PROGRAMA DE DOCTORADO EN PSICOLOGÍA: ESCUELA DE VERANO
Ciencia abierta, replicabilidad y evidencia científica en la investigación en Psicología
Junio 2019

3er Plan Propio de Docencia

Facultad de Psicología
UNIVERSIDAD D SEVILLA

Replicability Crisis in Psychology: causes and solutions

Pandelis Perakakis
Departamento de Psicología
Universidad Loyola Andalucía
https://pandelisperakakis.info

@ppandelis

I. Definitions

II. Is there a crisis and how serious is it?

III. Methodological & statistical causes (p-values, questionable research practices, etc)

IV. Solutions (methodological/statistical education, preregistered studies, etc)

V. Systemic causes (research validation, evaluation and communication)

VI. Solutions (peer review, preprints, repositories, etc)

VII. Discussion

I. Definitions

REPRODUCIBILITY

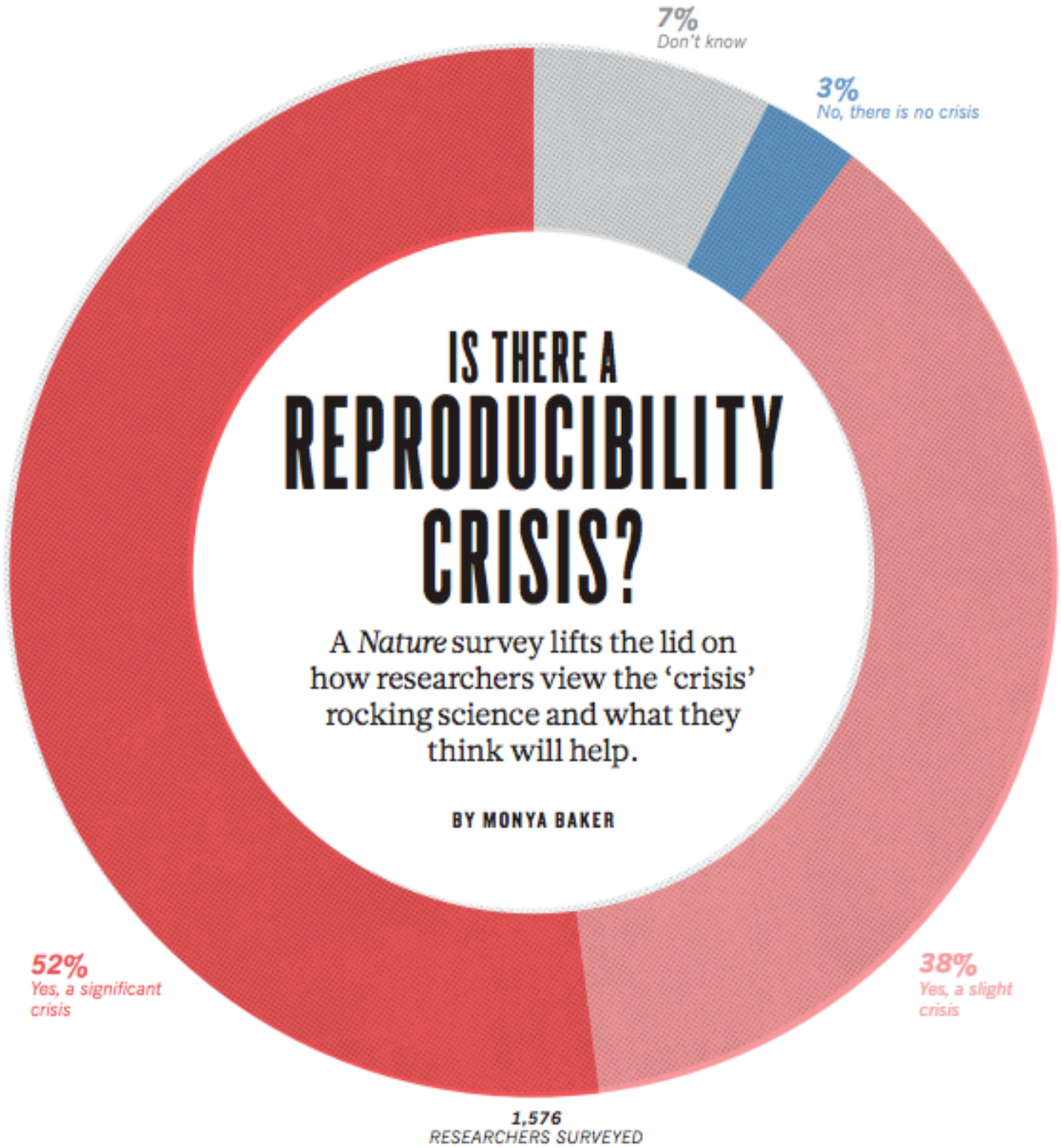
- Refers to the ability of a researcher to duplicate the results of a prior study using the same materials (incl. data) and procedures as were used by the original investigator

REPLICABILITY

- Refers to the ability of a researcher to duplicate the results of a prior study if the same procedures are followed but new data are collected (although 82% of replications in psychology are conceptual: Makel, Plucker & Hegarty, 2012)

Report of the Subcommittee on Replicability in Science of the SBE Advisory Committee to the National Science Foundation
13 May 2015 Presentation
at SBE AC Spring Meeting by K. Bollen

II. Is there a crisis and how serious is it?



IS THERE A REPRODUCIBILITY CRISIS?

A *Nature* survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

BY MONYA BAKER

WIKIPEDIA
The Free Encyclopedia

Article | Talk

Replication crisis

From Wikipedia, the free encyclopedia

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*

28 AUGUST 2015 • VOL 349 ISSUE 6251
SCIENCE sciencemag.org

nature human behaviour
LETTERS
https://doi.org/10.1038/s41562-018-0399-z

Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015

Colin F. Camerer[1,16], Anna Dreber[2,16], Felix Holzmeister[3,16], Teck-Hua Ho[4,16], Jürgen Huber[3,16], Magnus Johannesson[2,16], Michael Kirchler[3,5,16], Gideon Nave[6,16], Brian A. Nosek[7,8,16]*, Thomas Pfeiffer[9,16], Adam Altmejd[2], Nick Buttrick[7,8], Taizan Chan[10], Yiling Chen[11], Eskil Forsell[12], Anup Gampa[7,8], Emma Heikensten[2], Lily Hummer[8], Taisuke Imai[13], Siri Isaksson[2], Dylan Manfredi[6], Julia Rose[3], Eric-Jan Wagenmakers[14] and Hang Wu[15]

Marcia McNutt is Editor-in-Chief of *Science*.

EDITORIAL

SCIENCE VOL 343 17 JANUARY 2014

Reproducibility

SCIENCE ADVANCES ON A FOUNDATION OF TRUSTED DISCOVERIES. REPRODUCING AN EXPERIMENT is one important approach that scientists use to gain confidence in their conclusions. Recently, the scientific community was shaken by reports that a troubling proportion of

MENU

nature
International journal of science

Special | 18 October 2018

Challenges in irreproducible research

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

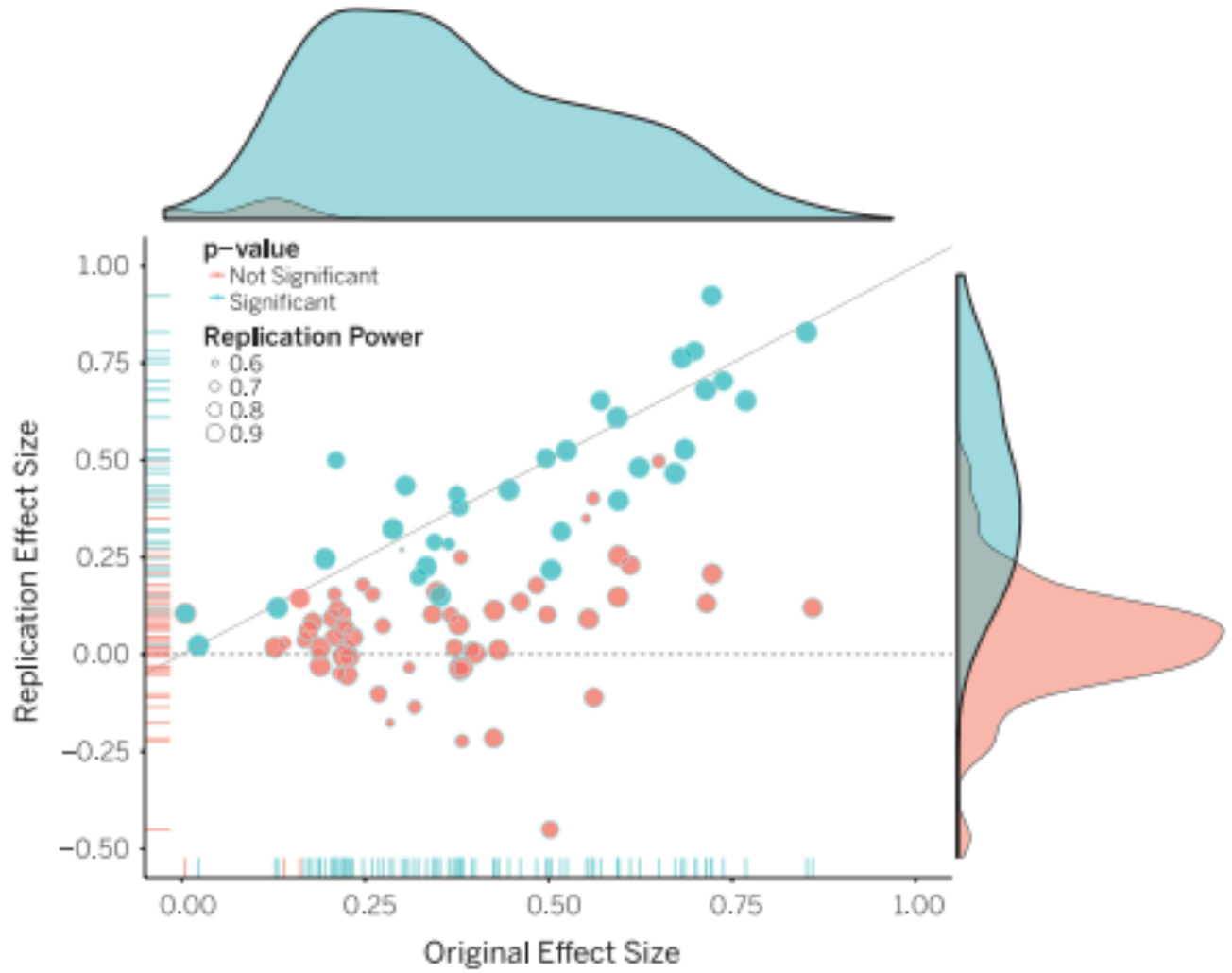
Open Science Collaboration[*,†]

[*]*All authors with their affiliations appear at the end of this paper.*

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

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

Thirty-six percent of replications had statistically significant results; 47% of original effect sizes were in the 95% confidence interval of the replication effect size; 39% of effects were subjectively rated to have replicated the original result; and if no bias in original results is assumed, combining original and replication results left 68% with statistically significant effects.



