

Open Science Q & A

Brian Nosek

University of Virginia and Center for Open Science

5 April 2019

University of Melbourne, Department of Psychology

Open Science is...

Showing and sharing your work

Work = process, content (data, materials, code), and outcomes



Question

What are our blind spots?

That is, what are the most important areas for improvement in scientific practice which are currently being ignored by the open science movement?

What are our blind spots in open science?

- Replicability *versus* Validity
- Statistical inferences *versus* Theoretical claims/conclusions
- Doing It *versus* Doing It “Right”
- All-or-None *versus* Incrementalism
- Social Identity *versus* Good Practice

Preregistration



Question

Do you think pre-registration can or should apply to computational model building?

If so, how can it be done in a way that is both feasible and preserves the essential theory-building component of model development?

Context of Justification

Confirmation

Data independent

Hypothesis testing

p -values interpretable

Context of Discovery

Exploration

Data contingent

Hypothesis generating

p -values not so interpretable

**Presenting exploratory as confirmatory
increases *publishability* of results at the
cost of *credibility* of results**

PREREGISTRATION

Context of Justification

Confirmation

Data independent

Hypothesis testing

p-values interpretable

Context of Discovery

Exploration

Data contingent

Hypothesis generating

p-values not so interpretable

Presenting exploratory results as confirmatory increases *publication* of results at the cost of *creativity* of results

Presenting exploratory results as confirmatory increases *publication* of results at the cost of *creativity* of results

Improving Practices

Better Theory, Better Reproducibility, just
Better?



Question

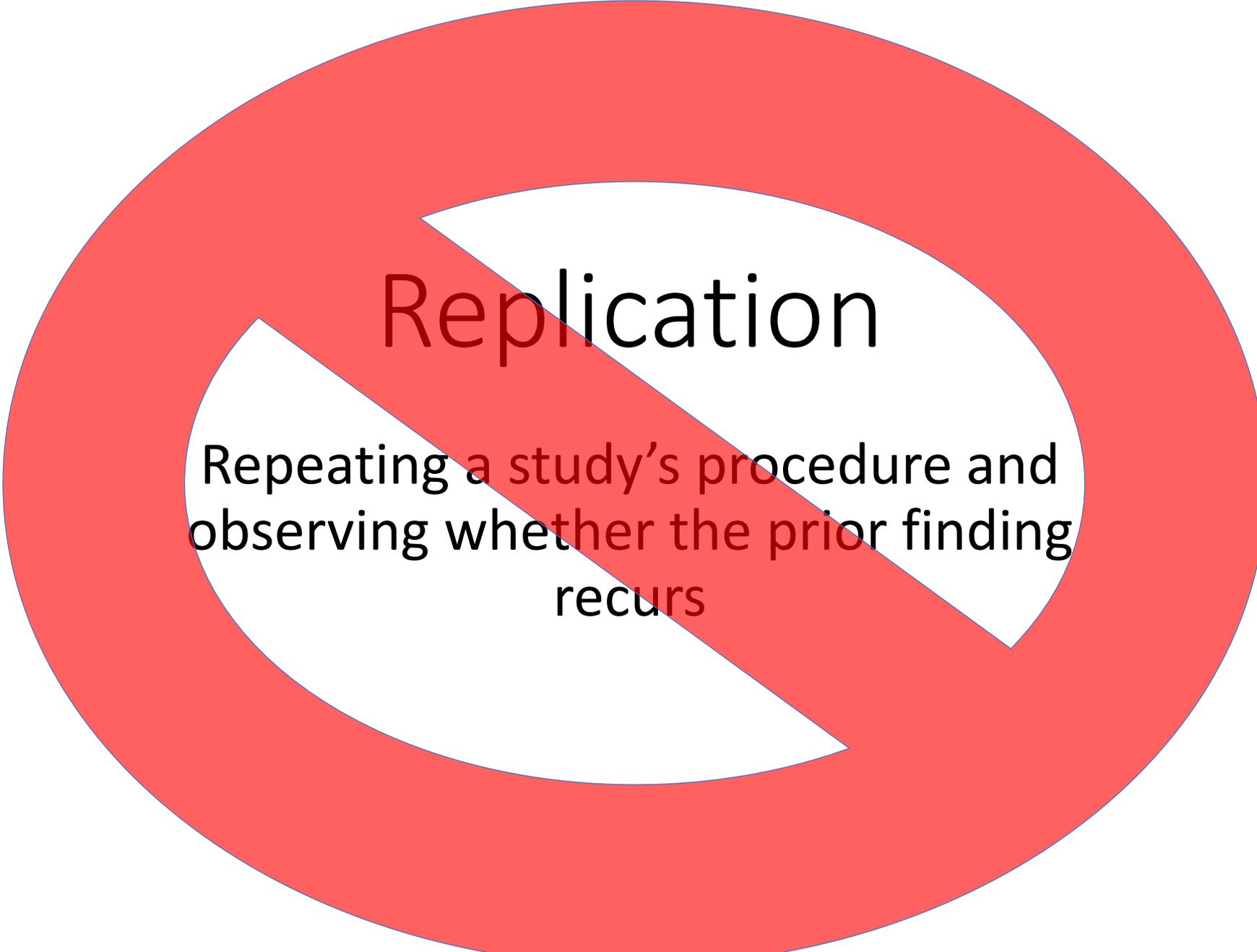
What is the role of theory in the open science movement? Can gathering more reliable "effects" make up for poor theory development?

