

Open Science and Data Transparency

Tim Errington
Center for Open Science
<http://cos.io/>



JOHN TEMPLETON FOUNDATION
SUPPORTING SCIENCE ~ INVESTING IN THE BIG QUESTIONS



NIH National Institute of Mental Health NIH National Institute on Aging



THE WILLIAM AND FLORA HEWLETT FOUNDATION



Mission: Improve
openness, integrity, and
reproducibility of
scientific research

Research Culture on Openness

Norms

Communality

Open sharing

Universalism

Evaluate research on own merit

Disinterestedness

Motivated by knowledge and discovery

Organized skepticism

Consider all new evidence, even against one's prior work

Quality

Counternorms

Secrecy

Closed

Particularism

Evaluate research by reputation

Self-interestedness

Treat science as a competition

Organized dogmatism

Invest career promoting one's own theories, findings

Quantity

Merton, 1942

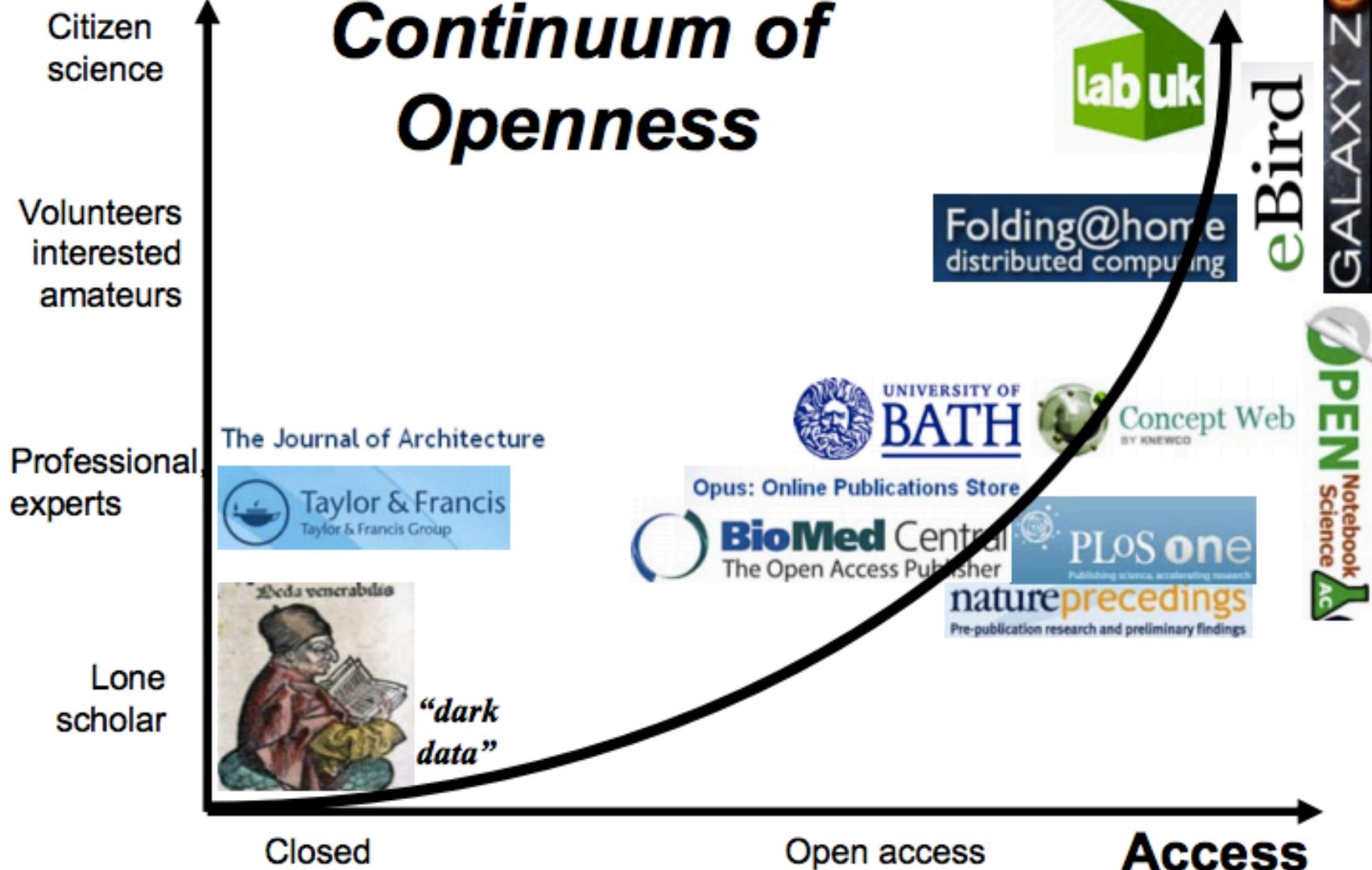


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

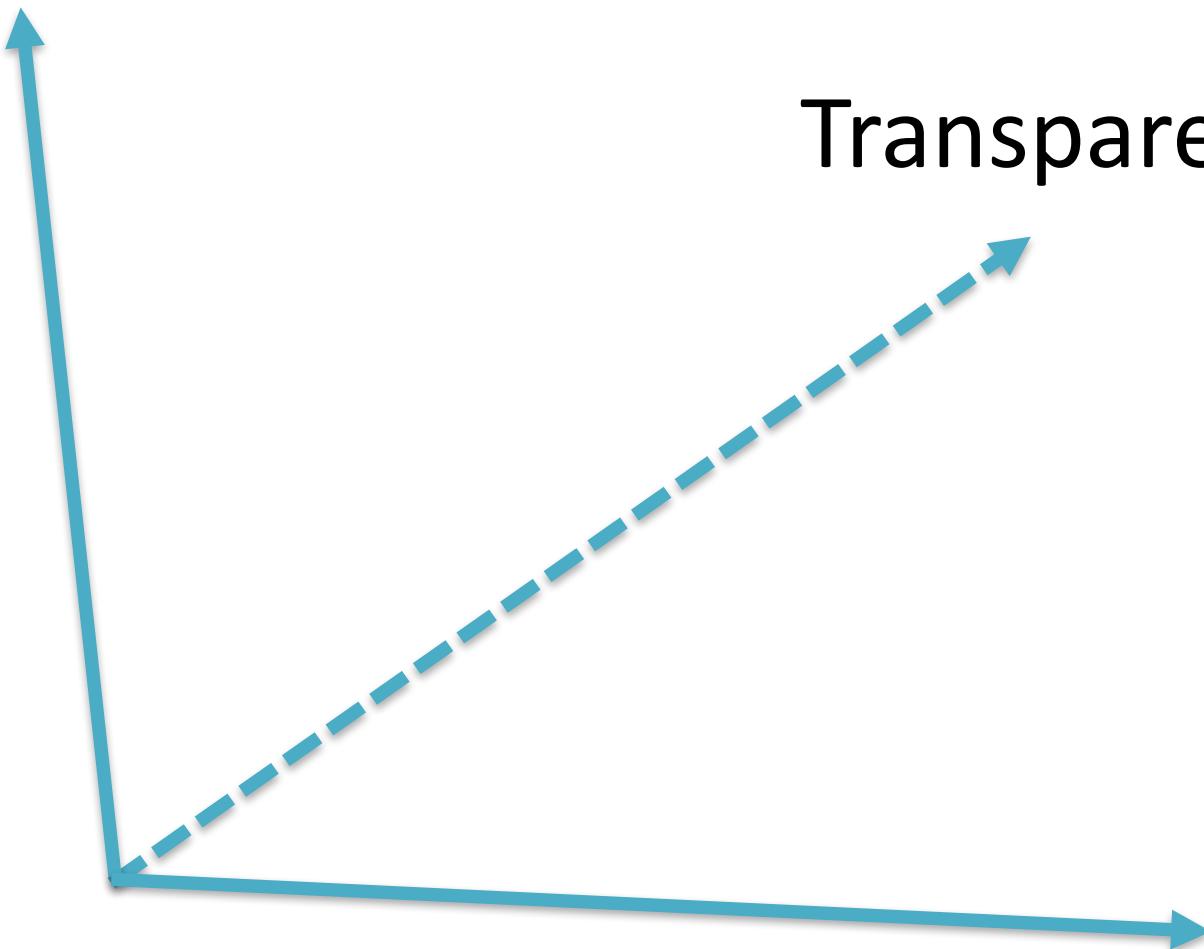
What is Open Science?