Original study effect size versus replication effect size (correlation coefficients). Diagonal line represents replication effect size equal to original effect size. Dotted line represents replication effect size of 0. Points below the dotted line were effects in the opposite direction of the original. Density plots are separated by significant (blue) and nonsignificant (red) effects.

nature human behaviour

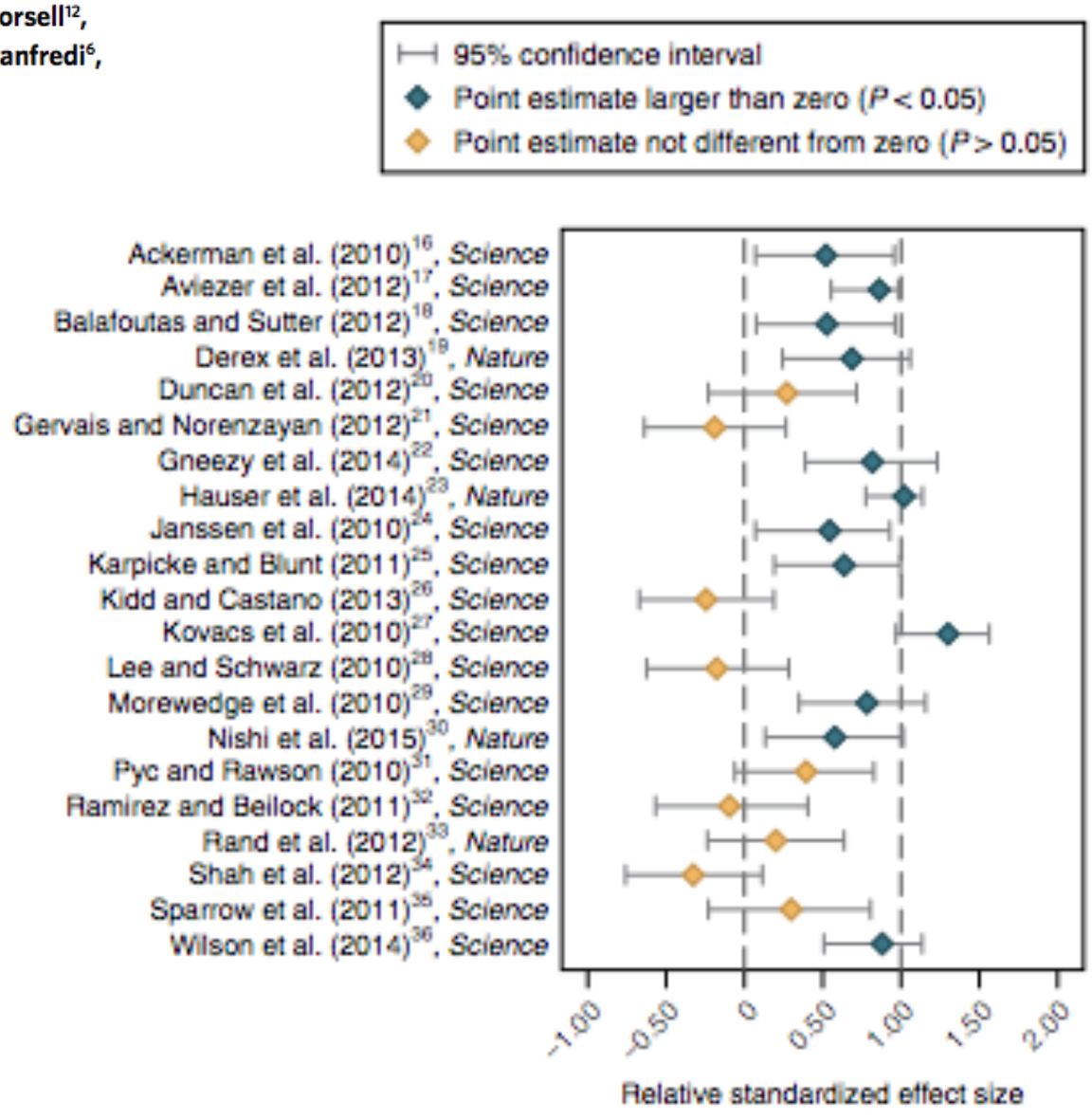
LETTERS

https://doi.org/10.1038/s41562-018-0399-z

Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015

Colin F. Camerer[1,16], Anna Dreber[2,16], Felix Holzmeister[3,16], Teck-Hua Ho[4,16], Jürgen Huber[3,16], Magnus Johannesson[2,16], Michael Kirchler[3,5,16], Gideon Nave[6,16], Brian A. Nosek[7,8,16*], Thomas Pfeiffer[9,16], Adam Altmejd[2], Nick Buttrick[7,8], Taizan Chan[10], Yiling Chen[11], Eskil Forsell[12], Anup Gampa[7,8], Emma Heikensten[2], Lily Hummer[8], Taisuke Imai[13], Siri Isaksson[2], Dylan Manfredi[6], Julia Rose[3], Eric-Jan Wagenmakers[14] and Hang Wu[15]

We find a significant effect in the same direction as the original study for 13 (62%) studies, and the effect size of the replications is on average about 50% of the original effect size.



Study	N	Replication rate
Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* 349, aac4716 (2015).	100	36 %
Camerer, C. F. et al. Evaluating the Replicability of Social Science Experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior* 2 (9): 637–44 (2018).	21	62 %
Camerer, C. F. et al. Evaluating replicability of laboratory experiments in economics. *Science* 351, 1433–1436 (2016).	18	61 %
Klein, R. A. et al. Investigating variation in replicability: a 'many labs' replication project. *Soc. Psychol*. 45, 142–152 (2014).	13	77 %
Klein, R. A. et al. Many Labs 2: Investigating Variation in Replicability Across Samples and Settings. *Advances in Methods and Practices in Psychological Science* 1, 443–490 (2018).	28	50 %
Ebersole, C. R. et al. Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *J. Exp. Soc. Psychol*. 67, 68–82 (2016).	10	30 %

The Washington Post
Democracy Dies in Darkness

Science

No, science's reproducibility problem is not limited to psychology

III. Methodological & statistical causes (p-values, questionable research practices, etc)

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

It can be proven that most claimed research findings are false.

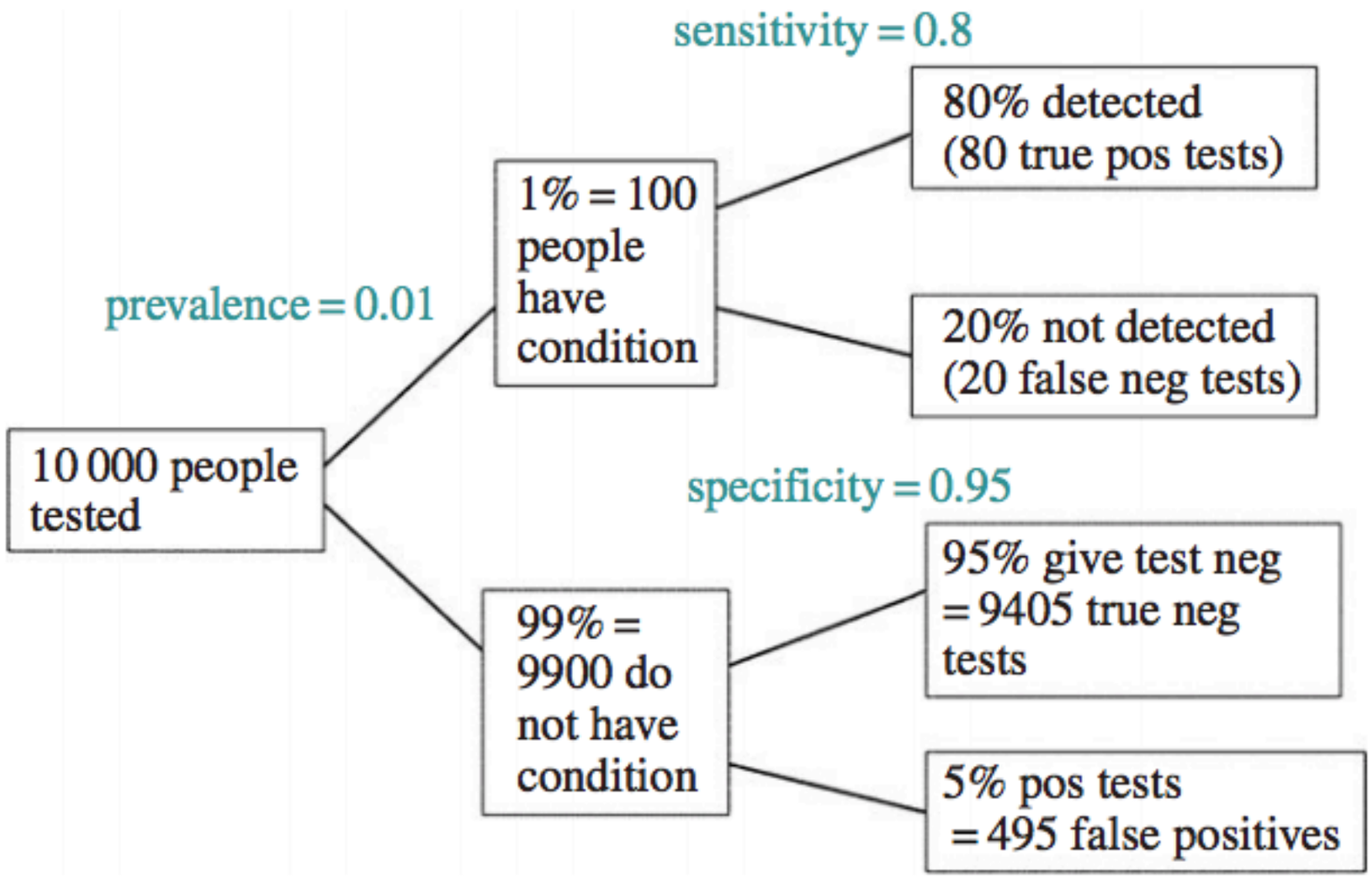
p-value

- The probability of observing a value for the difference equal to, or greater than, that actually observed, given that there is really no effect (H_o is true)

False Discovery Rate

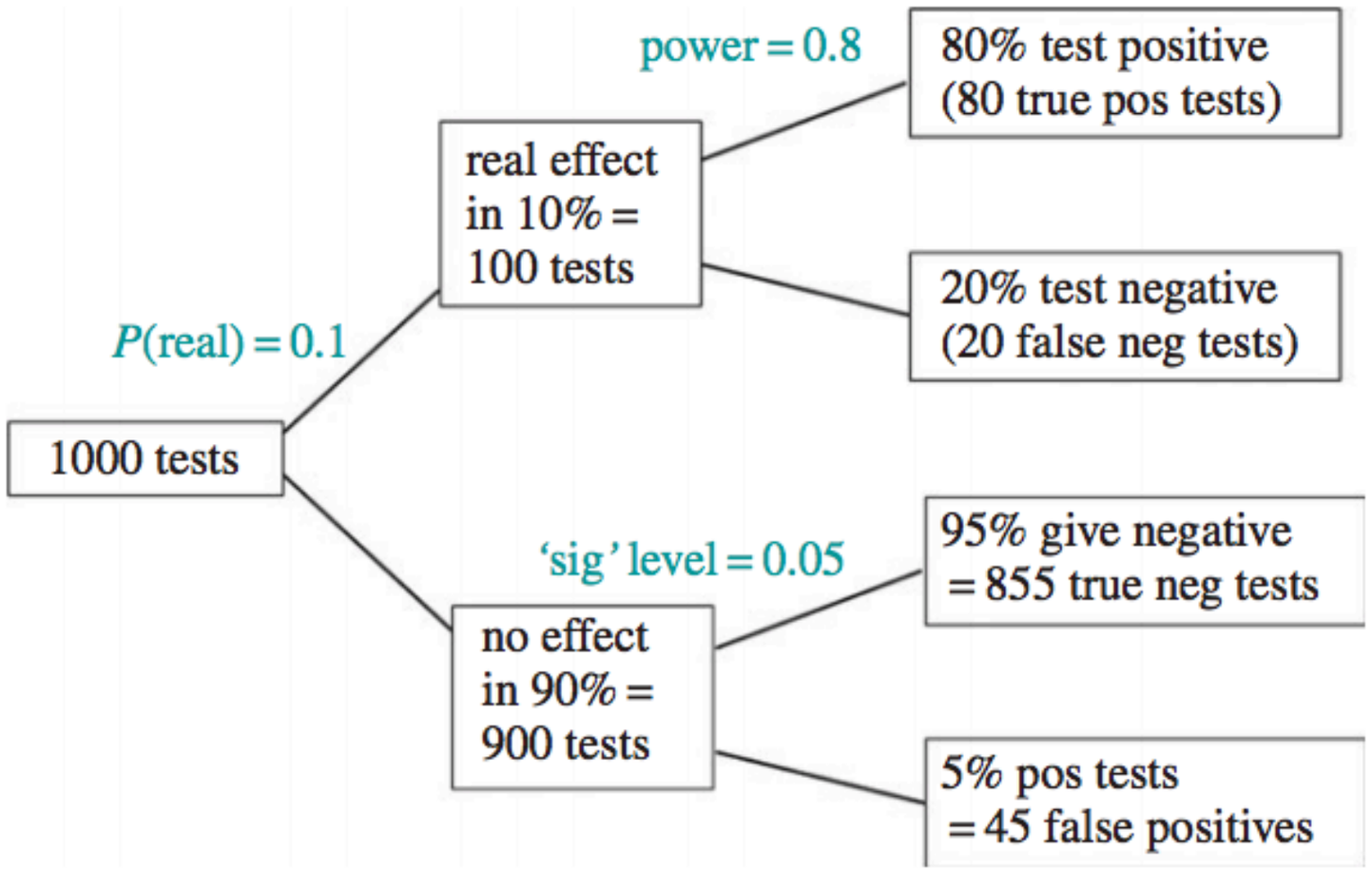
- The fraction of all positive tests that all false positives, i.e., there is no real effect and your observation is due to chance

Colquhoun, D. An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci* 1, 140216 (2014).



False discovery rate = 495/(495 + 80) = 86%

Colquhoun, D. An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci* 1, 140216 (2014).