Theory Development versus Effects Replicability

- False lead – Which comes first?
- False lead – Independent issues
- Better lead – Appreciate interdependence between reliable evidence and improving theoretical specificity
 - With lack of attention to replication, theories never needed to become specific

Replication

Repeating a study's procedure and observing whether the prior finding recurs



Replication

Repeating a study's procedure and observing whether the prior finding recurs

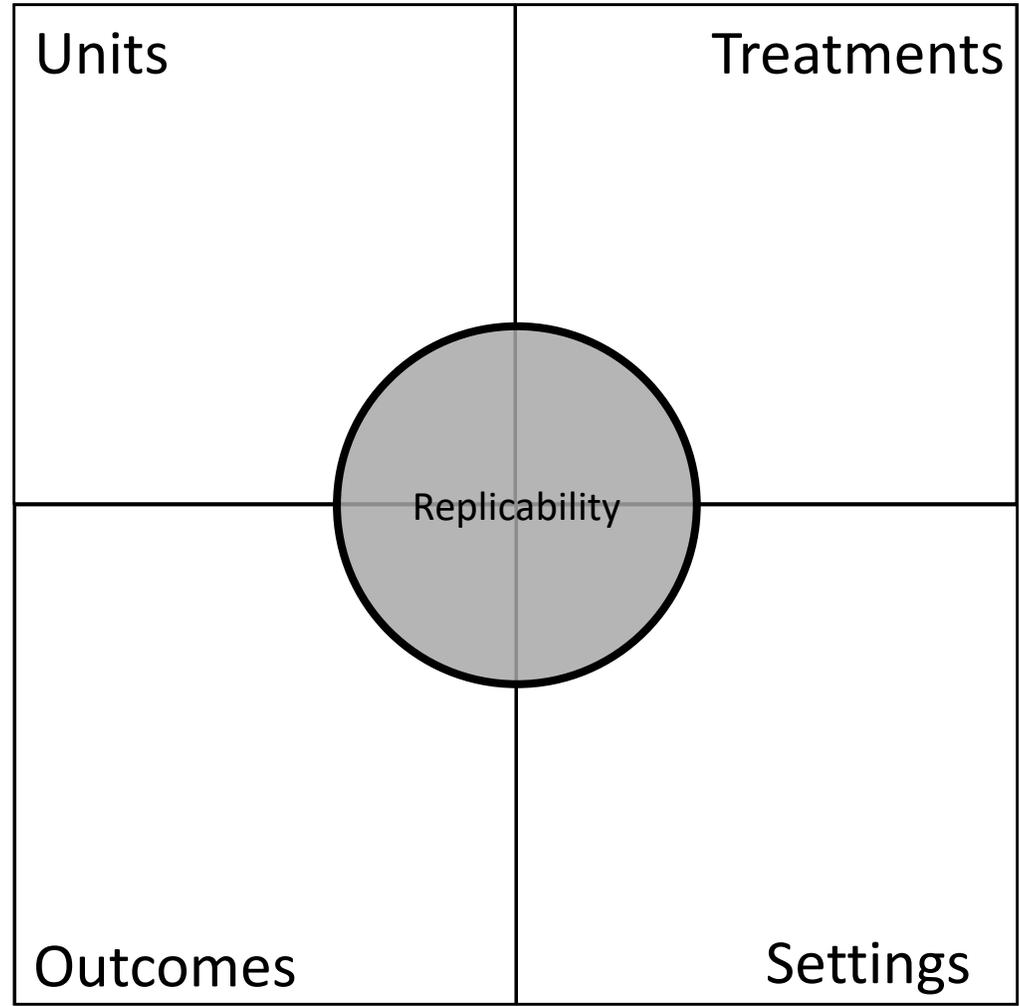


Replication

Attempting to reproduce a previously observed finding with no *a priori* reason to expect a different outcome



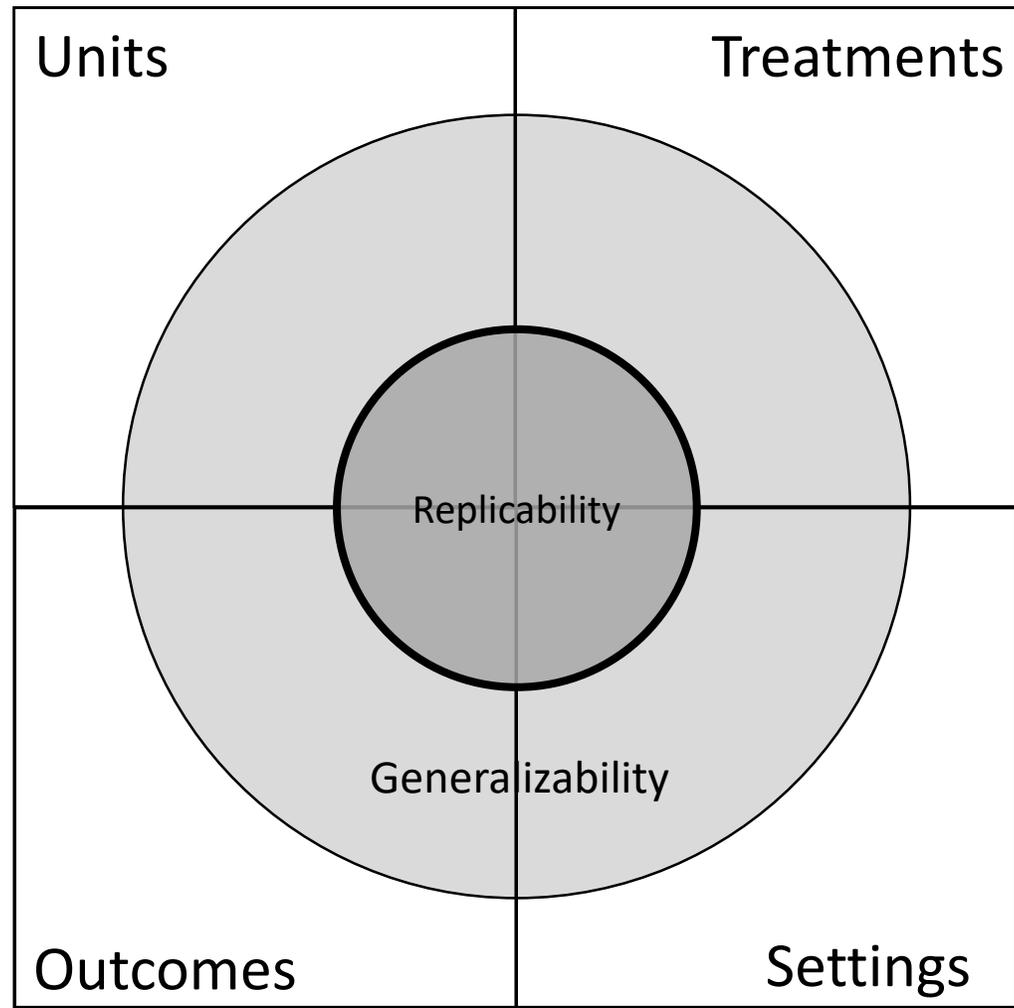
Claiming that a study is a replication is a theoretical claim





