report schemes
- **Policy-makers:** integrate transparency, reproducibility, and replicability into standards of evidence on “What Works”

Contact Details

Twitter: @GrantSeanP

Email: spgrant@iu.edu

Website: <https://fsph.iupui.edu/about/directory/grant-sean.html>



IUPUI

FULFILLING *the* PROMISE