

Assessing Reproducibility

Caren M. Rotello
Examining the Replication Crisis
ISSR Panel
April 8, 2016

Current Strategies in Psychology

- Pipeline Project
- Many Labs Project
- Reproducibility Project: Psychology
- Registered replications

Core Idea: Pipeline Project

- Before submitting your experiments for publication, have other labs replicate them for you
 - You are free to negotiate the details, including use of materials

The Pipeline Project

ARTICLE IN PRESS

YJESP-03368; No. of pages: 13; 4C: 5

Journal of Experimental Social Psychology xxx (2015) xxx-xxx



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

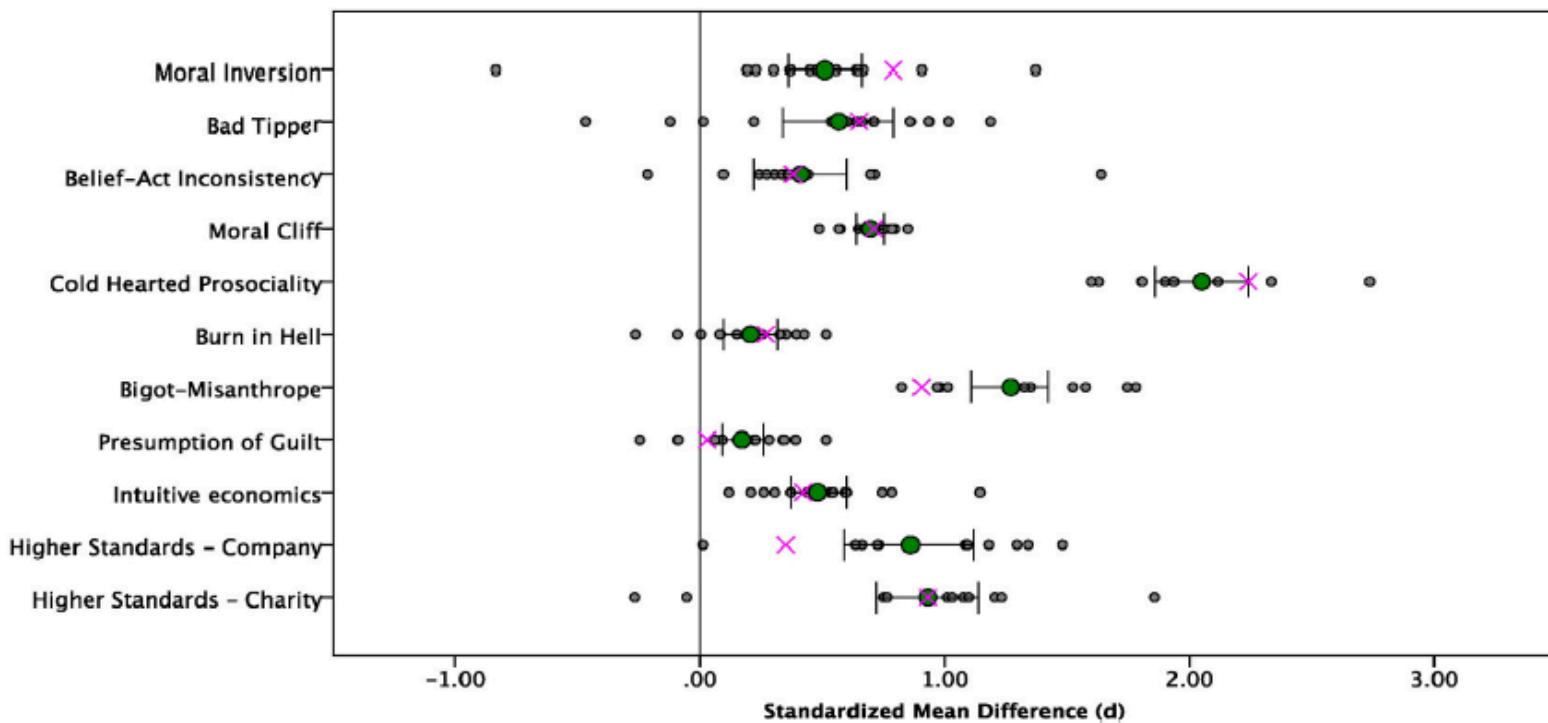
journal homepage: www.elsevier.com/locate/jesp