Why then procedural similarity?

- Lack of theoretical specificity
- Lack of understanding of methods
- Same procedures is a substitute for theoretical and methodological understanding





Replication

Attempting to reproduce a previously observed finding with no *a priori* reason to expect a different outcome

Implications

- Theoretical specificity enables effective replication
- Replication promotes theoretical specificity
- Understanding whether a study is a replication can change over time

Cultural Incentives



Question

What, if anything, can open science do to remove the problem of perverse incentives in academia?

Public response to science



Question

With rising skepticism towards science in the general public, who is responsible for restoring trust and belief in science? Is this a problem that should be addressed by institutions (Like COS), Universities or individual researchers?

Research Design and Statistical Practices



Question

How big of an issue do you think the continued use of convenience sampling in psychological research might be for reproducibility?

Many Labs 2

- Replications of 28 findings (2 slates)
- 124 samples (63 slate 1, 61 slate 2)
- >15,000 participants total
- 36 nations/regions (Serbia, Poland, New Zealand, Netherlands, Canada, Jamaica, UK, USA, Australia, South Africa, Colombia, Turkey, Costa Rica, Spain, Chile, Brazil, India, United Arab Emirates, Tanzania, Malaysia, Italy, Nigeria, Germany, Belgium, France, Czech Republic, Hong Kong, China, Uruguay, Sweden, Mexico, Portugal, Japan, Hungary, Switzerland, Taiwan)
- 186 co-authors