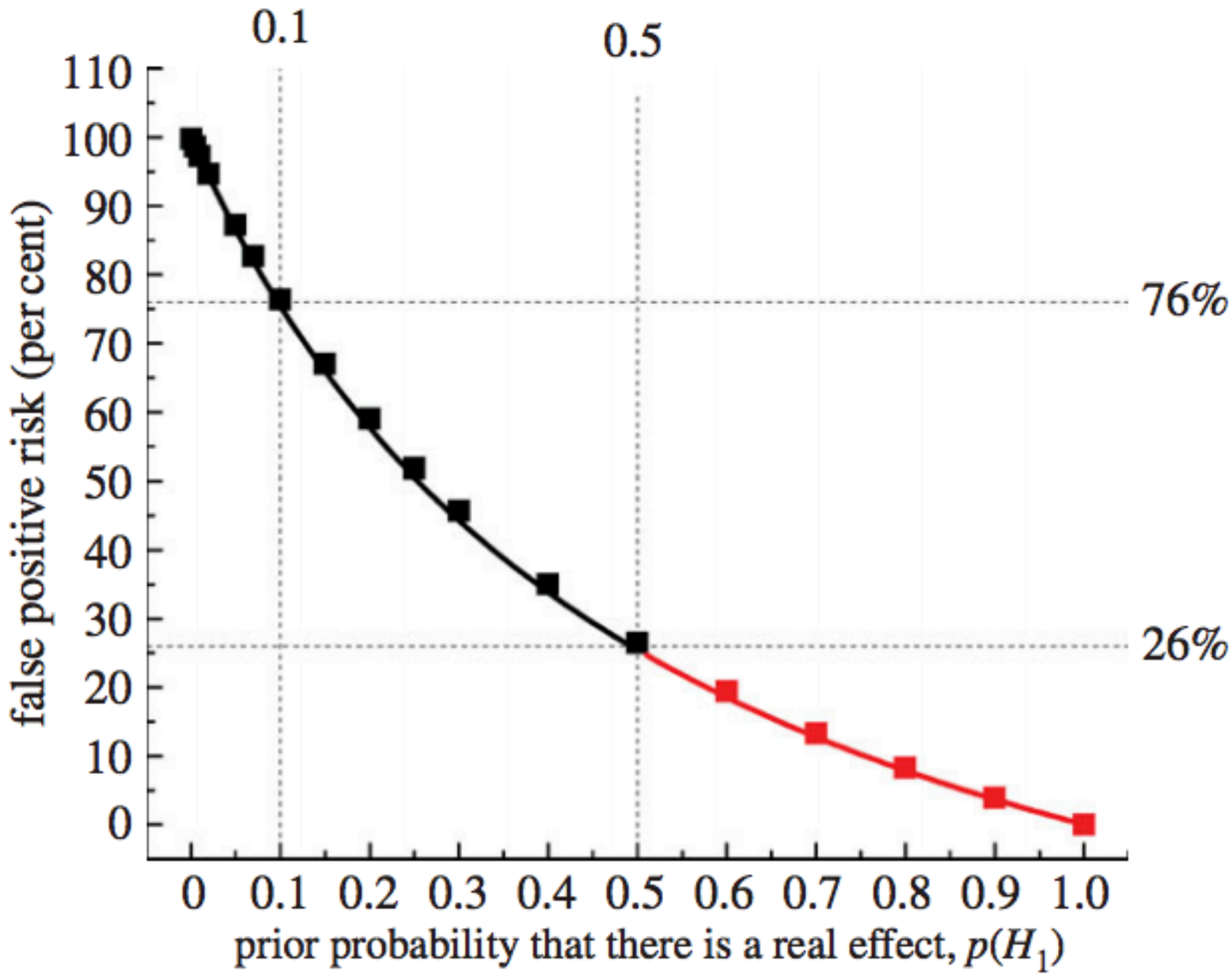


False discovery rate = 45/(45 + 80) = 36%
Power = 0.2 | FDR = 70%
P(real) = 0.5 | Power = 0.8 | FDR = 0.6%
P(real) = 0.5 | Power = 0.2 | FDR = 20%
P(real) = 0.5 | Power = 0.8 | *p*(equals) | FDR = 26%

We optimistically estimate the median statistical power of studies in the neuroscience field to be between about 8% and about 31%.

Button KS, et al. 2013 Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci.14, 365–376.

Colquhoun, D. An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci* 1, 140216 (2014).



False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Psychological Science
22(11) 1359–1366

DOI: 10.1177/0956797611417632
http://pss.sagepub.com
SAGE

Joseph P. Simmons[1], Leif D. Nelson[2], and Uri Simonsohn[1]
[1]The Wharton School, University of Pennsylvania, and [2]Haas School of Business, University of California, Berkeley

Table 1. Likelihood of Obtaining a False-Positive Result

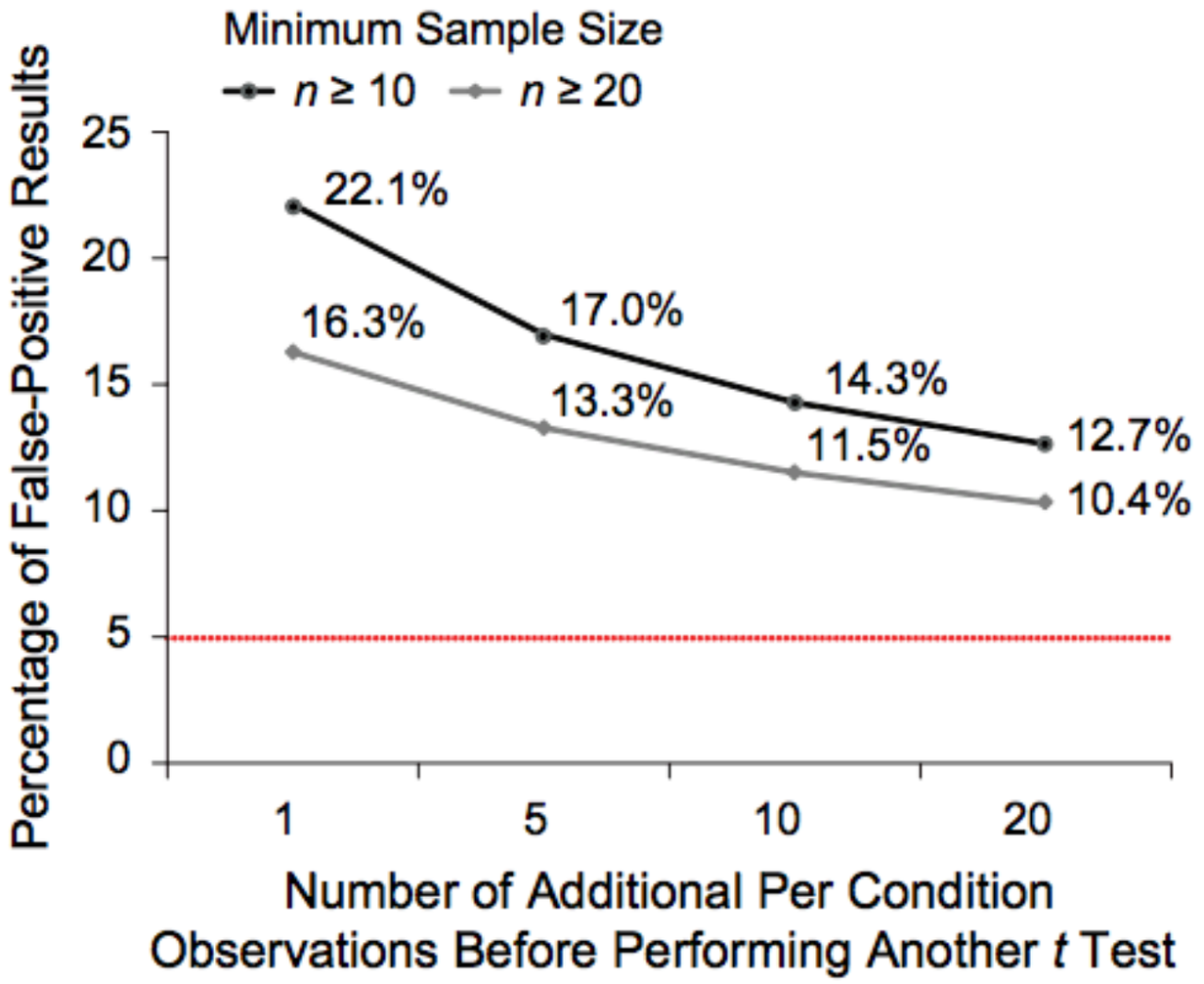
	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Psychological Science
22(11) 1359–1366

DOI: 10.1177/0956797611417632
http://pss.sagepub.com
SAGE

Joseph P. Simmons[1], Leif D. Nelson[2], and Uri Simonsohn[1]
[1]The Wharton School, University of Pennsylvania, and [2]Haas School of Business, University of California, Berkeley

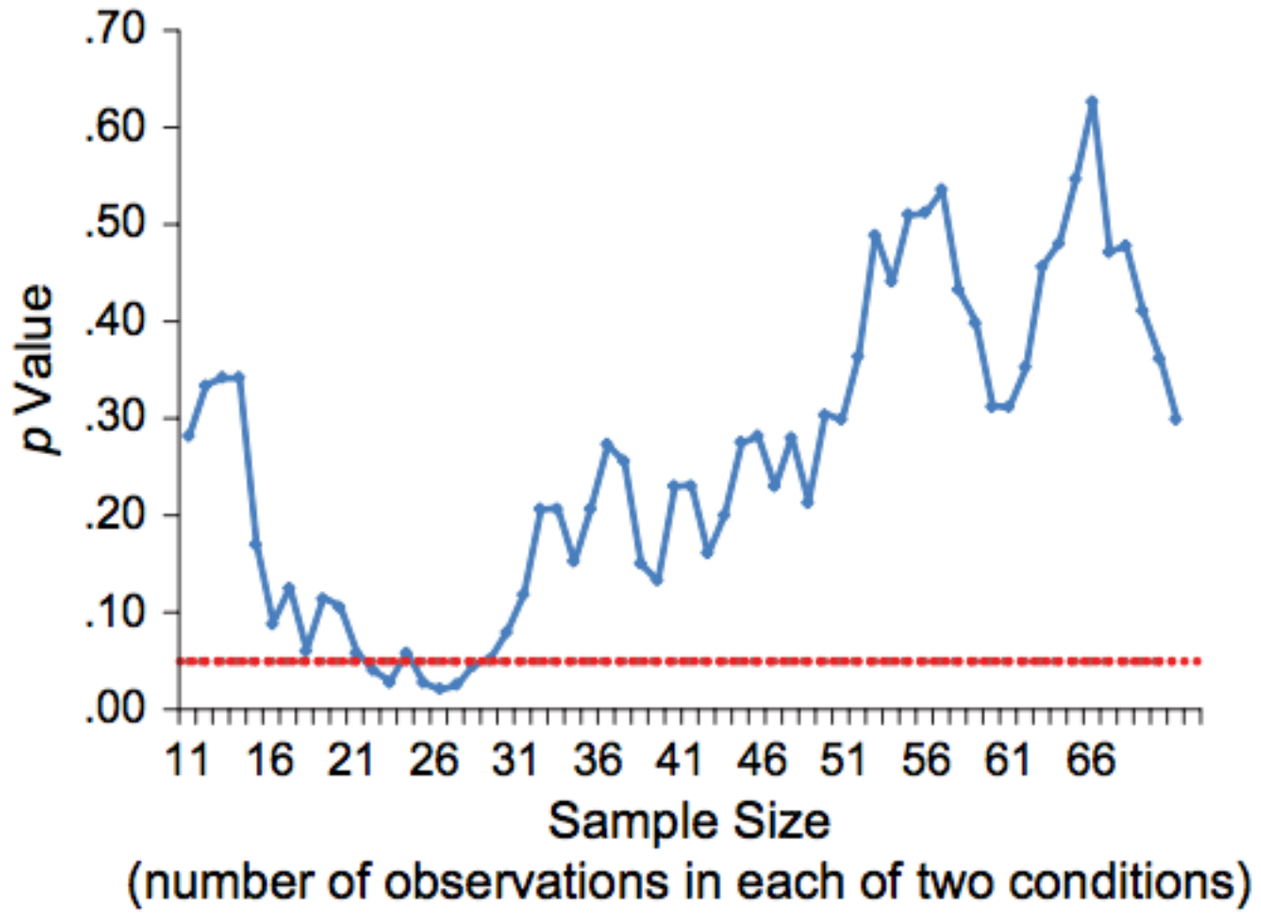


False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Psychological Science
22(11) 1359–1366

DOI: 10.1177/0956797611417632
http://pss.sagepub.com
SAGE

Joseph P. Simmons[1], Leif D. Nelson[2], and Uri Simonsohn[1]
[1]The Wharton School, University of Pennsylvania, and [2]Haas School of Business, University of California, Berkeley



If several thousand potential analytical pipelines can be applied to high-dimensional data, the generation of false-positive findings is highly likely. For example, applying almost 7,000 analytical pipelines to a single fMRI dataset resulted in over 90% of brain voxels showing significant activation in at least one analysis.