The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline

Martin Schweinsberg ^{a,*}, Nikhil Madan ^a, Michelangelo Vianello ^b, S. Amy Sommer ^c, Jennifer Jordan ^d, Warren Tierney ^a, Eli Awtrey ^e, Luke Lei Zhu ^f, Daniel Diermeier ^g, Justin E. Heinze ^h, Malavika Srinivasan ⁱ, David Tannenbaum ^g, Eliza Bivolaru ^a, Jason Dana ^j, Clintin P. Davis-Stober ^k, Christilene du Plessis ^l, Quentin F. Gronau ^m, Andrew C. Hafenbrack ⁿ, Eko Yi Liao ^o, Alexander Ly ^m, Maarten Marsman ^m, Toshio Murase ^p, Israr Qureshi ^q, Michael Schaefer ^a, Nico Thornley ^a, Christina M. Tworek ^r, Eric-Jan Wagenmakers ^m, Lynn Wong ^a, Tabitha Anderson ^s, Christopher W. Bauman ^t, Wendy L. Bedwell ^u, Victoria Brescoll ^j, Andrew Canavan ^s, Jesse J. Chandler ^v, Erik Cheries ^w, Sapna Cheryan ^e, Felix Cheung ^{x,y}, Andrei Cimpian ^r, Mark A. Clark ^z, Diana Cordon ^s, Fiery Cushman ⁱ, Peter H. Ditto ^t, Thomas Donahue ^s, Sarah E. Frick ^u, Monica Gamez-Djokic ^{aa}, Rebecca Hofstein Grady ^t, Jesse Graham ^{ab}, Jun Gu ^{ac}, Adam Hahn ^{ad}, Brittany E. Hanson ^{ae}, Nicole J. Hartwich ^{ad}, Kristie Hein ^s, Yoel Inbar ^{af}, Lily Jiang ^e, Tehlyr Kellogg ^s, Deanna M. Kennedy ^{ag}, Nicole Legate ^s, Timo P. Luoma ^{ad}, Heidi Maibuecher ^s, Peter Meindl ^{ab}, Jennifer Miles ^t, Alexandra Mislin ^z, Daniel C. Molden ^{aa}, Matt Motyl ^{ae}, George Newman ^j, Hoai Huong Ngo ^{ah}, Harvey Packham ^{ai}, Philip S. Ramsay ^u, Jennifer L. Ray ^{aj}, Aaron M. Sackett ^{ak}, Anne-Laure Sellier ^c, Tatiana Sokolova ^{cal}, Walter Sowden ^h, Daniel Storage ^r, Xiaomin Sun ^{am}, Jay J. Van Bavel ^{aj}, Anthony N. Washburn ^{ae}, Cong Wei ^{am}, Erik Wetter ^{an}, Carlos T. Wilson ^s, Sophie-Charlotte Darroux ^a, Eric Luis Uhlmann ^{a,*}

Pipeline Results



Challenges: Pipeline Project

- For replicating lab:
 - Labor-intensive, expensive
 - No direct payoff, no funding
- For replicated lab:
 - Not clear what to do in cases of failures to replicate
- Not yet clear the experiments are of interest to anyone else!