Participation

Continuum of Openness



Participation

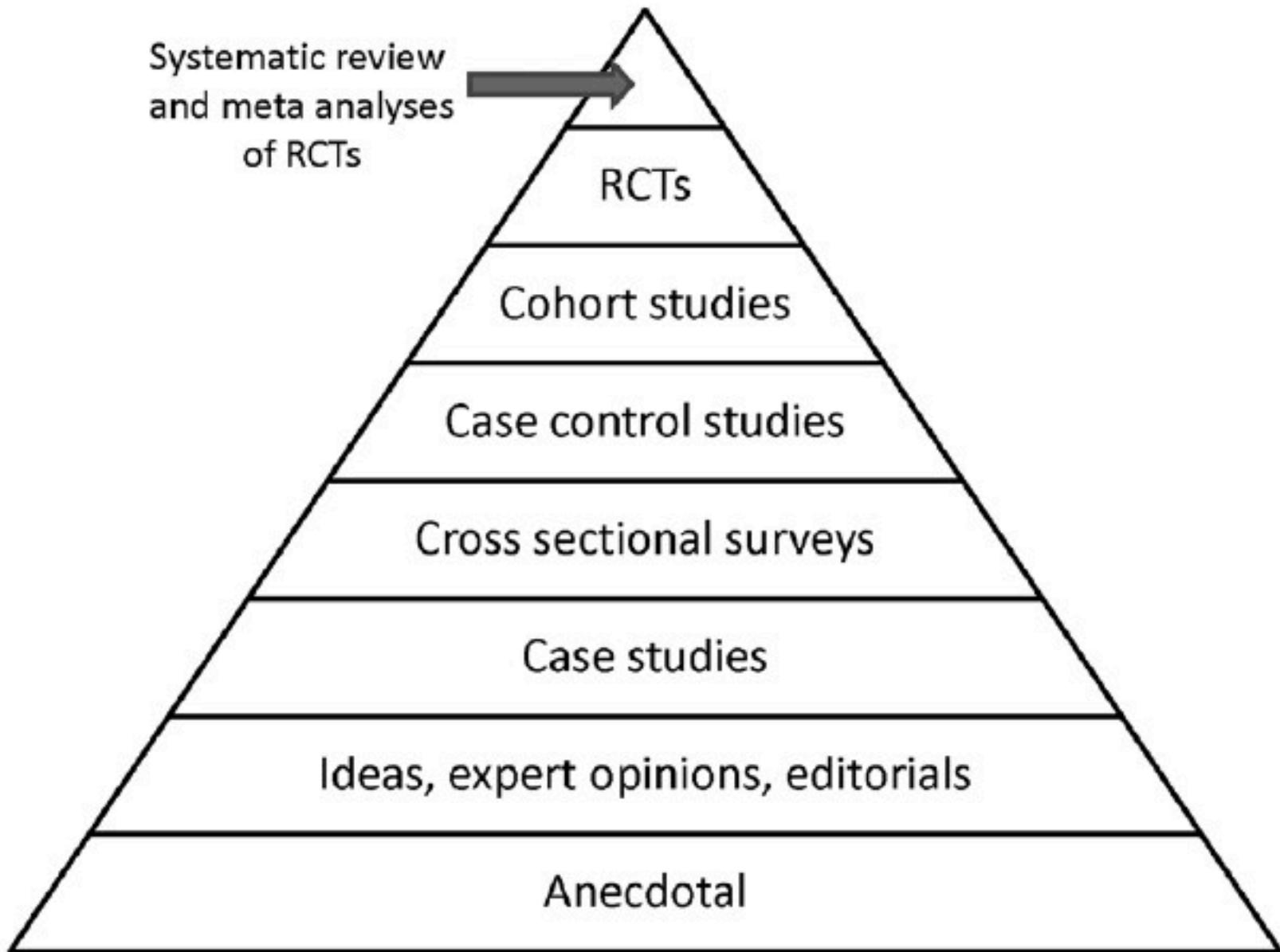


Transparency

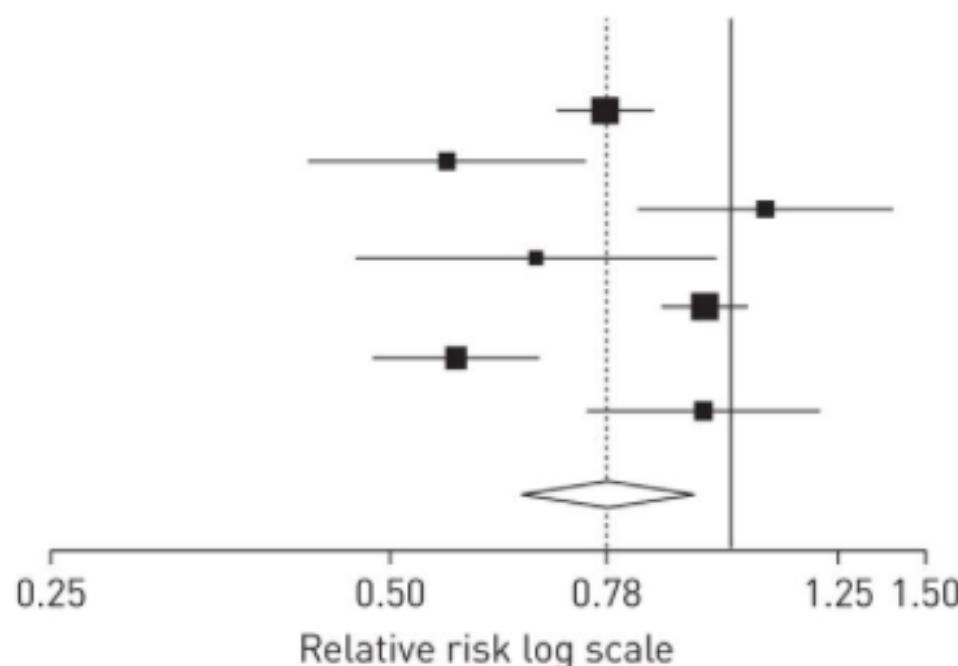
Access

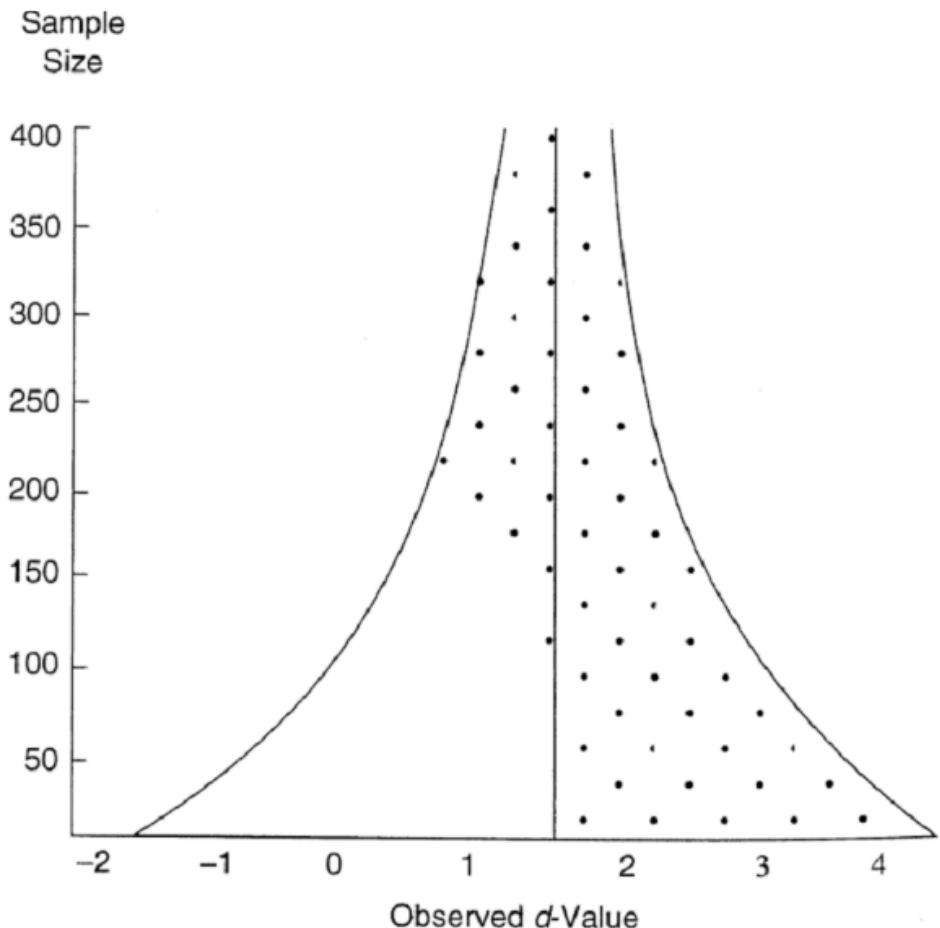
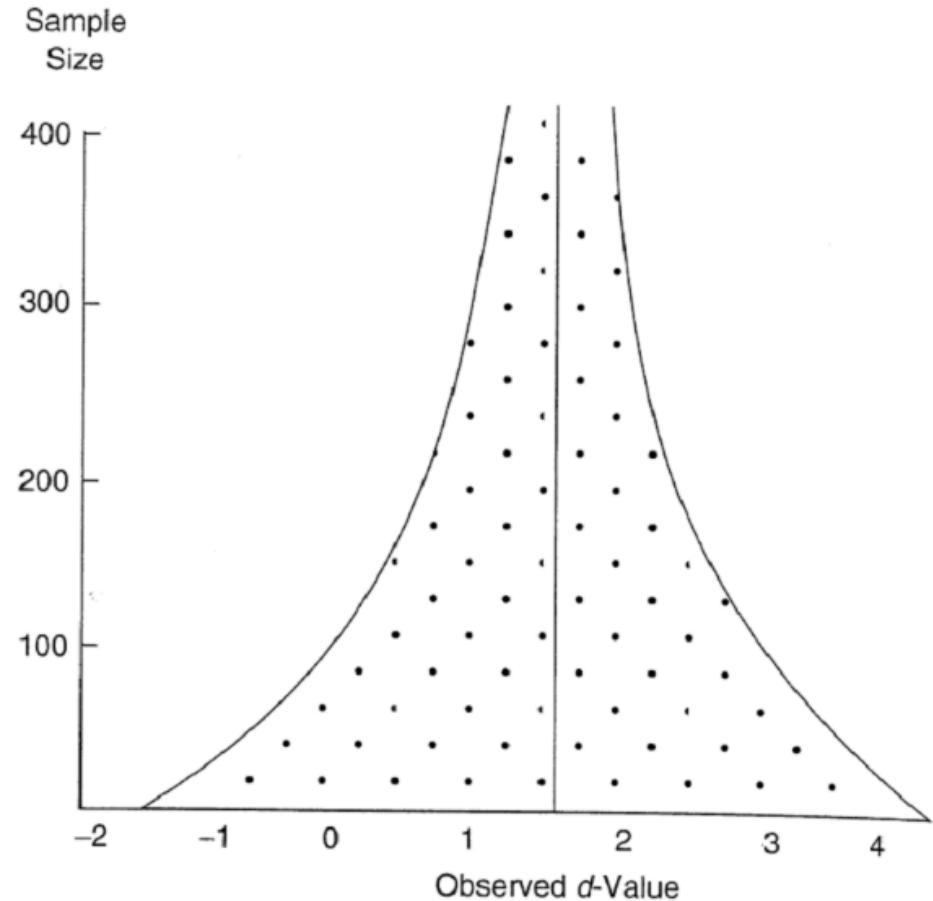
Why should we care?

- To increase the efficiency of research
- Fulfill ethical obligations to participants
- Provide information to potential participants and the research community
- Facilitate systematic reviews and other analyses



Study [ref.], year	Estimate	[95% CI]
ZHENG <i>et al.</i> [1], 2014	0.78	(0.70–0.86)
TSE <i>et al.</i> [20], 2013	0.56	(0.42–0.75)
SCHERMER <i>et al.</i> [26], 2009	1.08	(0.83–1.40)
BACHH <i>et al.</i> [27], 2007	0.67	(0.47–0.97)
DECRAMER <i>et al.</i> [3], 2005	0.95	(0.87–1.04)
PELA <i>et al.</i> [24], 1999	0.57	(0.48–0.68)
MCGAVIN <i>et al.</i> [22], 1985	0.95	(0.75–1.21)
Overall ($I^2=86\%$, $p<0.01$)	0.78	(0.65–0.93)





67TRIALS
CHECKED**9**TRIALS WERE
PERFECT**354**OUTCOMES NOT
REPORTED**357**NEW
OUTCOMES
SILENTLY
ADDED

On average, each trial reported just 58.2% of its specified outcomes. And on average, each trial silently added 5.3 new outcomes.