Carp, J. On the plurality of (methodological) worlds: estimating the analytic flexibility of FMRI experiments. Front. Neurosci. 6, 149 (2012).

Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction

Craig M. Bennett[1], Abigail A. Baird[2], Michael B. Miller[1], and George L. Wolford[3]

[1] Psychology Department, University of California Santa Barbara, Santa Barbara, CA; [2] Department of Psychology, Vassar College, Poughkeepsie, NY; [3] Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH

INTRODUCTION

With the extreme dimensionality of functional neuroimaging data comes extreme risk for false positives. Across the 130,000 voxels in a typical fMRI volume the probability of a false positive is almost certain. Correction for multiple comparisons should be completed with these datasets, but is often ignored by investigators. To illustrate the magnitude of the problem we carried out a real experiment that demonstrates the danger of not correcting for chance properly.

METHODS

Subject. One mature Atlantic Salmon (Salmo salar) participated in the fMRI study. The salmon was approximately 18 inches long, weighed 3.8 lbs, and was not alive at the time of scanning.

Task. The task administered to the salmon involved completing an open-ended mentalizing task. The salmon was shown a series of photographs depicting human individuals in social situations with a specified emotional valence. The salmon was asked to determine what emotion the individual in the photo must have been experiencing.

Design. Stimuli were presented in a block design with each photo presented for 10 seconds followed by 12 seconds of rest. A total of 15 photos were displayed. Total scan time was 5.5 minutes.

Preprocessing. Image processing was completed using SPM2. Preprocessing steps for the functional imaging data included a 6-parameter rigid-body affine realignment of the fMRI timeseries, coregistration of the data to a T_1-weighted anatomical image, and 8 mm full-width at half-maximum (FWHM) Gaussian smoothing.

Analysis. Voxelwise statistics on the salmon data were calculated through an ordinary least-squares estimation of the general linear model (GLM). Predictors of the hemodynamic response were modeled by a boxcar function convolved with a canonical hemodynamic response. A temporal high pass filter of 128 seconds was include to account for low frequency drift. No autocorrelation correction was applied.

Voxel Selection. Two methods were used for the correction of multiple comparisons in the fMRI results. The first method controlled the overall false discovery rate (FDR) and was based on a method defined by Benjamini and Hochberg (1995). The second method controlled the overall familywise error rate (FWER) through the use of Gaussian random field theory. This was done using algorithms originally devised by Friston et al. (1994).

DISCUSSION

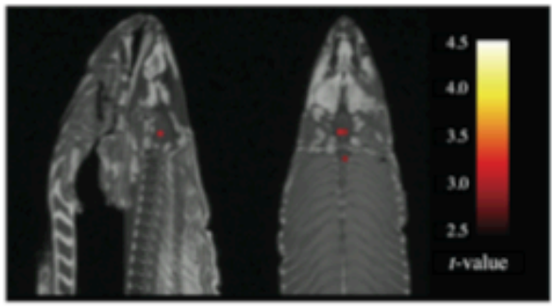
Can we conclude from this data that the salmon is engaging in the perspective-taking task? Certainly not. What we can determine is that random noise in the EPI timeseries may yield spurious results if multiple comparisons are not controlled for. Adaptive methods for controlling the FDR and FWER are excellent options and are widely available in all major fMRI analysis packages. We argue that relying on standard statistical thresholds ($p < 0.001$) and low minimum cluster sizes ($k > 8$) is an ineffective control for multiple comparisons. We further argue that the vast majority of fMRI studies should be utilizing multiple comparisons correction as standard practice in the computation of their statistics.

REFERENCES

Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57:289-300.

Friston K.J, Worsley K.J, Frackowiak RSJ, Mazziotta JC, and Evans AC. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:214-220.

GLM RESULTS

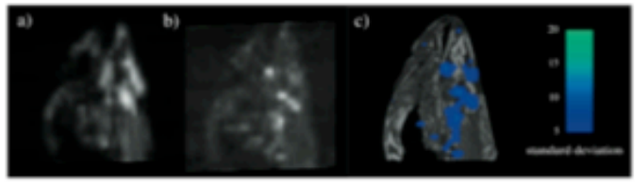


A *t*-contrast was used to test for regions with significant BOLD signal change during the photo condition compared to rest. The parameters for this comparison were $t(131) > 3.15$, p(uncorrected) < 0.001, 3 voxel extent threshold.

Several active voxels were discovered in a cluster located within the salmon's brain cavity (Figure 1, see above). The size of this cluster was 81 mm^3 with a cluster-level significance of $p = 0.001$. Due to the coarse resolution of the echo-planar image acquisition and the relatively small size of the salmon brain further discrimination between brain regions could not be completed. Out of a search volume of 8064 voxels a total of 16 voxels were significant.

Identical *t*-contrasts controlling the false discovery rate (FDR) and familywise error rate (FWER) were completed. These contrasts indicated no active voxels, even at relaxed statistical thresholds ($p = 0.25$).

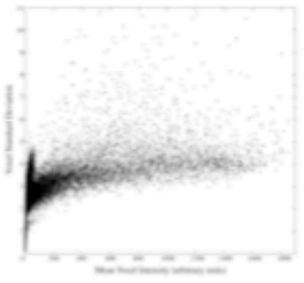
VOXELWISE VARIABILITY



To examine the spatial configuration of false positives we completed a variability analysis of the fMRI timeseries. On a voxel-by-voxel basis we calculated the standard deviation of signal values across all 140 volumes.

We observed clustering of highly variable voxels into groups near areas of high voxel signal intensity. Figure 2a shows the mean EPI image for all 140 image volumes. Figure 2b shows the standard deviation values of each voxel. Figure 2c shows thresholded standard deviation values overlaid onto a high-resolution T_1-weighted image.

To investigate this effect in greater detail we conducted a Pearson correlation to examine the relationship between the signal in a voxel and its variability. There was a significant positive correlation between the mean voxel value and its variability over time ($r = 0.54$, $p < 0.001$). A scatterplot of mean voxel signal intensity against voxel standard deviation is presented to the right.



Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction

Craig M. Bennett[1], Abigail A. Baird[2], Michael B. Miller[1], and George L. Wolford[3]

[1] Psychology Department, University of California Santa Barbara, Santa Barbara, CA; [2] Department of Psychology, Vassar College, Poughkeepsie, NY; [3] Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH

INTRODUCTION

With the extreme dimensionality of functional neuroimaging data comes

METHODS

Subject. One mature Atlantic Salmon (Salmo salar) participated in the fMRI study. The salmon was approximately 18 inches long, weighed 3.8 lbs, and was not alive at the time of scanning.

Task. The task administered to the salmon involved completing an open-ended mentalizing task. The salmon was shown a series of photographs depicting human individuals in social situations with a specified emotional valence. The salmon was asked to determine what emotion the individual in the photo must have been experiencing.

Design. Stimuli were presented in a block design with each photo presented for 10 seconds followed by 12 seconds of rest. A total of 15 photos were displayed. Total scan time was 5.5 minutes.

in the fMRI results. The first method controlled the overall false discovery rate (FDR) and was based on a method defined by Benjamini and Hochberg (1995). The second method controlled the overall familywise error rate (FWER) through the use of Gaussian random field theory. This was done using algorithms originally devised by Friston et al. (1994).

DISCUSSION

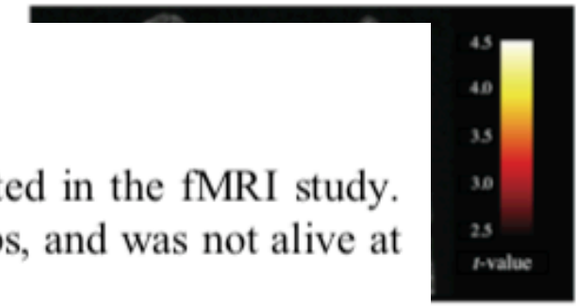
Can we conclude from this data that the salmon is engaging in the perspective-taking task? Certainly not. What we can determine is that random noise in the EPI timeseries may yield spurious results if multiple comparisons are not controlled for. Adaptive methods for controlling the FDR and FWER are excellent options and are widely available in all major fMRI analysis packages. We argue that relying on standard statistical thresholds ($p < 0.001$) and low minimum cluster sizes ($k > 8$) is an ineffective control for multiple comparisons. We further argue that the vast majority of fMRI studies should be utilizing multiple comparisons correction as standard practice in the computation of their statistics.

REFERENCES

Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57:289-300.

Friston K.J, Worsley K.J, Frackowiak RSJ, Mazziotta JC, and Evans AC. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:214-220.

GLM RESULTS

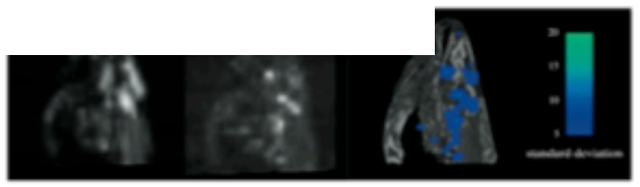


...nt BOLD signal change
...he parameters for this
... 0.001, 3 voxel extent

...ated within the salmon's
...aster was 81 mm³ with a
...coarse resolution of the
...nall size of the salmon
...ould not be completed.
...xels were significant.

...e (FDR) and familywise
...sts indicated no active
...

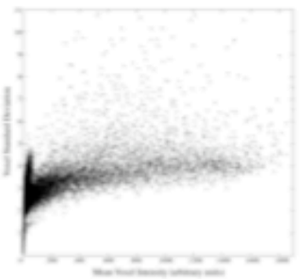
...ITY



To examine the spatial configuration of false positives we completed a variability analysis of the fMRI timeseries. On a voxel-by-voxel basis we calculated the standard deviation of signal values across all 140 volumes.

We observed clustering of highly variable voxels into groups near areas of high voxel signal intensity. Figure 2a shows the mean EPI image for all 140 image volumes. Figure 2b shows the standard deviation values of each voxel. Figure 2c shows thresholded standard deviation values overlaid onto a high-resolution T_1-weighted image.

To investigate this effect in greater detail we conducted a Pearson correlation to examine the relationship between the signal in a voxel and its variability. There was a significant positive correlation between the mean voxel value and its variability over time ($r = 0.54$, $p < 0.001$). A scatterplot of mean voxel signal intensity against voxel standard deviation is presented to the right.



Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction

Craig M. Bennett[1], Abigail A. Baird[2], Michael B. Miller[1], and George L. Wolford[3]

[1] Psychology Department, University of California Santa Barbara, Santa Barbara, CA; [2] Department of Psychology, Vassar College, Poughkeepsie, NY; [3] Department of Psychological & Brai

INTRODUCTION

With the extreme dimensionality of functional ne
extreme risk for false positives. Across the 130,000
volume the probability of a false positive is almost
multiple comparisons should be completed with the
ignored by investigators. To illustrate the magnit
carried out a real experiment that demonstrates the
for chance properly.

METHODS

Subject. One mature Atlantic Salmon (Salmo salar) parti
The salmon was approximately 18 inches long, weighed 3
the time of scanning.

Task. The task administered to the salmon involved
mentalizing task. The salmon was shown a series of ph
individuals in social situations with a specified emotional
asked to determine what emotion the individual in t
experiencing.

Design. Stimuli were presented in a block design with e
seconds followed by 12 seconds of rest. A total of 15 ph
scan time was 5.5 minutes.

Preprocessing. Image processing was completed using S
for the functional imaging data included a 6-parameter rig
of the fMRI timeseries, coregistration of the data to a T_1-
and 8 mm full-width at half-maximum (FWHM) Gaussian

Analysis. Voxelwise statistics on the salmon data w
ordinary least-squares estimation of the general linear m
the hemodynamic response were modeled by a boxcar
canonical hemodynamic response. A temporal high pass
include to account for low frequency drift. No auto
applied.

Voxel Selection. Two methods were used for the correcti
in the fMRI results. The first method controlled the
(FDR) and was based on a method defined by Benjamini
second method controlled the overall familywise error ra
of Gaussian random field theory. This was done using al
by Friston et al. (1994).