The Many Labs Project

- 13 “classic and contemporary” effects selected for replication
 - Short, simple task (combine into single session)
 - Easy to administer online (MTurk)
 - Diversity of topic, study age, and impact
- 36 different samples run (min. sample size = 80)
 - US, international
 - Lab based, online

The Many Labs Project

Replication

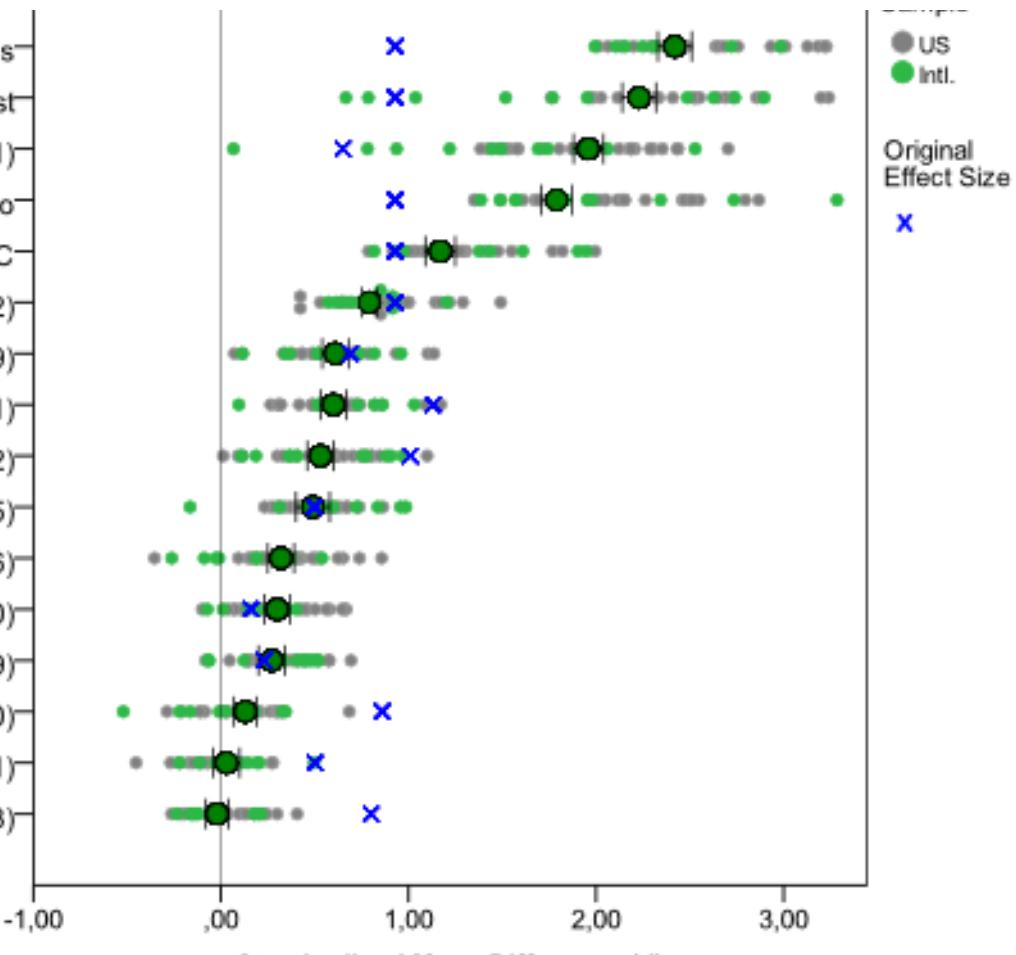
Investigating Variation in Replicability