Home

September

< Previous

Genetics in Medicine

Official Journal of the American College of Medical Genetics and Genomics

[Home](#) | [Current Issue](#) | [Archive](#) | [Podcasts](#) | [For Authors & Referees](#) | [About the journal](#)

Original Inv

Advance Online Publication

Article

Reana

GENETICS IN MEDICINE | ORIGINAL RESEARCH ARTICLE



Shanil Ebral
John P. A. Ioannidis

[+] Author A

JAMA. 2014

Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers

Aaron M. Wenger PhD, Harendra Guturu PhD, Jonathan A. Bernstein MD, PhD & Gill Bejerano PhD

[Affiliations](#) | [Corresponding author](#)

Genetics in Medicine (2016) | doi:10.1038/gim.2016.88

Received 29 February 2016 | Accepted 16 May 2016 | Published online 21 July 2016

Open Access

Outcomes

Open Data

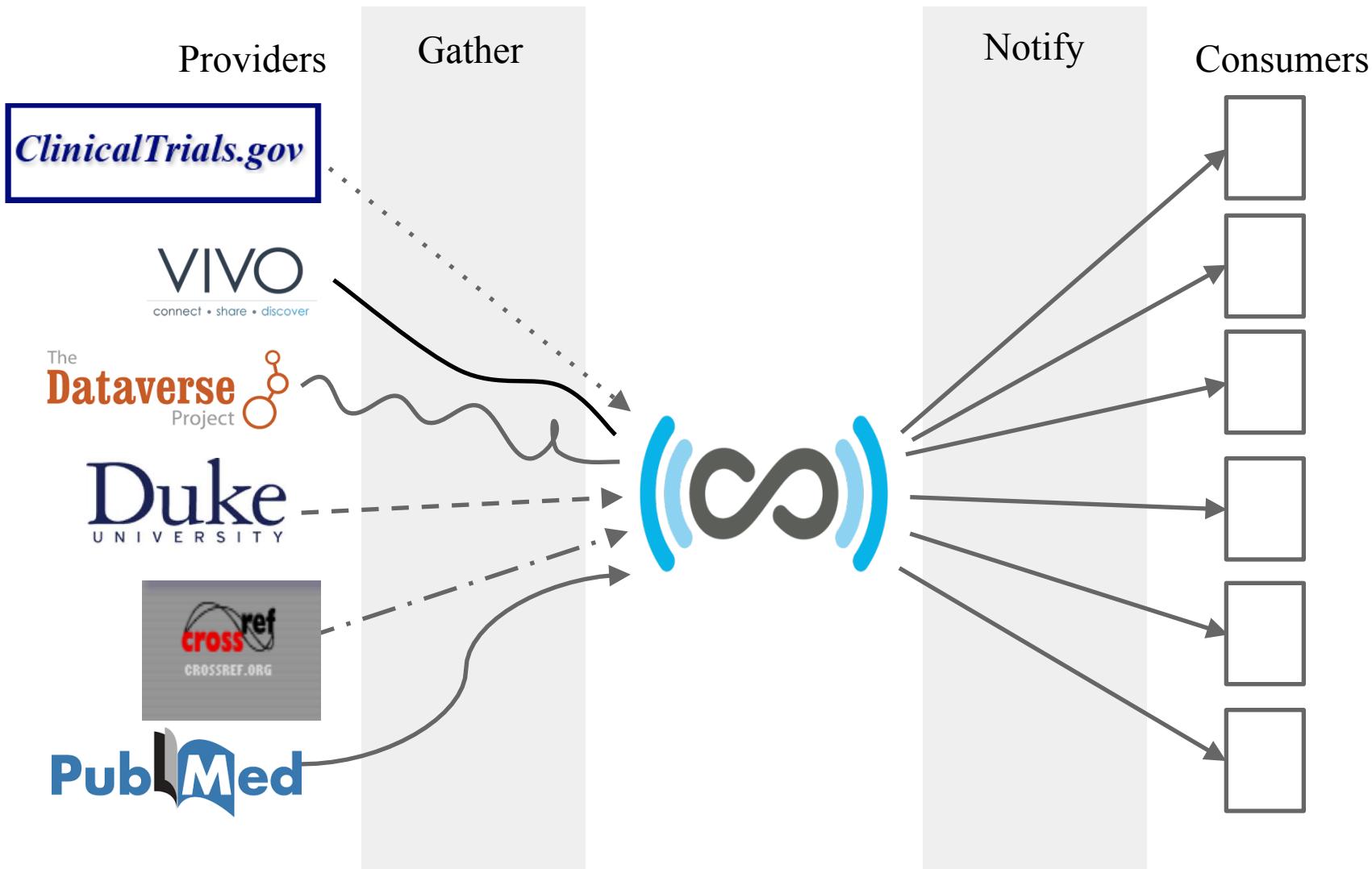
Content



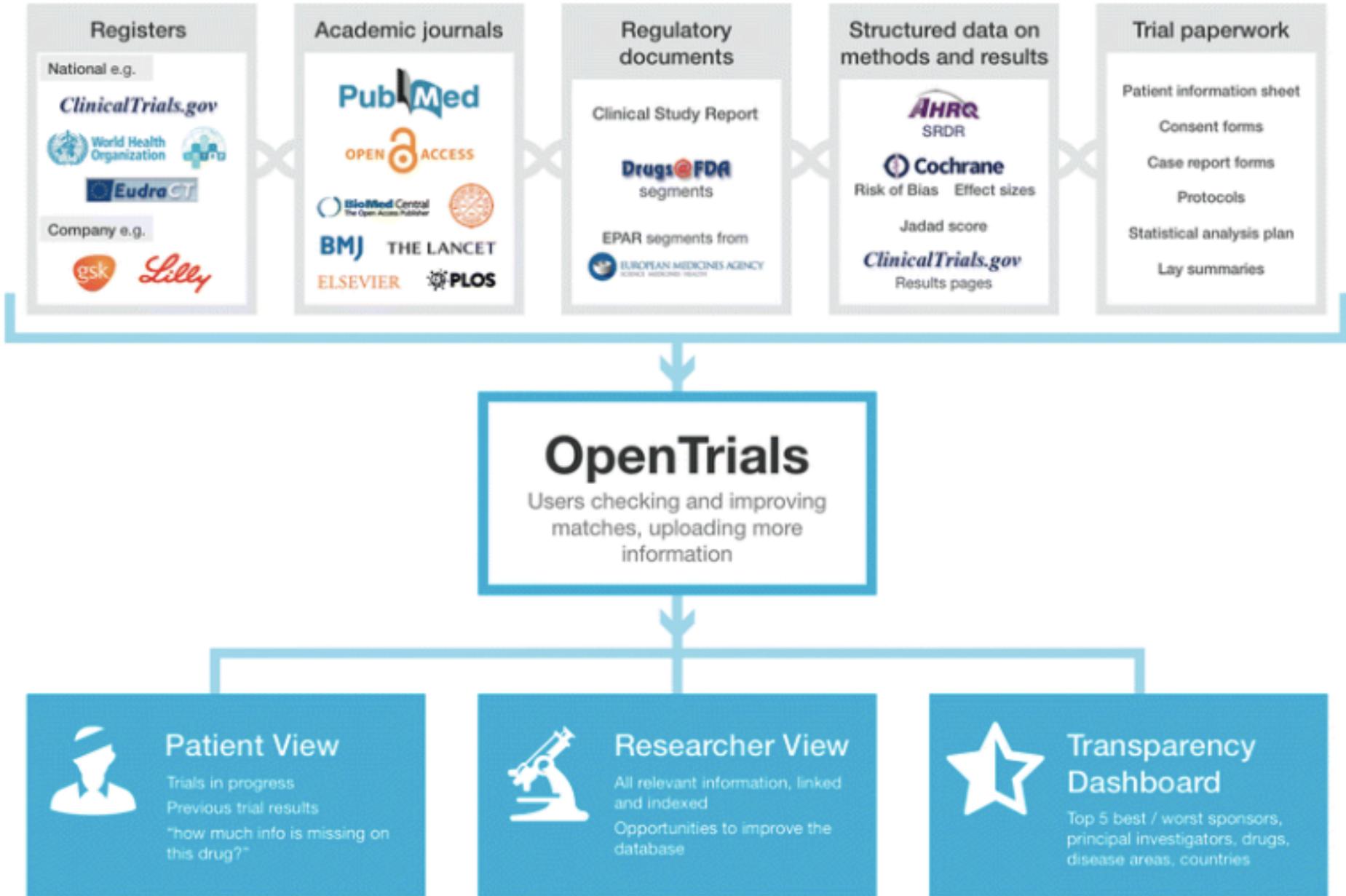
More than just data access, sharing, and compliance