DISCUSSION

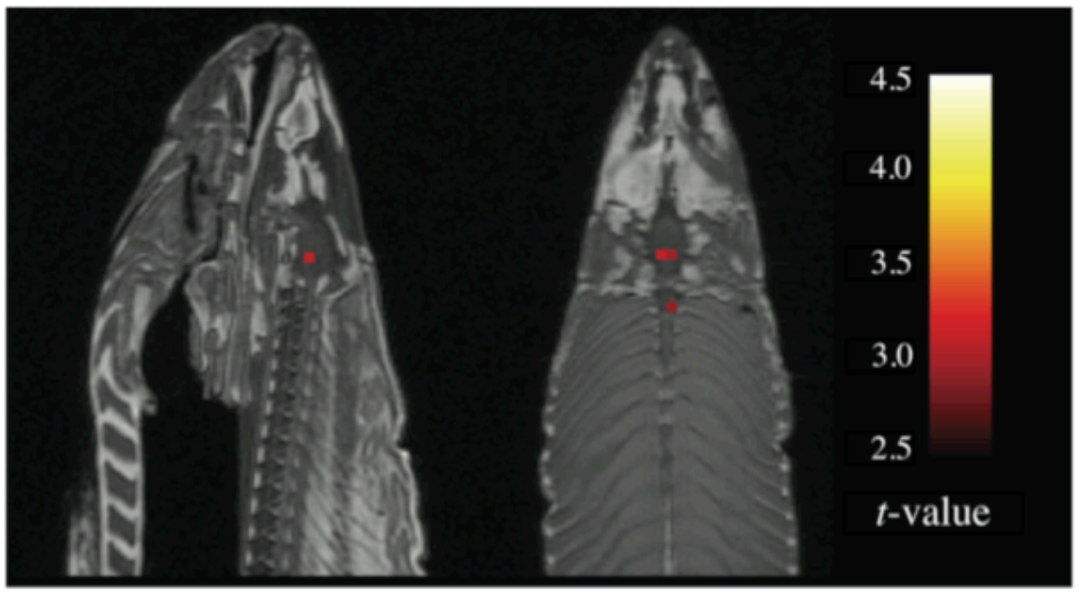
Can we conclude from this data that the salm
perspective-taking task? Certainly not. What we can
noise in the EPI timeseries may yield spurious results
are not controlled for. Adaptive methods for controll
are excellent options and are widely available in
packages. We argue that relying on standard statistic
and low minimum cluster sizes ($k > 8$) is an ineffec
comparisons. We further argue that the vast majorit
be utilizing multiple comparisons correction as s
computation of their statistics.

REFERENCES

Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57:289-300.

Friston K.J, Worsley K.J, Frackowiak RSJ, Mazziotta JC, and Evans AC. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:214-220.

GLM RESULTS

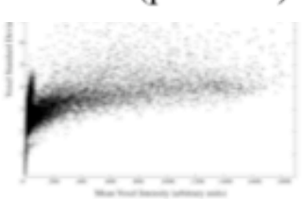


A *t*-contrast was used to test for regions with significant BOLD signal change during the photo condition compared to rest. The parameters for this comparison were $t(131) > 3.15$, p(uncorrected) < 0.001, 3 voxel extent threshold.

Several active voxels were discovered in a cluster located within the salmon's brain cavity (Figure 1, see above). The size of this cluster was 81 mm^3 with a cluster-level significance of $p = 0.001$. Due to the coarse resolution of the echo-planar image acquisition and the relatively small size of the salmon brain further discrimination between brain regions could not be completed. Out of a search volume of 8064 voxels a total of 16 voxels were significant.

Identical *t*-contrasts controlling the false discovery rate (FDR) and familywise error rate (FWER) were completed. These contrasts indicated no active voxels, even at relaxed statistical thresholds ($p = 0.25$).

variability. There was a significant positive correlation between the mean voxel value and its variability over time ($r = 0.54$, $p < 0.001$). A scatterplot of mean voxel signal intensity against voxel standard deviation is presented to the right.



Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction

Craig M. Bennett[1], Abigail A. Baird[2], Michael B. Miller[1], and George L. Wolford[3]

[1] Psychology Department, University of California Santa Barbara, Santa Barbara, CA; [2] Department of Psychology, Vassar College, Poughkeepsie, NY; [3] Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH

INTRODUCTION

With the extreme dimensionality of functional neuroimaging data comes extreme risk for false positives. Across the 130,000 voxels in a typical fMRI volume the probability of a false positive is almost certain. Correction for multiple comparisons should be completed with these datasets, but is often ignored by investigators. To illustrate the magnitude of the problem we carried out a real experiment that demonstrates the danger of not correcting for chance properly.

METHODS

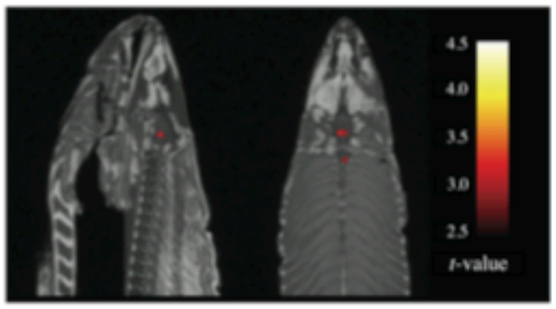
Subject. One mature Atlantic Salmon (Salmo salar) participated in the fMRI study. The salmon was approximately 18 inches long, weighed 3.8 lbs, and was not alive at the time of scanning.

Task. The task administered to the salmon involved completing an open-ended mentalizing task. The salmon was shown a series of photographs depicting human individuals in social situations with a specified emotional valence. The salmon was asked to determine what emotion the individual in the photo must have been experiencing.

Design. Stimuli were presented in a block design with each photo presented for 10 seconds followed by 12 seconds of rest. A total of 15 photos were displayed. Total scan time was 5.5 minutes.

Preprocessing. Image processing was completed using SPM2. Preprocessing steps for the functional imaging data included a 6-parameter rigid body affine realignment

GLM RESULTS



A *t*-contrast was used to test for regions with significant BOLD signal change during the photo condition compared to rest. The parameters for this comparison were $t(131) > 3.15$, p(uncorrected) < 0.001, 3 voxel extent threshold.

Several active voxels were discovered in a cluster located within the salmon's brain cavity (Figure 1, see above). The size of this cluster was 81 mm^3 with a cluster-level significance of $p = 0.001$. Due to the coarse resolution of the echo-planar image acquisition and the relatively small size of the salmon brain further discrimination between brain regions could not be completed. Out of a search volume of 8064 voxels a total of 16 voxels were significant.

DISCUSSION

Can we conclude from this data that the salmon is engaging in the perspective-taking task? Certainly not. What we can determine is that random noise in the EPI timeseries may yield spurious results if multiple comparisons are not controlled for. Adaptive methods for controlling the FDR and FWER are excellent options and are widely available in all major fMRI analysis packages. We argue that relying on standard statistical thresholds ($p < 0.001$) and low minimum cluster sizes ($k > 8$) is an ineffective control for multiple comparisons. We further argue that the vast majority of fMRI studies should be utilizing multiple comparisons correction as standard practice in the computation of their statistics.

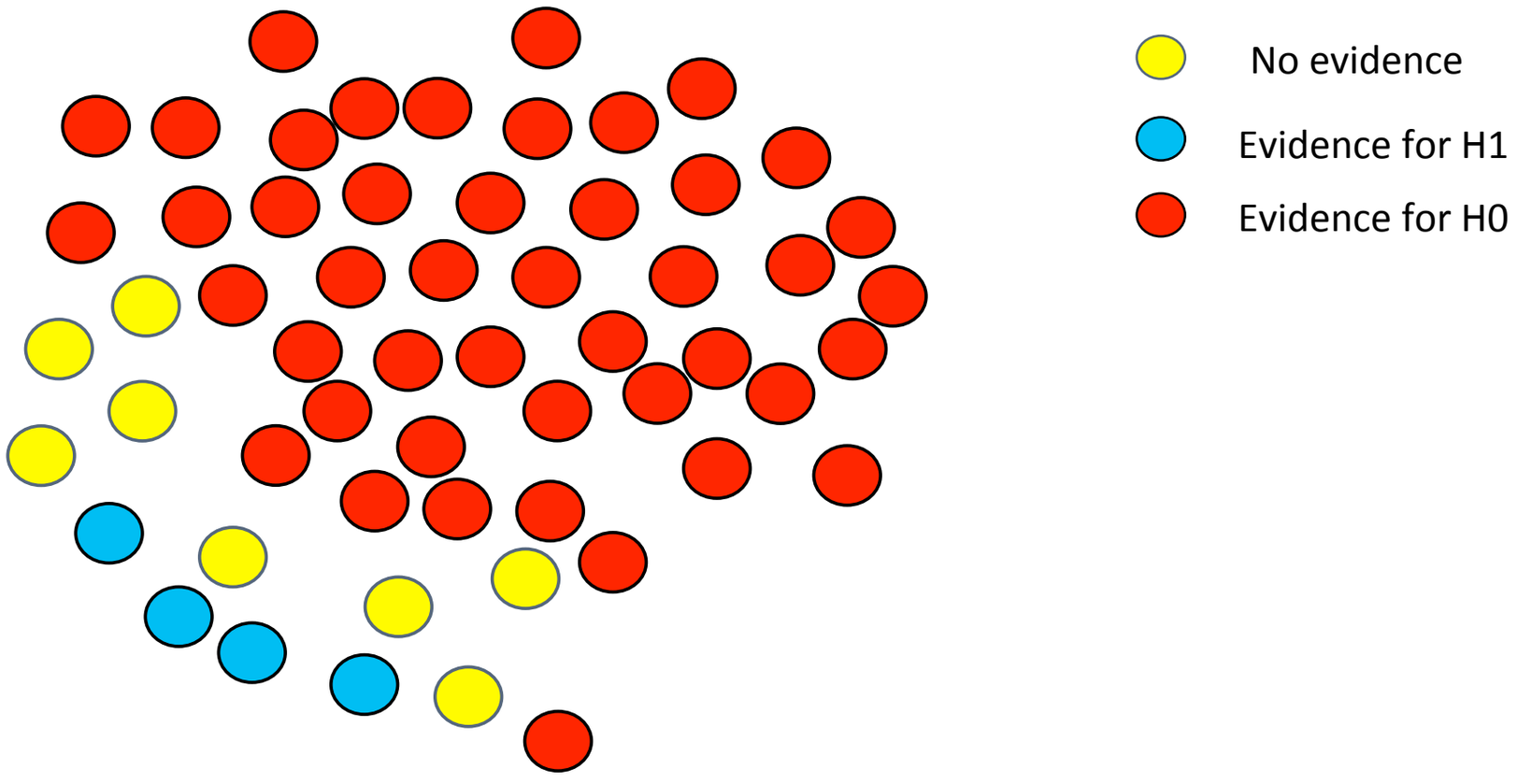
Friston K.J, Worsley K.J, Frackowiak RSJ, Mazziotta JC, and Evans AC. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*. 1:214-220.

deviation is presented to the right.

How is variable 1 coded? How is variable 2 coded? Exclude if …? How ….

5 X 3 X 4 X ….

= the **multiverse**

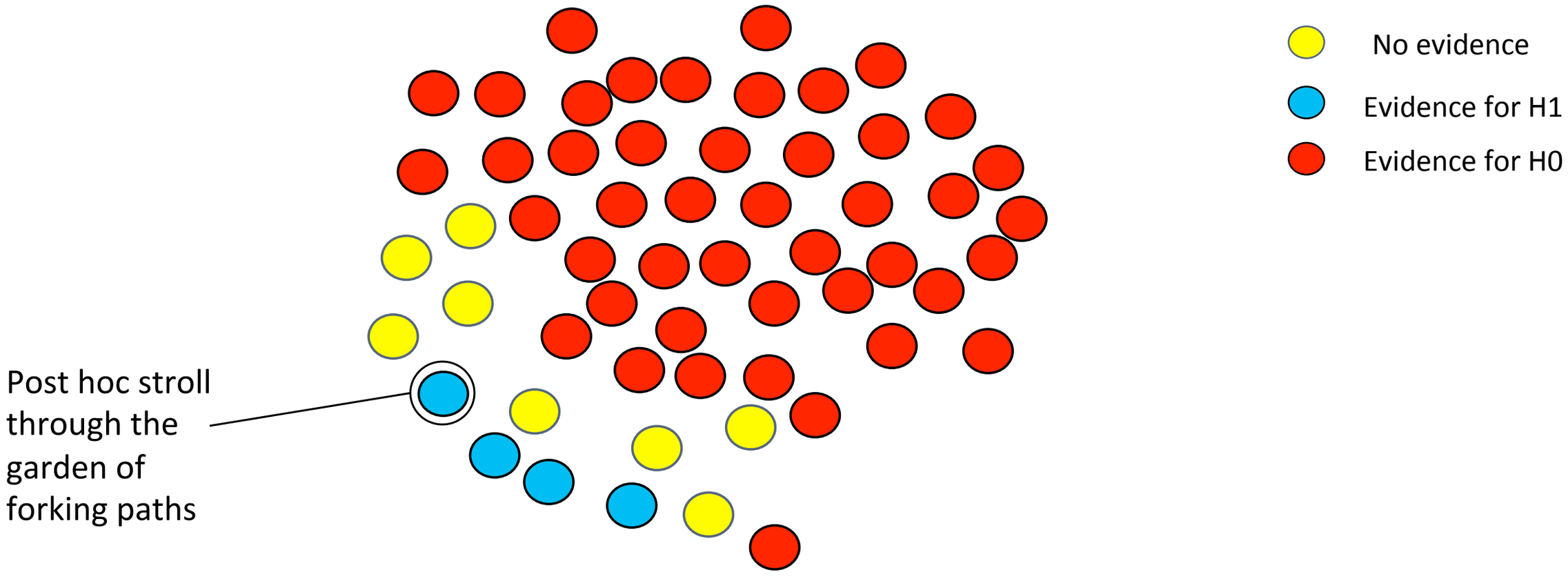


Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing Transparency Through a Multiverse Analysis. *Perspect. Psychol. Sci.* 11, 702–712 (2016).