A “Many Labs” Replication Project

Richard A. Klein,¹ Kate A. Ratliff,¹ Michelangelo Vianello,² Reginald B. Adams Jr.,³ Štěpán Bahník,⁴ Michael J. Bernstein,⁵ Konrad Bocian,⁶ Mark J. Brandt,⁷ Beach Brooks,¹ Claudia Chloe Brumbaugh,⁸ Zeynep Cemalcilar,⁹ Jesse Chandler,^{10,36} Winnee Cheong,¹¹ William E. Davis,¹² Thierry Devos,¹³ Matthew Eisner,¹⁰ Natalia Frankowska,⁶ David Furrow,¹⁵ Elisa Maria Galliani,² Fred Hasselman,^{16,37} Joshua A. Hicks,¹² James F. Hovermale,¹⁷ S. Jane Hunt,¹⁸ Jeffrey R. Huntsinger,¹⁹ Hans IJzerman,⁷ Melissa-Sue John,²⁰ Jennifer A. Joy-Gaba,¹⁷ Heather Barry Kappes,²¹ Lacy E. Krueger,¹⁸ Jaime Kurtz,²² Carmel A. Levitan,²³ Robyn K. Mallett,¹⁹ Wendy L. Morris,²⁴ Anthony J. Nelson,³ Jason A. Nier,²⁵ Grant Packard,²⁶ Ronaldo Pilati,²⁷ Abraham M. Rutchick,²⁸ Kathleen Schmidt,²⁹ Jeanine L. Skorinko,²⁰ Robert Smith,¹⁴ Troy G. Steiner,³ Justin Storbeck,⁸ Lyn M. Van Swol,³⁰ Donna Thompson,¹⁵ A. E. van ‘t Veer,⁷ Leigh Ann Vaughn,³¹ Marek Vranka,³² Aaron L. Wichman,³³ Julie A. Woodzicka,³⁴ and Brian A. Nosek^{29,35}

The Many Labs Project

- Anchoring (Jacowitz & Kahneman, 1995) - Babies
- Anchoring (Jacowitz & Kahneman, 1995) - Everest
- Allowed/Forbidden (Rugg, 1941)
- Anchoring (Jacowitz & Kahneman, 1995) - Chicago
- Anchoring (Jacowitz & Kahneman, 1995) - NYC
- Corr. between I and E math attitudes (Nosek et al., 2002)
- Retro. gambler's fallacy (Oppenheimer & Monin, 2009)
- Gain vs loss framing (Tversky & Kahneman, 1981)
- Sex diff. in implicit math attitudes (Nosek et al., 2002)
- Low-vs.-high category scales (Schwarz et al., 1985)
- Quote Attribution (Lorge & Curtis, 1936)
- Norm of reciprocity (Hyman and Sheatsley, 1950)
- Sunk costs (Oppenheimer et al., 2009)
- Imagined contact (Husnu & Crisp, 2010)
- Flag Priming (Carter et al., 2011)
- Currency priming (Caruso et al., 2013)



Challenges: Many Labs Project

- Non-random sample of effects, not all of general interest
- Nothing systematically influenced whether replication was successful

Reproducibility Project: Psychology (RPP)

RESEARCH ARTICLE SUMMARY

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*

INTRODUCTION: Reproducibility is a defining feature of science, but the extent to which