Tau

.14

.10

.10

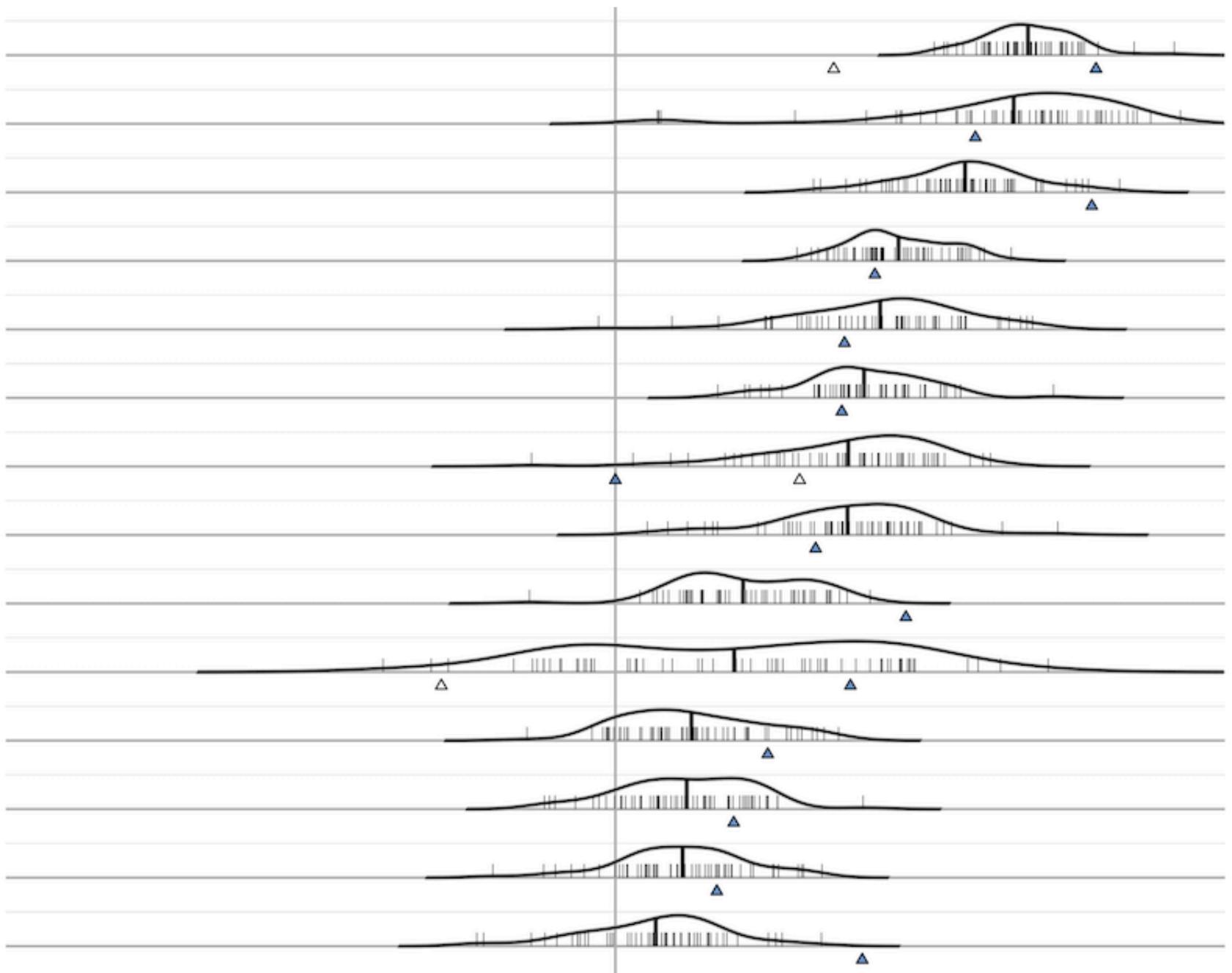
.10

.10

.24

.09

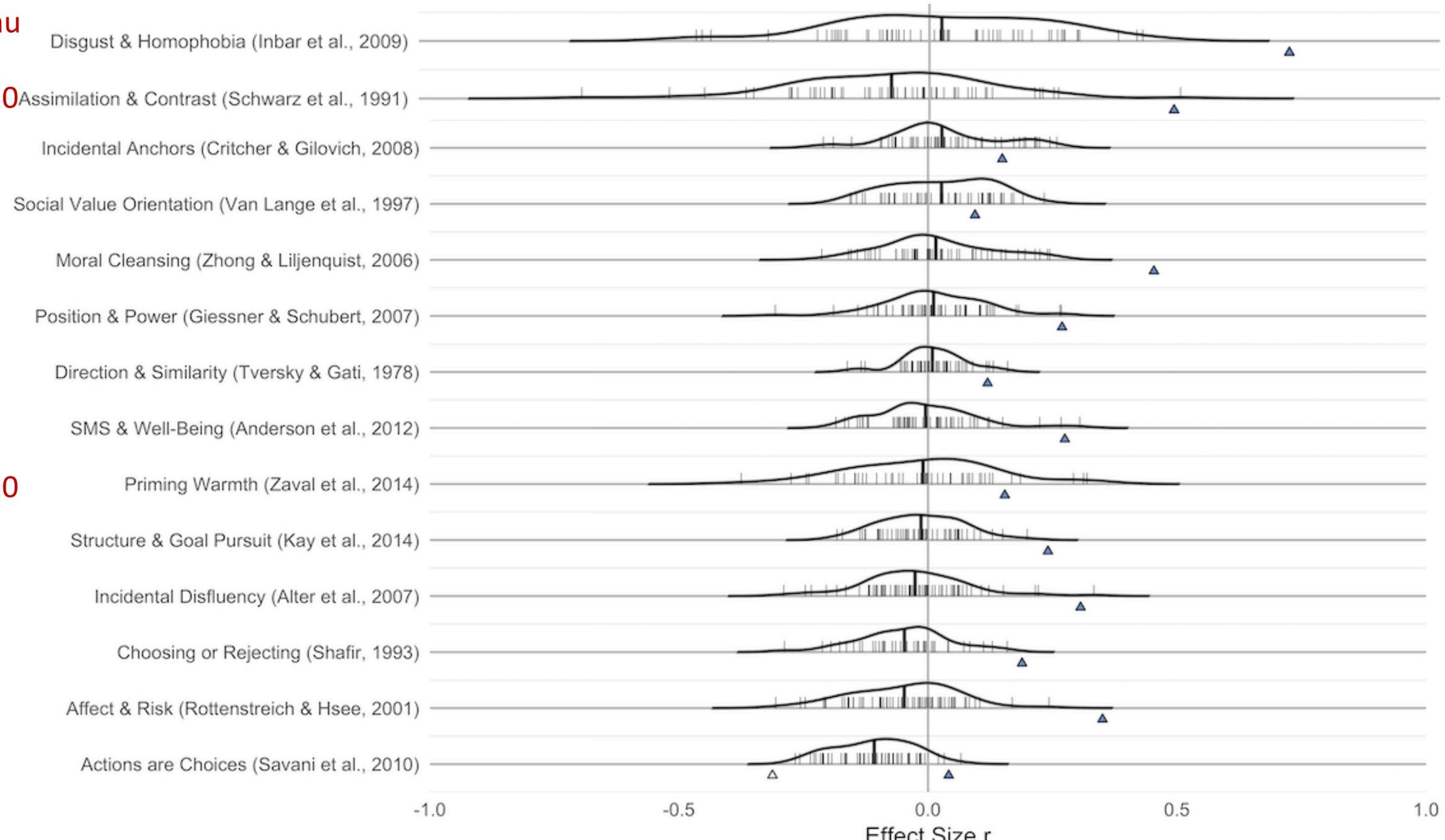
- Correspondence Bias (Miyamoto & Kitayama, 2002)
- Intentional Side-Effects (Knobe, 2003)
- Trolley Dilemma 1 (Hauser et al., 2007)
- False Consensus 1 (Ross et al., 1977)
- Moral Typecasting (Gray & Wegner, 2009)
- False Consensus 2 (Ross et al., 1977)
- Intuitive Reasoning (Norenzayan et al., 2002)
- Less is Better (Hsee, 1998)
- Framing (Tversky & Kahneman, 1981)
- Direction & SES (Huang et al., 2014)
- Moral Foundations (Graham et al., 2009)
- Tempting Fate (Risen & Gilovich, 2008)
- Trolley Dilemma 2 (Hauser et al., 2007)
- Priming Consumerism (Bauer et al., 2012)



Tau

.10

.10





Question

Has there been any up-take in the recommendation to change alpha to .005?

Data Sharing and Use



Question

While sharing of ideas, data, etc is great and noble from the sharer's perspective, how do we ensure better practice from the "sharee"?

That is, how do we stop people misusing open data (cherry picking, interpreting out of context, claiming ownership, etc)?



Question

Openly sharing data that contains personally sensitive information violates people's right to privacy.

However, computational privacy experts emphasise that de-identification is really not possible in many cases.

Privacy concerns seem to be in direct conflict with open science principles. How do you see this evolving?



Question

Our ethics protocols require that participants provide informed consent. An important part of informed consent is that the use to which the data will be put is revealed.

However, once the data has been published it seems like the horse has bolted. How can we ensure that the data is only put to the purposes for which people have provided consent?

These slides available at:

<https://osf.io/y28bf/>