Access is restricted/hidden by

- Licenses
- Terms of service
- Lack of programmatic access
- Pay walls
- Unavailability
- Inconsistency
- Low quality



<https://share.osf.io>



Evidence to *encourage* change

Technology to *enable* change

Training to *enact* change

Incentives to *embrace* change

- Reporting guidelines
 - <http://www.equator-network.org/reporting-guidelines/>
- Funder/government mandates
 - <http://datasharing.sparcopen.org>
- Journal policies
 - <https://cos.io/top/>
- Rewarding open practices
 - Badges to acknowledge open practices (Kidwell, 2016)

Promoting an open research culture

Author guidelines for journals could help to promote transparency, openness, and reproducibility

By B. A. Nosek,* G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. A. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. Levy Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, T. Yarkoni

Transparency, openness, and reproducibility are readily recognized as



749 Journals
62 Organizations

TOP Design Principles

Low barrier to entry: 3 levels
(disclose, require, verify)

Modular (8 standards)

Agnostic to discipline



TOP Guidelines

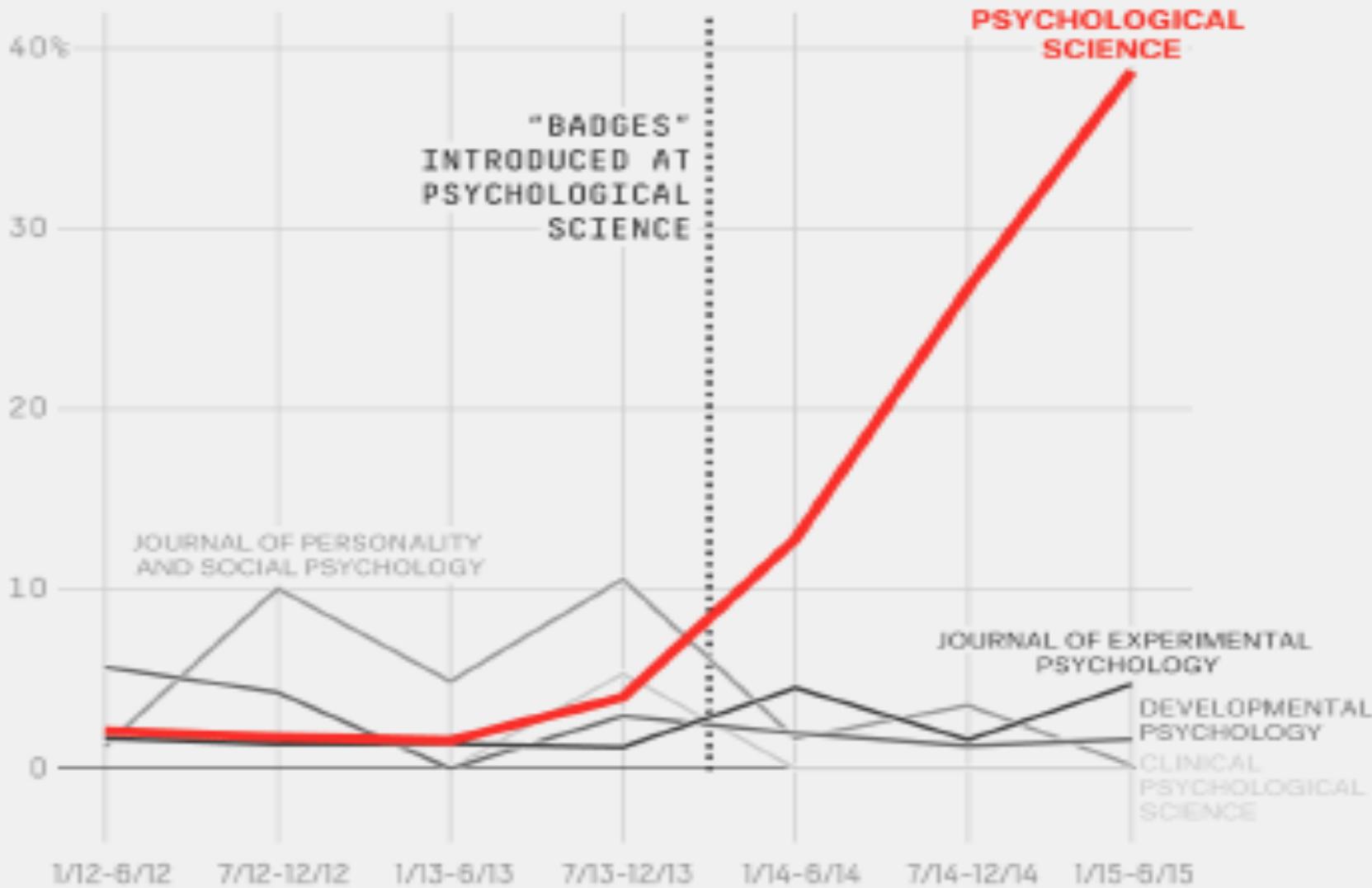
1. Data citation
2. Design transparency
3. Research materials transparency
4. Data transparency
5. Analytic methods (code) transparency
6. Preregistration of studies
7. Preregistration of analysis plans
8. Replication