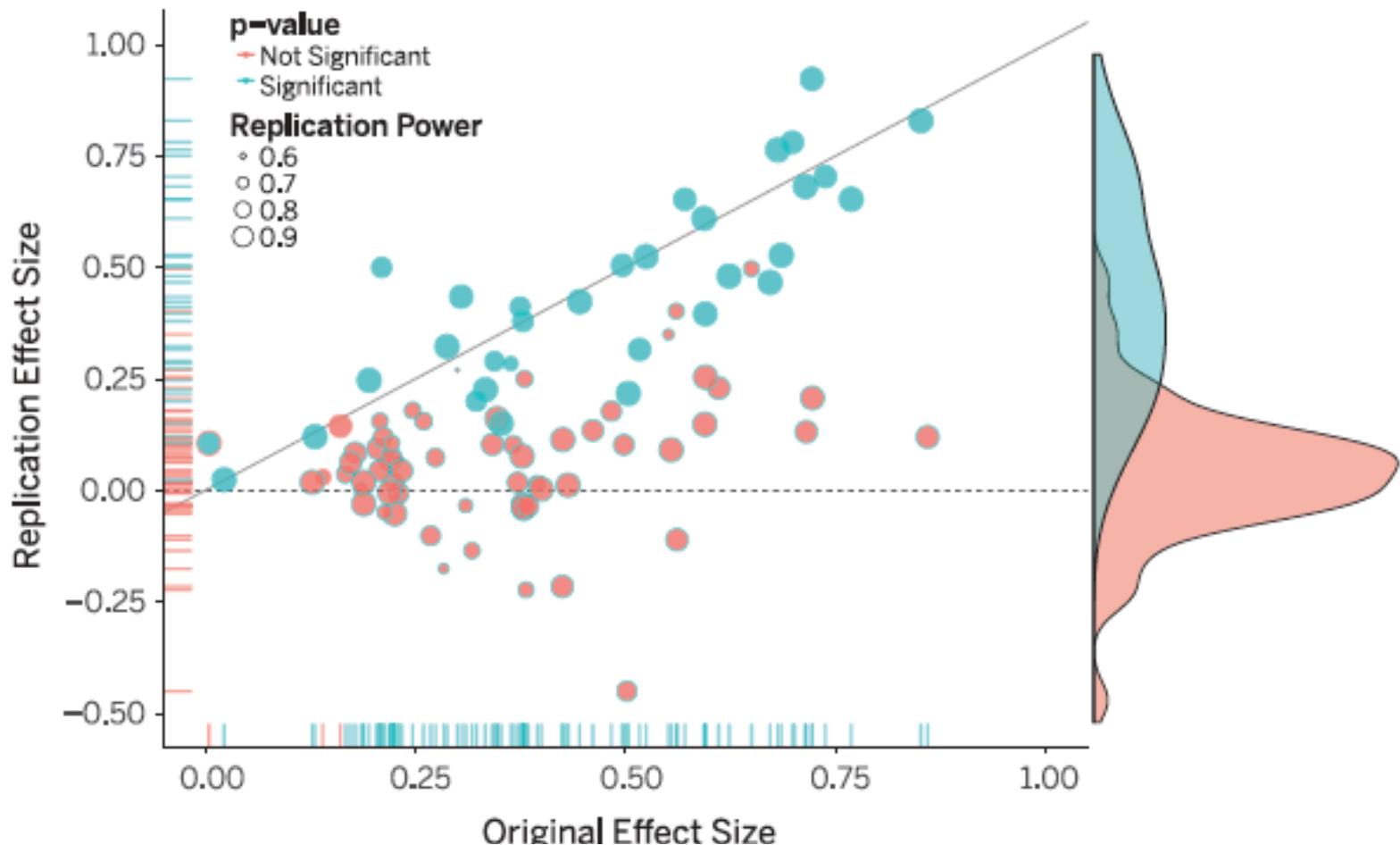
previously observed finding and is the means of establishing reproducibility of a finding with

Science, 2015

Core Idea: RPP

- 2008 volumes of 3 major journals
- Volunteers replicated an experiment from every paper

Reproducibility Project: Psychology



Challenges: RPP

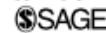
- Selection of experiments
 - Ease of implementation
 - Small original sample size
 - NOT based on theoretical significance

The first registered replication report, 2014



Registered Replication Report: Schooler and Engstler-Schooler (1990)

Perspectives on Psychological Science
2014, Vol. 9(5) 556–578
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691614545653
pps.sagepub.com



Proposing Authors: This proposal was initiated by the editors

Contributing authors (alphabetical order): Alogna, V. K., Attaya, M. K., Aucoin, P., Bahnik, Š., Birch, S., Birt, A. R., Bornstein, B. H., Bouwmeester, S., Brandimonte, M. A., Brown, C., Buswell, K., Carlson, C., Carlson, M., Chu, S., Cislak, A., Colarusso, M., Colloff, M. F., Dellapaoleta, K. S., Delvenne, J.-F., Di Domenico, A., Drummond, A., Echterhoff, G., Edlund, J. E., Eggleston, C. M., Fairfield, B., Franco, G., Gabbert, F., Gamblin, B. W., Garry, M., Gentry, R., Gilbert, E. A., Greenberg, D. L., Halberstadt, J., Hall, L., Hancock, P. J. B., Hirsch, D., Holt, G., Jackson, J. C., Jong, J., Kehn, A., Koch, C., Kopietz, R., Körner, U., Kunar, M. A., Lai, C. K., Langton, S. R. H., Leite, F. P., Mammarella, N., Marsh, J. E., McConaughy, K. A., McCoy, S., McIntyre, A. H., Meissner, C. A., Michael, R. B., Mitchell, A. A., Mugayar-Baldocchi, M., Musselman, R., Ng, C., Nichols, A. L., Nunez, N. L., Palmer, M. A., Pappagianopoulos, J. E., Petro, M. S., Poirier, C. R., Portch, E., Rainsford, M., Rancourt, A., Romig, C., Rubínová, E., Sanson, M., Satchell, L., Sauer, J. D., Schweitzer, K., Shaheed, J., Skelton, F., Sullivan, G. A., Susa, K. J., Swanner, J. K., Thompson, W. B., Todaro, R., Ulatowska, J., Valentine, T., Verkoeijen, P. P. J. L., Vranka, M., Wade, K. A., Was, C. A., Weatherford, D., Wiseman, K., Zaksaitė, T., Zuj, D. V., Zwaan, R. A.