How is variable 1 coded? How is variable 2 coded? Exclude if …? How ….

5 X 3 X 4 X ….

= the **multiverse**

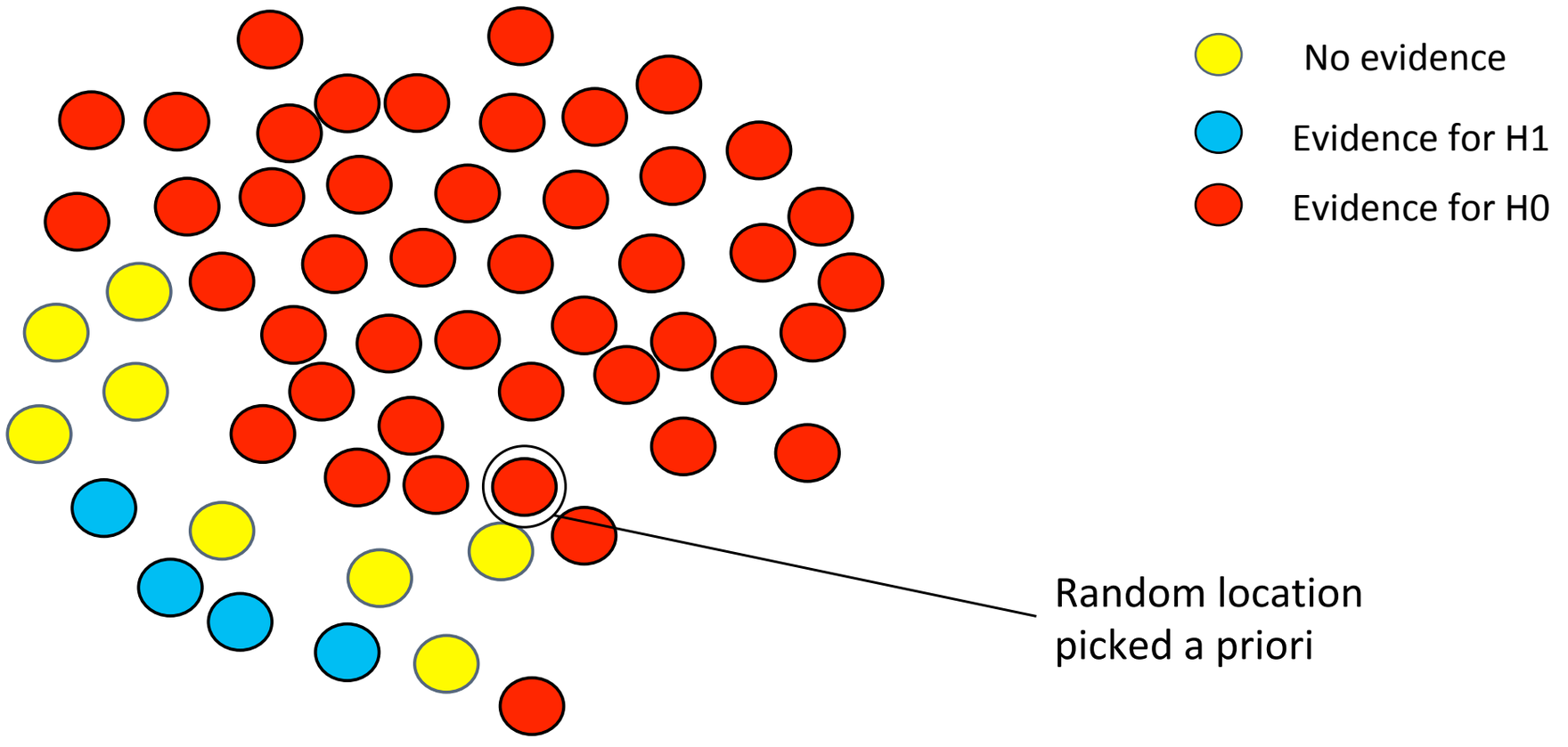


Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing Transparency Through a Multiverse Analysis. *Perspect. Psychol. Sci.* 11, 702–712 (2016).

How is variable 1 coded? How is variable 2 coded? Exclude if …? How ….

5 X 3 X 4 X ….

= the **multiverse**



Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing Transparency Through a Multiverse Analysis. *Perspect. Psychol. Sci.* 11, 702–712 (2016).

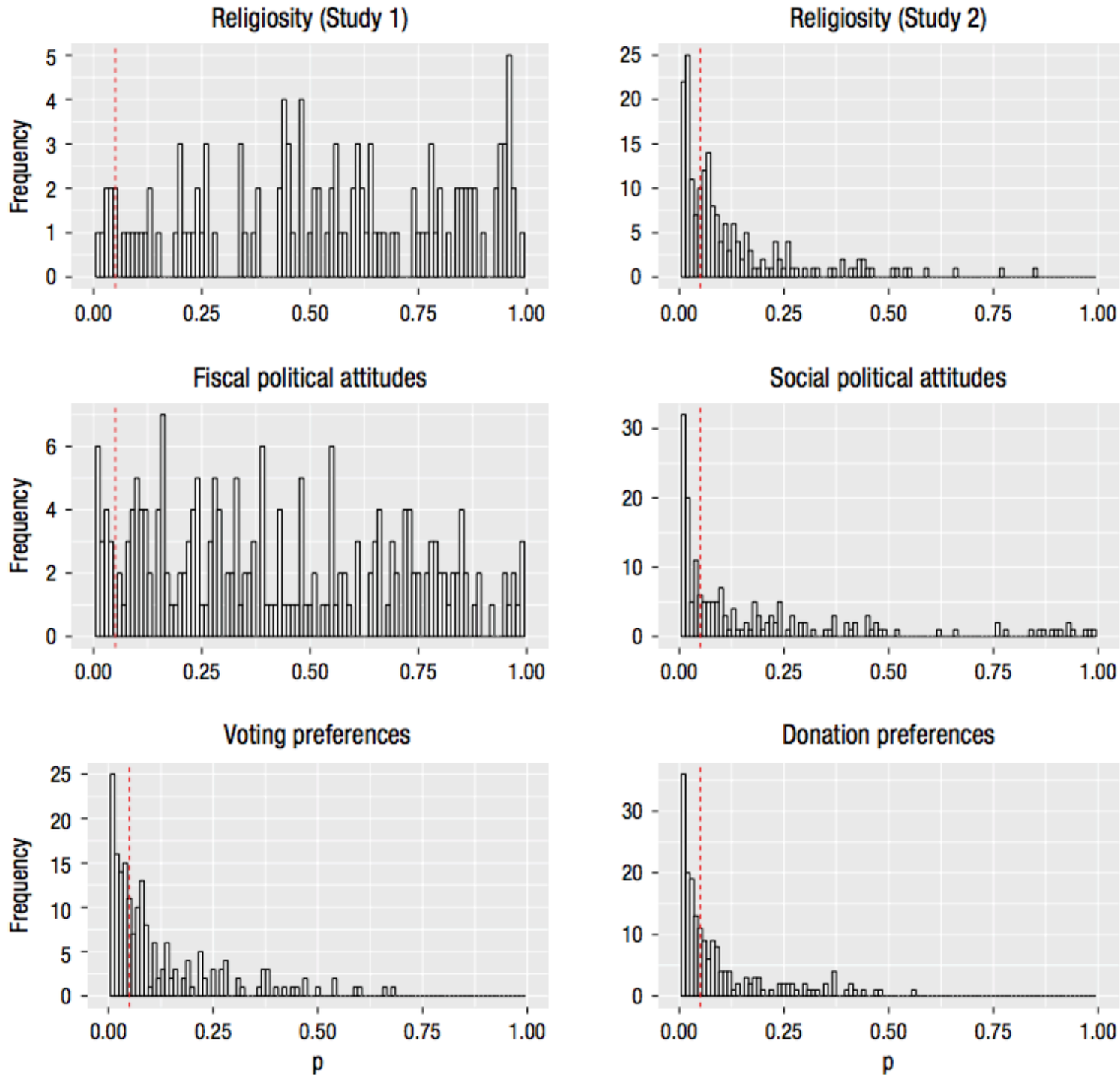


Fig. 1. Histogram of *p* values of the Fertility × Relationship status interaction on religiosity for the multiverse of 120 data sets in Study 1 and 210 data sets in Study 2 (Panels A and B), on fiscal and social political attitudes for the multiverse of 210 data sets in Study 2 (Panels C and D), and on voting and donation preferences for the multiverse of 210 data sets in Study 2 (Panels E and F). The dashed line indicates $p = .05$.

Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing Transparency Through a Multiverse Analysis. *Perspect. Psychol. Sci.* 11, 702–712 (2016).

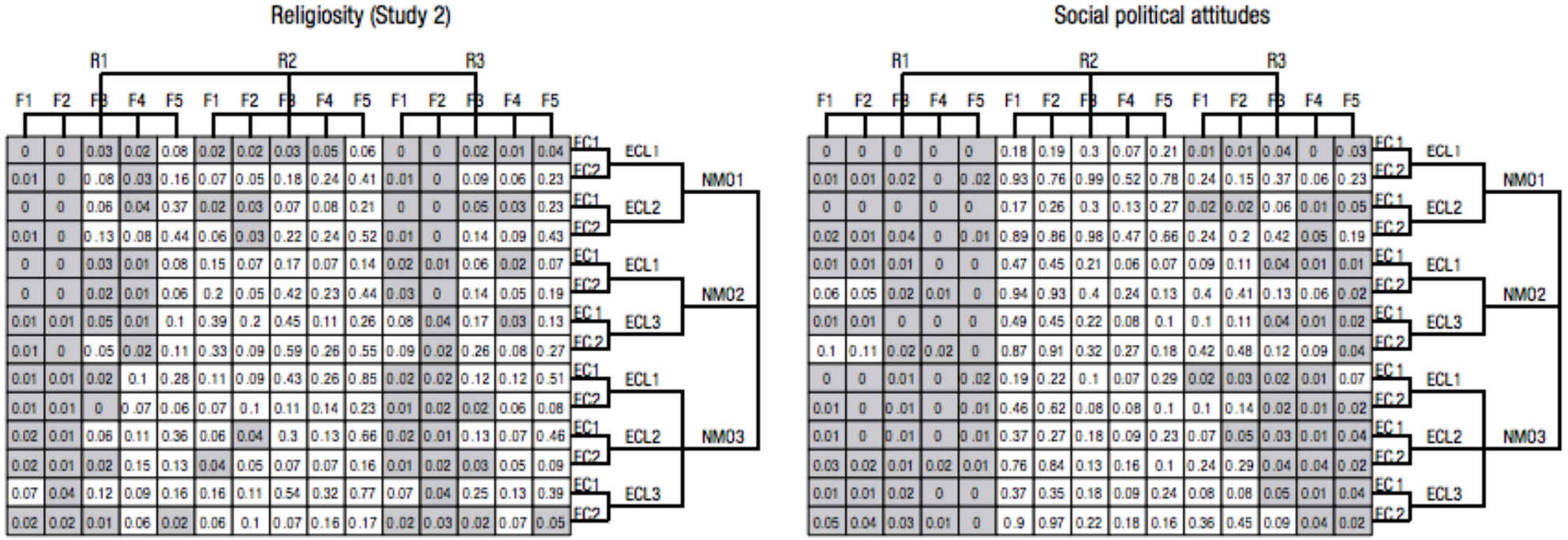


Fig. 2. Visualization of the multiverse of *p* values of the Fertility × Relationship status interaction on religiosity (Panel A), on social political attitudes (Panel B), on voting preferences (Panel C), and on donation preferences (Panel D) in Study 2, showing the dependence of the results on data processing choices. See Table 1 for an explanation of the acronyms.

Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing Transparency Through a Multiverse Analysis. *Perspect. Psychol. Sci.* 11, 702–712 (2016).

Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling

Psychological Science
23(5) 524–532

DOI: 10.1177/0956797611430953
http://pss.sagepub.com
SAGE

Leslie K. John[1], George Loewenstein[2], and Drazen Prelec[3]
[1]Marketing Unit, Harvard Business School; [2]Department of Social & Decision Sciences, Carnegie Mellon University; and [3]Sloan School of Management and Departments of Economics and Brain & Cognitive Sciences, Massachusetts Institute of Technology

Table 1. Results of the Main Study: Mean Self-Admission Rates, Comparison of Self-Admission Rates Across Groups, and Mean Defensibility Ratings

Item	Self-admission rate (%)		Odds ratio (BTS/control)	Two-tailed *p* (likelihood ratio test)	Defensibility rating (across groups)
	Control group	BTS group			
1. In a paper, failing to report all of a study's dependent measures	63.4	66.5	1.14	.23	1.84 (0.39)
2. Deciding whether to collect more data after looking to see whether the results were significant	55.9	58.0	1.08	.46	1.79 (0.44)
3. In a paper, failing to report all of a study's conditions	27.7	27.4	0.98	.90	1.77 (0.49)
4. Stopping collecting data earlier than planned because one found the result that one had been looking for	15.6	22.5	1.57	.00	1.76 (0.48)
5. In a paper, "rounding off" a *p* value (e.g., reporting that a *p* value of .054 is less than .05)	22.0	23.3	1.07	.58	1.68 (0.57)
6. In a paper, selectively reporting studies that "worked"	45.8	50.0	1.18	.13	1.66 (0.53)
7. Deciding whether to exclude data after looking at the impact of doing so on the results	38.2	43.4	1.23	.06	1.61 (0.59)
8. In a paper, reporting an unexpected finding as having been predicted from the start	27.0	35.0	1.45	.00	1.50 (0.60)
9. In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that they do)	3.0	4.5	1.52	.16	1.32 (0.60)
10. Falsifying data	0.6	1.7	2.75	.07	0.16 (0.38)

N=2000 psychologists

IV. Solutions (methodological/statistical education, preregistered studies, etc)

Colquhoun, D. The reproducibility of research and the misinterpretation of p-values. *R Soc Open Sci* 4, 171085 (2017).

- Continue to give p-values, but also effect size and confidence intervals

- Never, ever, use the words ‘significant’ and ‘non-significant’ to describe the results. This wholly arbitrary dichotomy has done untold mischief to the integrity of science

- Consider using a reverse Bayesian approach where you calculate the prior probability required to achieve a specified false positive rate (e.g., 5%)

Colquhoun, D. The reproducibility of research and the misinterpretation of p-values. *R Soc Open Sci* 4, 171085 (2017).

Example:

- A study of transcranial electromagnetic stimulation, published in Science, concluded that it ‘improved associative memory performance’, $p = 0.043$; $N = 8$

- Assuming adequate power, to achieve a false positive risk of 5% when we observe p = 0.043, we would have to assume a prior probability of 0.85 that the effect on memory was genuine

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Psychological Science
22(11) 1359–1366

DOI: 10.1177/0956797611417632
http://pss.sagepub.com
SAGE

Joseph P. Simmons[1], Leif D. Nelson[2], and Uri Simonsohn[1]
[1]The Wharton School, University of Pennsylvania, and [2]Haas School of Business, University of California, Berkeley

Table 2. Simple Solution to the Problem of False-Positive Publications

Requirements for authors

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Guidelines for reviewers

1. Reviewers should ensure that authors follow the requirements.
2. Reviewers should be more tolerant of imperfections in results.
3. Reviewers should require authors to demonstrate that their results do not hinge on arbitrary analytic decisions.
4. If justifications of data collection or analysis are not compelling, reviewers should require the authors to conduct an exact replication.

The preregistration revolution

Brian A. Nosek[a,b,1], Charles R. Ebersole[b], Alexander C. DeHaven[a], and David T. Mellor[a]

[a]Center for Open Science, Charlottesville, VA 22903; and [b]Department of Psychology, University of Virginia, Charlottesville, VA 22904

2600–2606 | PNAS | March 13, 2018 | vol. 115 | no. 11

All too often, we find ourselves unable to predict what will happen; yet after the fact we explain what did happen with a great deal of confidence. This "ability" to explain that which we cannot predict, even in the absence of any additional information, represents an important, though subtle, flaw in our reasoning. It leads us to believe that there is a less uncertain world than there actually is....

— Amos Tversky

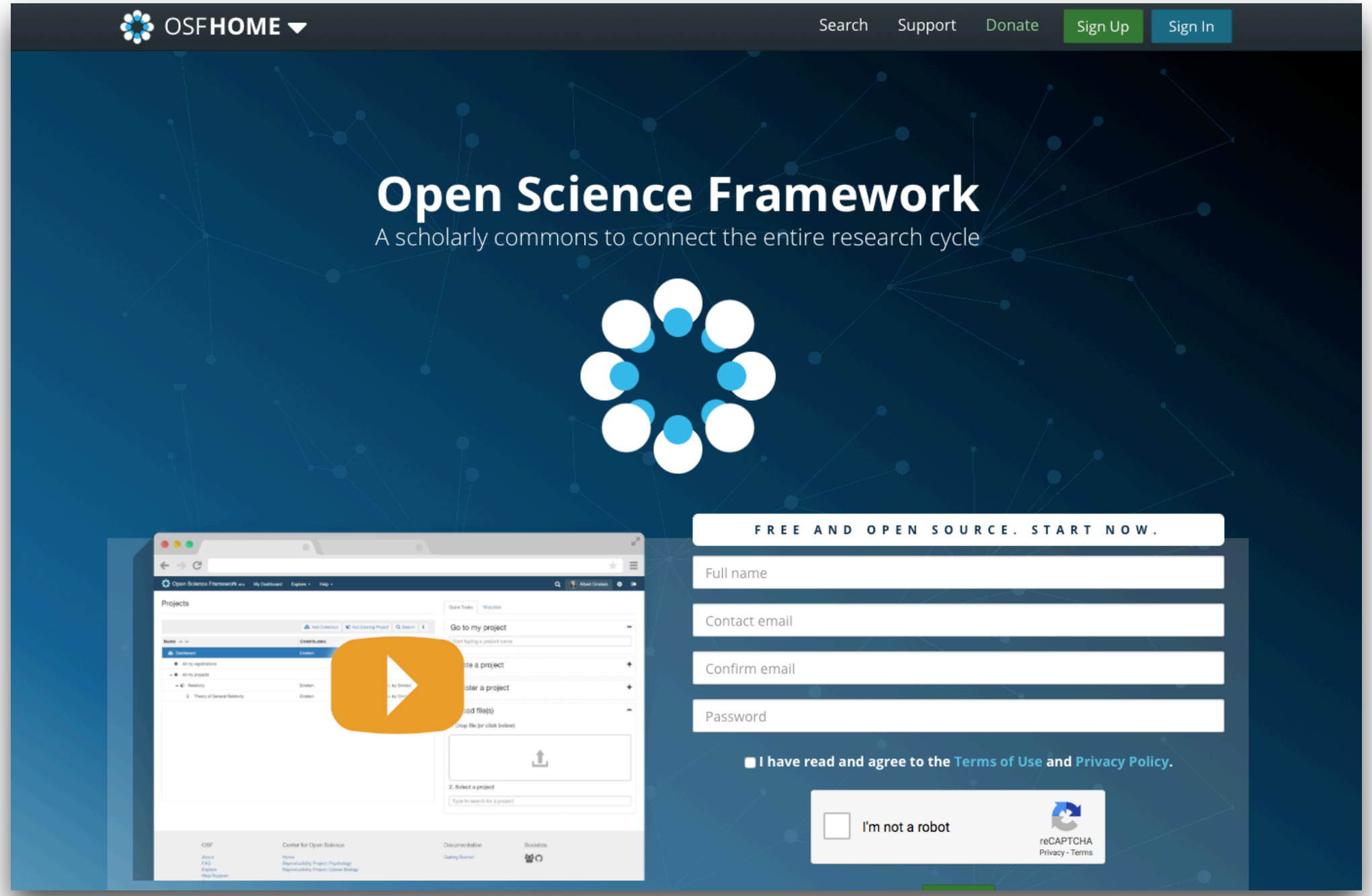
Preregistration benefits

- Makes distinction between confirmatory and exploratory research

- Reduces the effect of publication bias on effect-size estimation

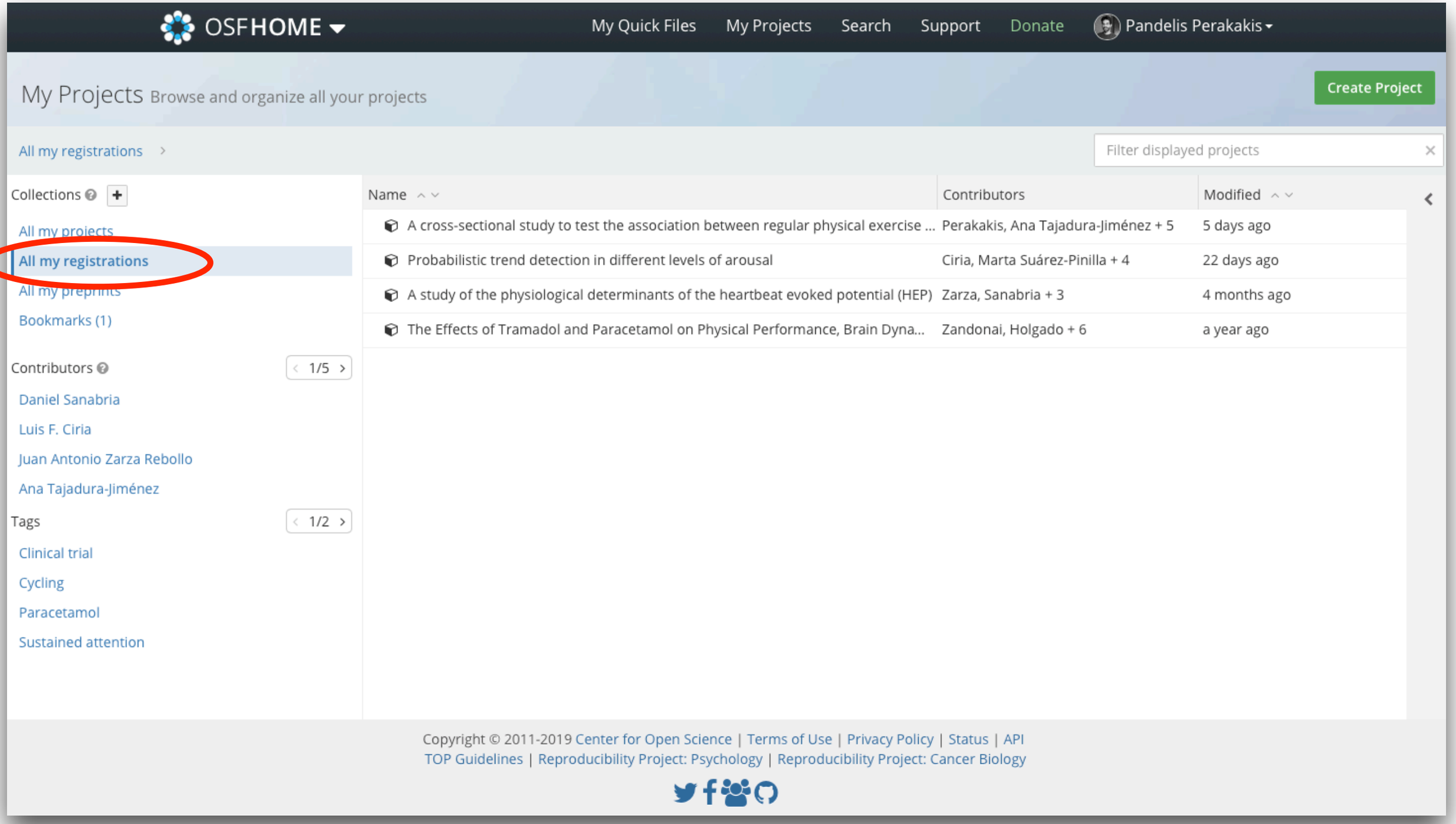
Nosek, B. A. & Lindsay, D. S. Preregistration becoming the norm in psychological science. *APS Obs*. 31, (2018).