Incentives: Making Behaviors Visible Promotes Adoption



Data sharing rose when it was rewarded

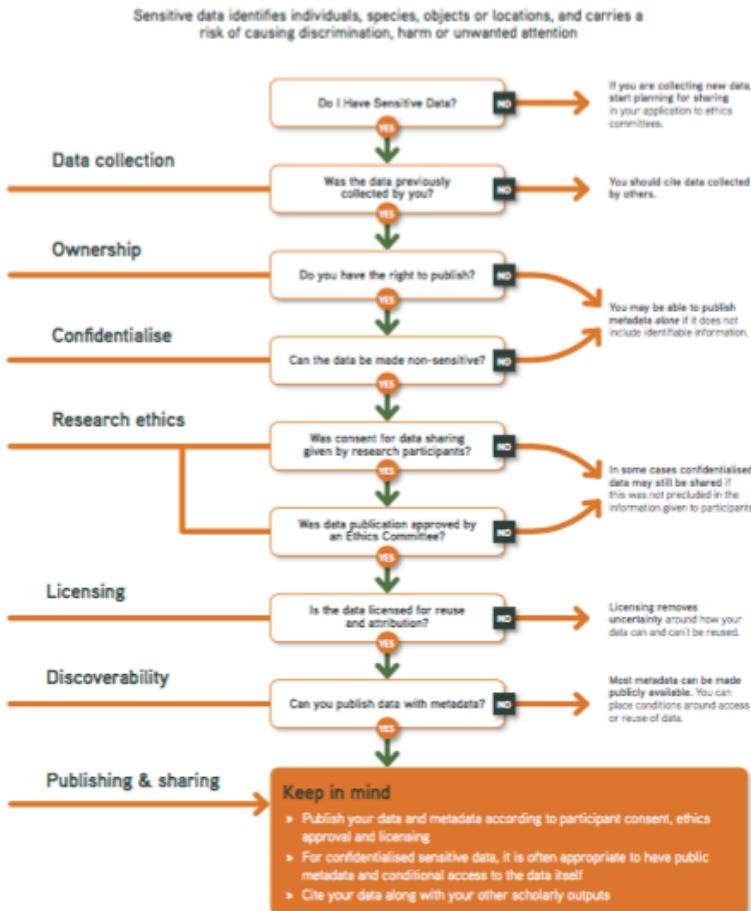
Share of papers in four psych journals reporting open data



Sharing Sensitive Data

PUBLISHING AND SHARING SENSITIVE DATA

When and how to publish sensitive data as openly and ethically as possible
For more information see: ands.org.au/sensitivedata



Tag Type	Description	Security Features	Access Credentials
Blue	Public	Clear storage, Clear transmit	Open
Green	Controlled public	Clear storage, Clear transmit	Email- or OAuth Verified Registration
Yellow	Accountable	Clear storage, Encrypted transmit	Password, Registered, Approval, Click-through DUA
Orange	More accountable	Encrypted storage, Encrypted transmit	Password, Registered, Approval, Signed DUA
Red	Fully accountable	Encrypted storage, Encrypted transmit	Two-factor authentication, Approval, Signed DUA
Crimson	Maximally restricted	Multi-encrypted storage, Encrypted transmit	Two-factor authentication, Approval, Signed DUA

Find this presentation at:

<https://osf.io/v38j5/>

Errington HESI 2016.06.08.pptx

[Delete](#) [Check out](#) [Share](#) [Download](#) [View](#) [Revisions](#)

The screenshot shows a file listing for 'OSF Storage' containing various presentation files. On the right, the presentation slide for 'Errington HESI 2016.06.08.pptx' is displayed. The slide features the Center for Open Science logo, the title 'Improving Openness and Reproducibility of Scientific Research', the author 'Tim Errington', and the URL 'http://cos.io/'. Logos for the Alfred P. Sloan Foundation, John Templeton Foundation, Association of Research Libraries, National Institute of Mental Health, National Institute on Aging, National Science Foundation, and the Laura and John Arnold Foundation are also visible.

Presentations

- OSF Storage

- 2015.10.GHC.general.share.pptx
- 20160107_cendl_spies.pptx
- 20160128_uva_dev_psych_spies.pptx
- 20160205_rpl_rcos_spies.pptx
- 20160308_sparc_spies.pptx
- 20160324_cos_spies.pptx
- 20160408_af_spies.pptx
- 20160408_ec_spies.pptx
- 20160408_samr_spies.pptx
- 20160427_arl_walters_spies.pptx
- 20160607_lwsg_spies.pptx
- Bowman.ACS.2015.08.17.pptx

Tags

add a tag

Page: 1 of 56

Automatic Zoom

COS CENTER FOR OPEN SCIENCE

Improving Openness and Reproducibility of Scientific Research

Tim Errington
Center for Open Science
<http://cos.io/>

ALFRED P. SLOAN FOUNDATION JOHN TEMPLETON FOUNDATION SUPPORTING SCIENCE - INVESTING IN THE BIG QUESTIONS

ASSOCIATION OF RESEARCH LIBRARIES NIH National Institute of Mental Health NIH National Institute on Aging NSF THE WILLIAM AND FLORA HEWLETT FOUNDATION

Questions: **tim@cos.io**

Sources:

- Merton, 1942: Journal of Legal and Political Sociology, 1, 115-126, 1942
- Anderson, Martinson, & DeVries, 2007: DOI: 10.1525/JERHRE.2007.2.4.3
- Lyon, 2009: <http://opus.bath.ac.uk/30056/1/open-science-report-6nov09-final-sentojisc.pdf>
- Mhaskar, 2009: doi: [10.4103/0253-7184.62770](https://doi.org/10.4103/0253-7184.62770)
- Cazzola, 2015: DOI: 10.1183/16000617.00002215
- Scherer, 2012: <http://blog.efpsa.org/2012/06/01/falsification-of-previous-results/>
- www.COMPare-trials.org
- Ebrahim, 2014: doi:10.1001/jama.2014.9646.
- Wenger, 2016: doi:10.1038/gim.2016.88
<https://share.osf.io>
- Goldacre & Gray, 2016: DOI: 10.1186/s13063-016-1290-8
- opentrials.net
<http://www.equator-network.org/reporting-guidelines/>
- <http://datasharing.sparcopen.org>
- <https://cos.io/top/>
- Kidwell, 2016: <http://dx.doi.org/10.1371/journal.pbio.1002456>
- <http://fivethirtyeight.com/features/even-psychologists-respond-to-meaningless-rewards/>
- Sweeney, 2015: <http://techscience.org/a/2015101601/>
- <http://www.ands.org.au>