Protocol vetted by: Jonathan W. Schooler

Protocol edited by: Daniel J. Simons

Multilab direct replication of: Study 4 (modified) and Study 1 from Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22, 36–71.

Data and registered protocols: <https://osf.io/ybeur/>

Citation: Alogna, V. K., Attaya, M. K., Aucoin, P., Bahnik, Š., Birch, S., Birt, A. R., ... Zwaan, R. A. (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science*, 9, 556–578.

Target a single experiment

- Original study author provides materials and approves design
- 31 labs ran one replication study
- 22 of those ran a second version
- In total, >4000 participants involved

Challenges: registered replication

- In first report, all studies replicated the flaws the original design
 - Would have been easy to fix
 - Proposed fix was rejected
- In general, we learn only about this particular experimental effect

Registered research plans

- Who supervises?
- How do we interpret drop-out rate?
- What about being “scooped”?

Simpler solution? Increase sample sizes in our own studies

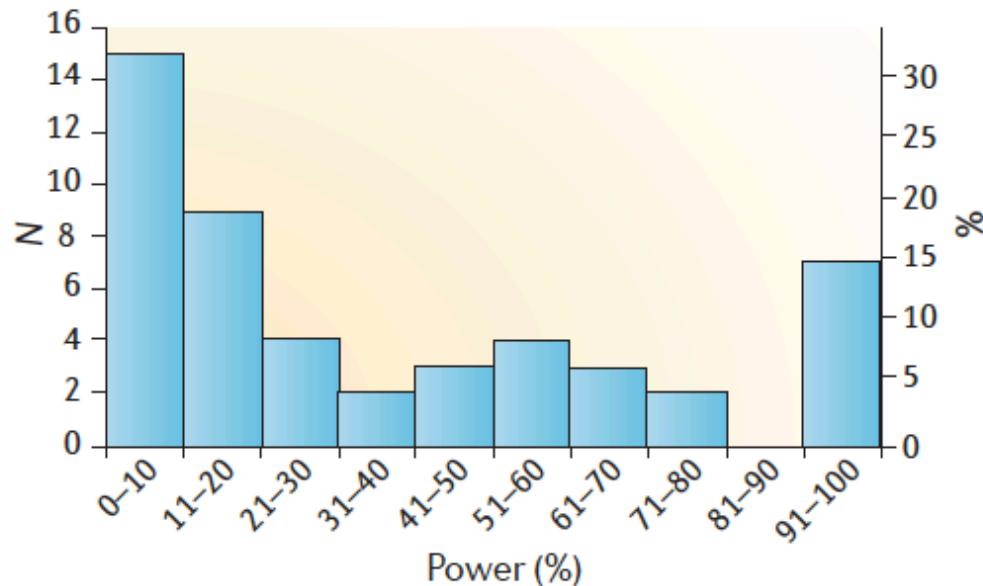


Figure 3 | Median power of studies included in neuroscience meta-analyses. The figure shows a histogram of median study power calculated for each of the $n = 49$ meta-analyses included in our analysis, with the number of meta-analyses (N) on the left axis and percent of meta-analyses (%) on the right axis. There is a clear