RESEARCH ARTICLE

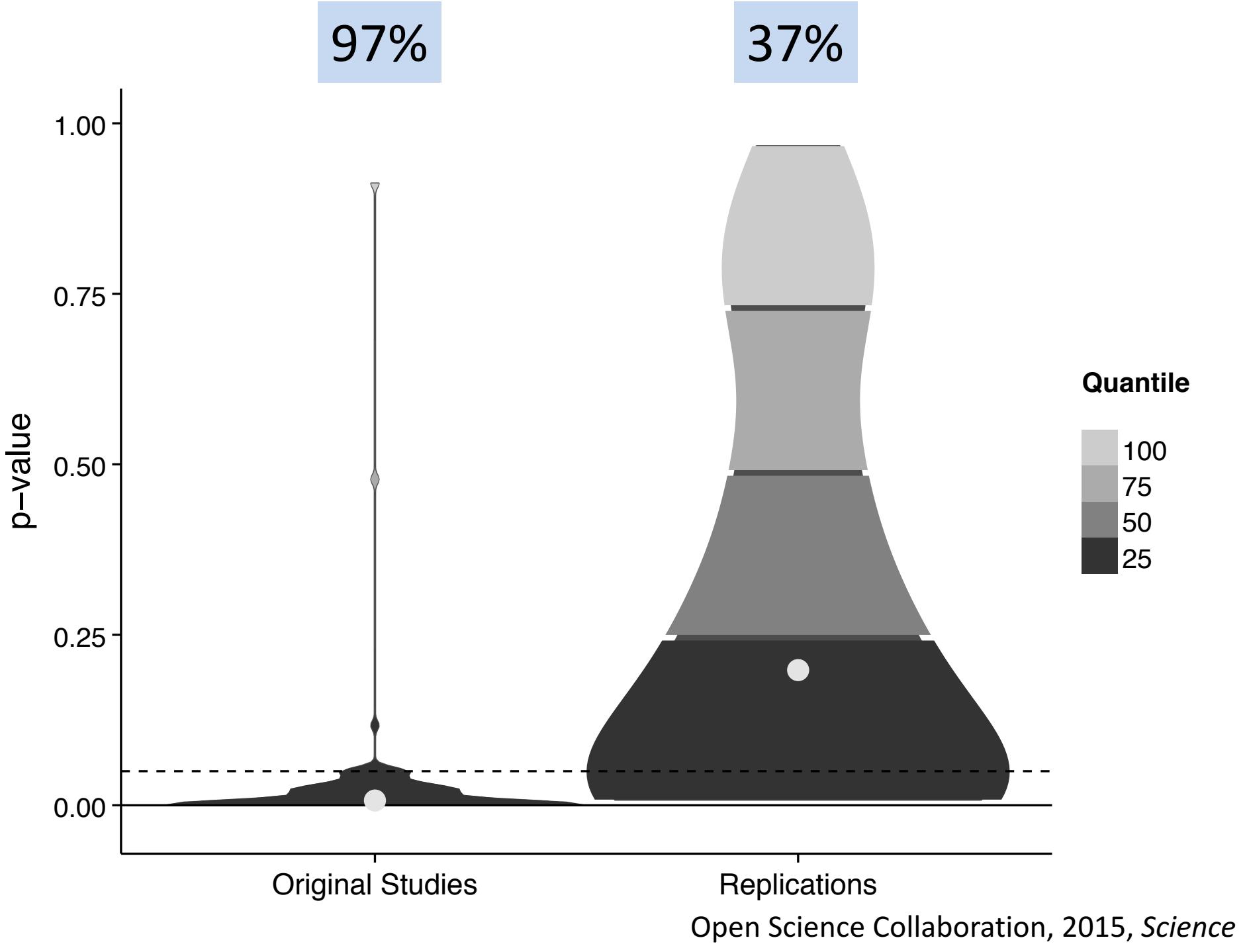
Estimating the reproducibility of psychological science

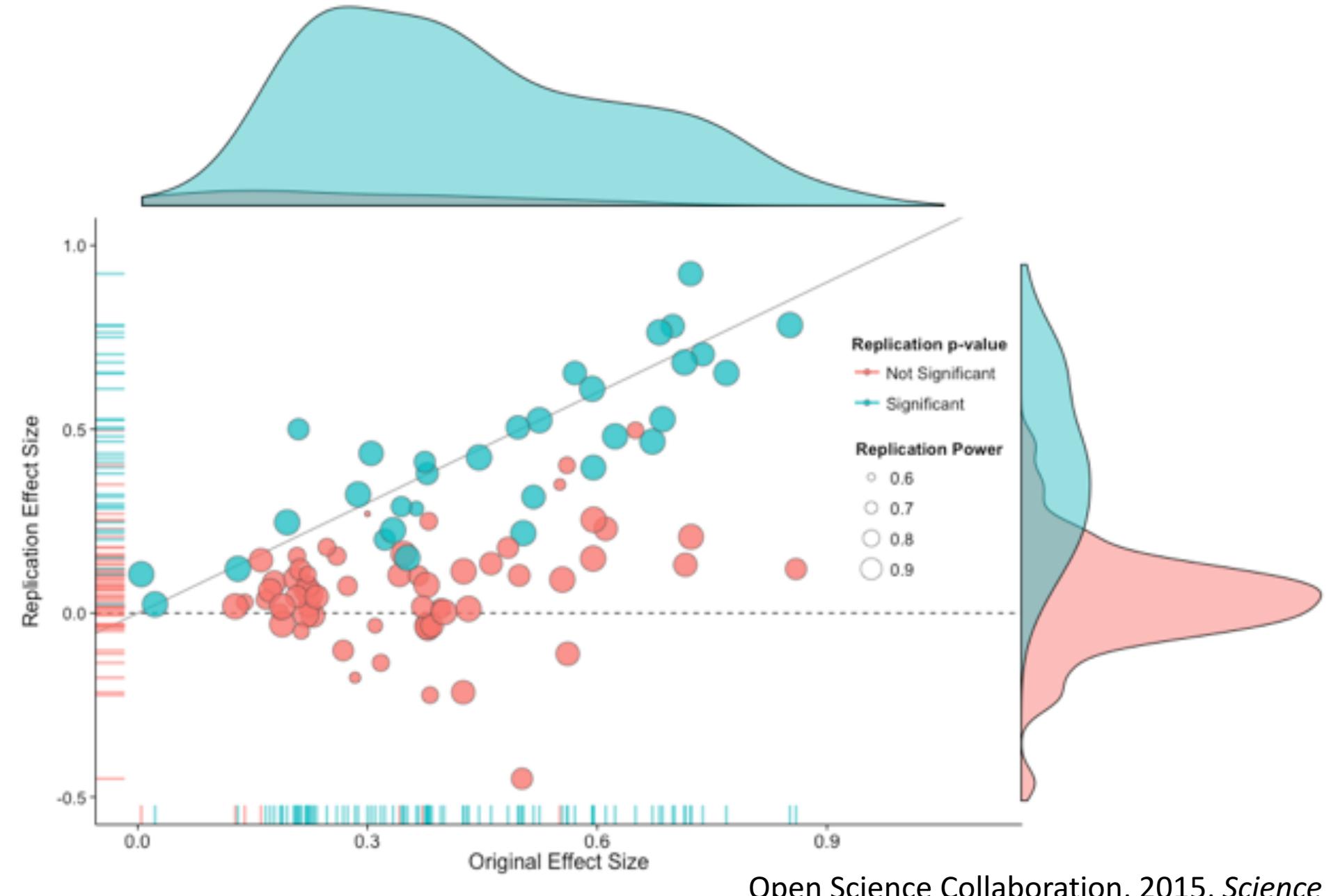
Open Science Collaboration^{*,†}

+ Author Affiliations

✉[†]Corresponding author. E-mail: nosek@virginia.edu

Science 28 Aug 2015:
Vol. 349, Issue 6251, pp.
DOI: 10.1126/science.aac4716

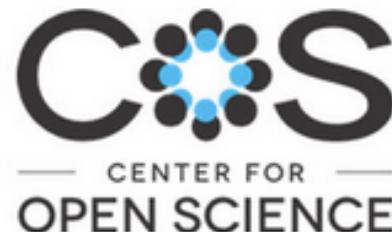






REPRODUCIBILITY PROJECT Cancer Biology

The Reproducibility Project: Cancer Biology is a collaboration between the [Center for Open Science](#) and [Science Exchange](#) to independently replicate selected results from [50 papers in cancer biology](#). For each paper a Registered Report detailing the proposed experimental designs and protocols for the replications is peer reviewed and published prior to data



A blue-tinted photograph of several laboratory glass containers, including beakers and flasks, arranged in a blurred, overlapping composition.

REPRODUCIBILITY PROJECT

Cancer Biology

Registered Reports

- The team replicating a previously published study first submits a Registered Report that explains how it intends to replicate selected experiments from the original paper
- Each Registered Report is peer reviewed by several experts, including a biostatistician
- Once the Registered Report has been revised satisfactorily, it will be published

REPRODUCIBILITY PROJECT
Cancer Biology

Registered Reports

- The replication team then starts to replicate the experiments, following the protocols detailed in the Registered Report
- Irrespective of the outcome, the results will be published as a Replication Study after peer review to check that the experiments were carried out in accordance with the protocols outlined in the Registered Report

CELL BIOLOGY | HUMAN BIOLOGY AND MEDICINE

Melanoma genome sequencing reveals frequent *PREX2* mutations



ORIGINAL ARTICLE

May 9, 2012

MF Berger, E Hodis, TP Heffernan,
YL Deribe, MS Lawrence, A Protopopov,
E Ivanova, IR Watson, E Nickerson,
P Ghosh et al.

Nature 2012;485:502-506
[10.1038/nature11071](https://doi.org/10.1038/nature11071)



REGISTERED REPORT

December 10, 2014

Denise Chroscinski, Darryl Sampey,
Alex Hewitt,
Reproducibility Project: Cancer Biology

eLife 2014;3:e04180
[10.7554/eLife.04180](https://doi.org/10.7554/eLife.04180)



REPLICATION STUDY

Forthcoming

Preregistration

Purposes

1. Discoverability: Study exists
2. Interpretability: Distinguish exploratory and confirmatory approaches

Why needed?

Mistaking exploratory as confirmatory *increases publishability* and *decreases credibility* of results

Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time

Robert M. Kaplan^{1*}, Veronica L. Irvin²

1 Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, Rockville, Maryland, United States of America, 2 Oregon State University, Corvallis, Oregon, United States of America

*  Positive Result Rate dropped from 57% to 8% after preregistration required.

compared treatment to placebo versus active comparator. Industry co-sponsorship was unrelated to the probability of reporting a significant benefit. Pre-registration in clinicaltrials.gov was strongly associated with the trend toward null findings.

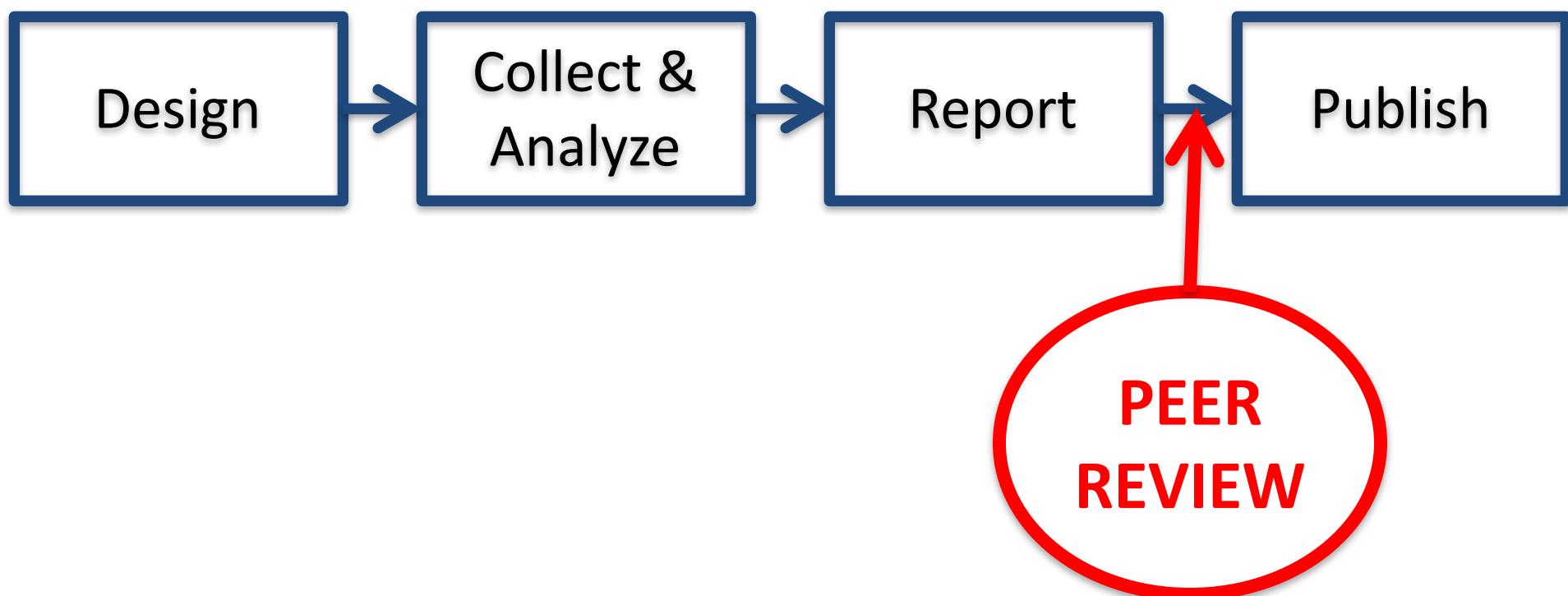
The Pre-registration Challenge



One thousand scientists will win \$1,000 each for publishing the results of their preregistered research.

<https://cos.io/prereg>

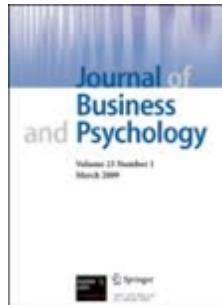
Registered Reports



Who Publishes Registered Reports?

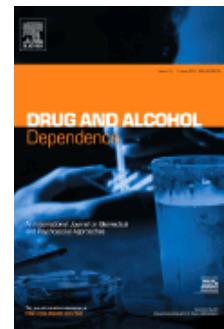


Neuroscience



Attention,
Perception, &
Psychophysics

AP&P



eLIFE

(just to name a few)

See the full list and compare features: osf.io/8mpji