Preregistration in practice



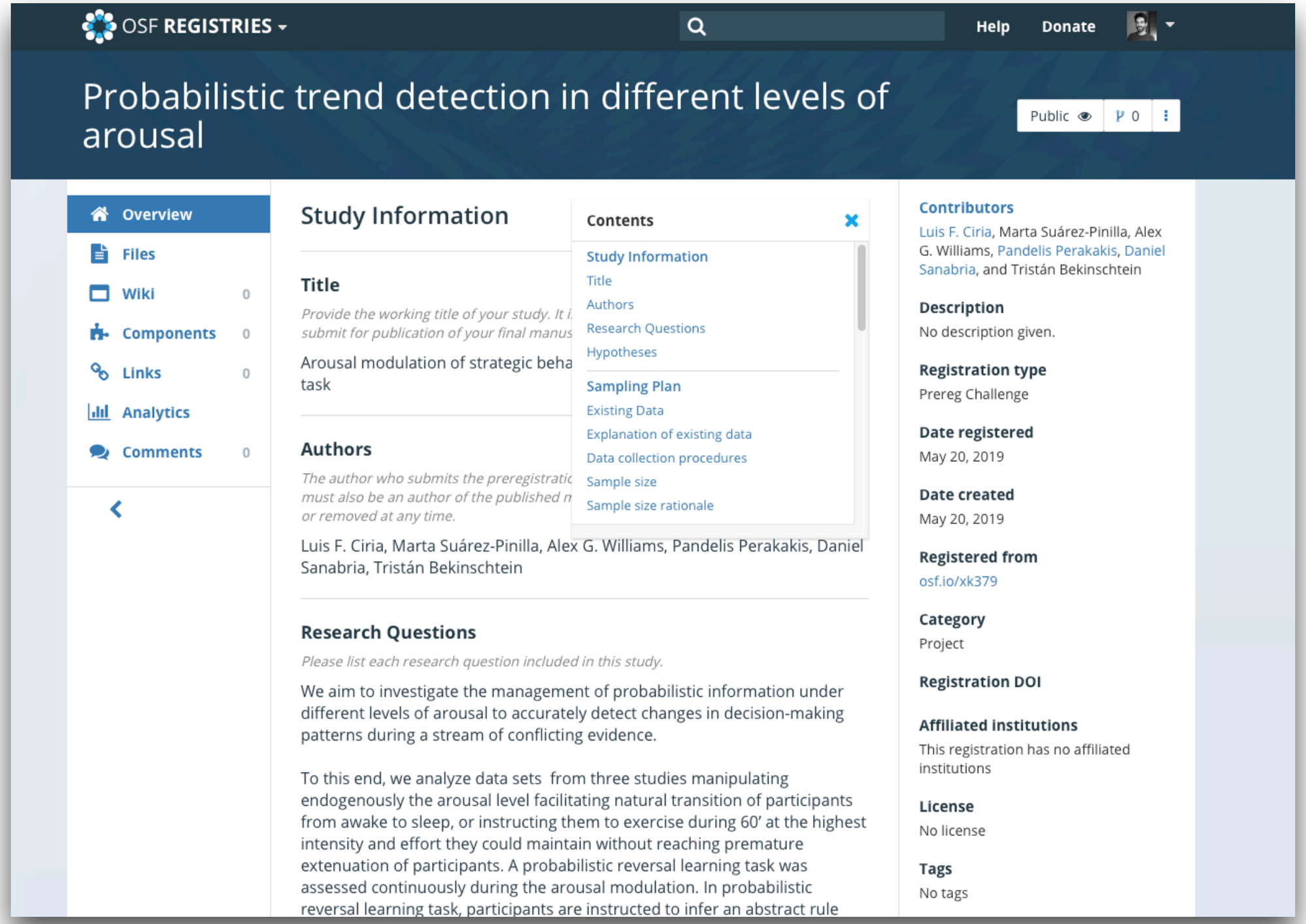
https://osf.io

Preregistration in practice



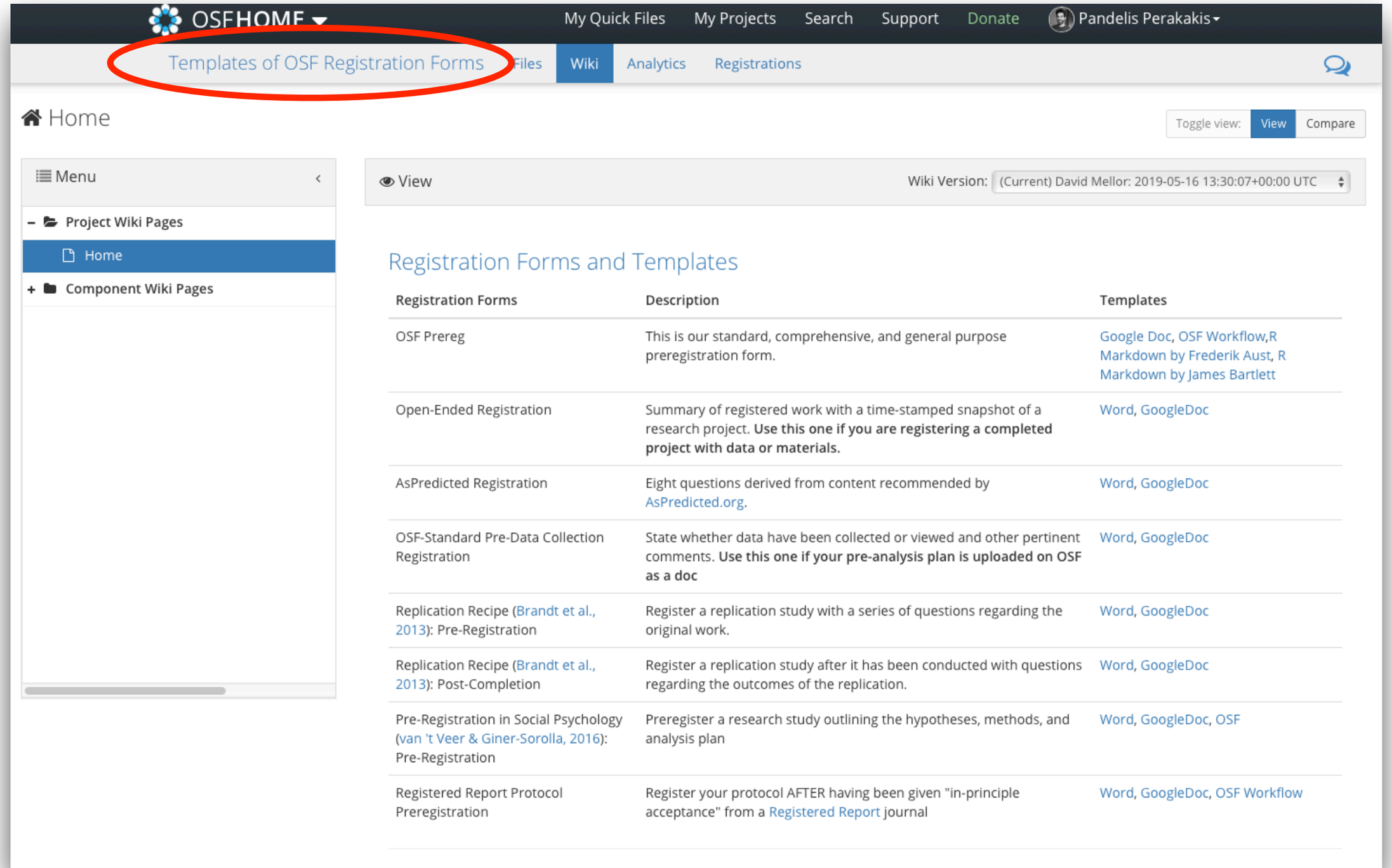
https://osf.io

Preregistration in practice



https://osf.io

Preregistration in practice



https://osf.io/zab38/wiki/home/

Incentives for preregistration

Preregistration Challenge http://cos.io/prereg/

An education campaign for preregistration with $1,000 awards to 1,000 scientists for publishing the results of preregistered research.

Preregistration Badges http://cos.io/badges/

Signals of preregistered research in published articles offered by Psychological Science, Clinical Psychological Science, and xx other psychology journals. In 2015, four Psychological Science papers earn a preregistration badge; in 2016, three did; and in 2017, 19 did.

Registered Reports http://cos.io/rr/

A publishing model in which peer review occurs prior to conducting the research. Offered by APS journals Advances in Methods and Practices in Psychological Science, Psychological Science, and more than 85 other journals.

Badges

Research Report

Writing a Sample Article

And Writing Other Things Too

Alfred R. Wallace[1] and Charles R. Darwin[2]

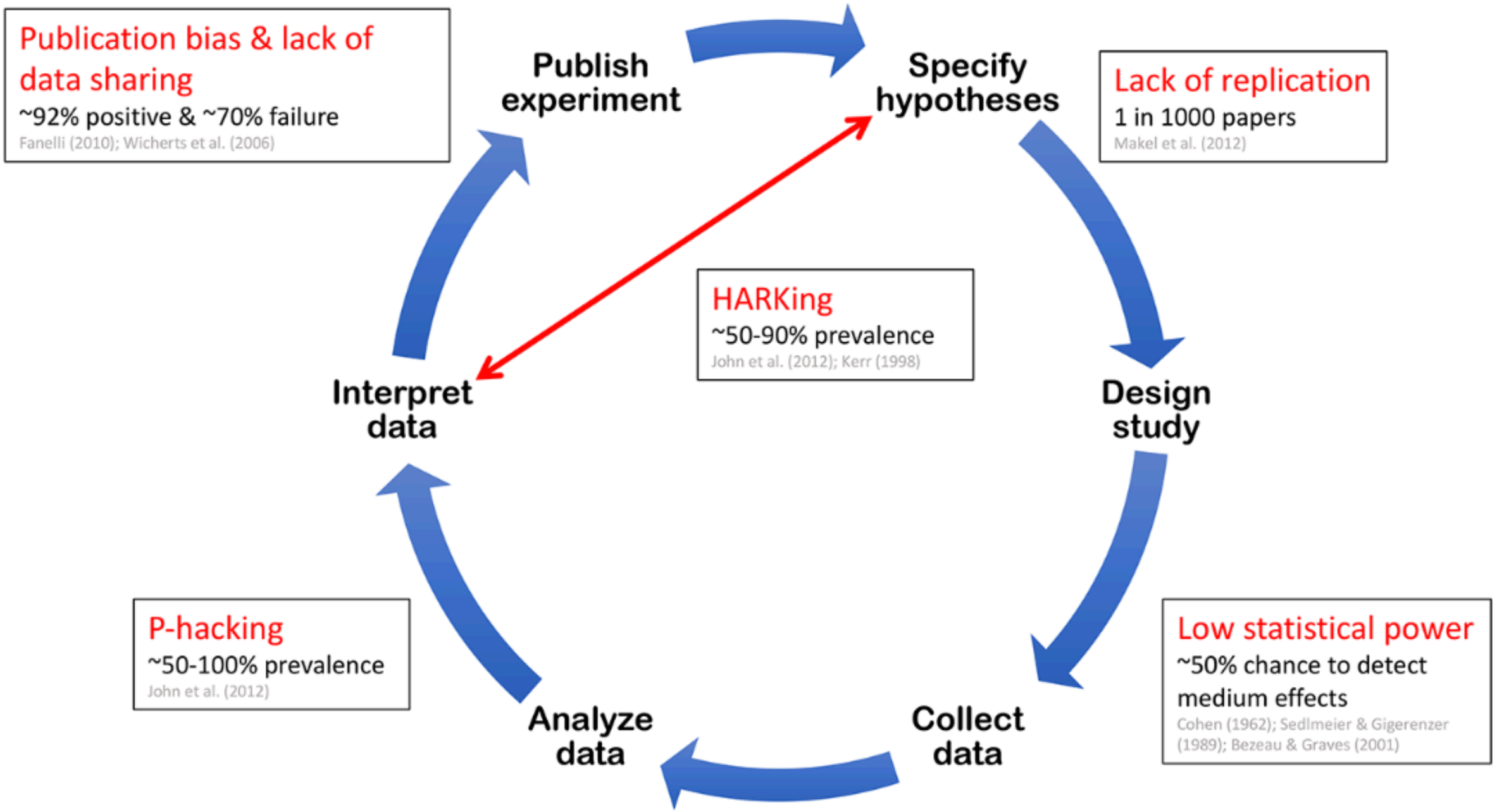
[1]*Collegiate School, Leicester, England, and* [2]*University of Edinburgh Medical School, Edinburgh, Scotland*

Figure 1. Badges earned for open practices: Open Data, Open Materials, and Preregistered. Experiment materials and data are available at http://dvn.iq.harvard.edu/dvn/study?globalId=hdl:1902.1/18708, and the preregistered design and analysis plan is accessible at http://opensciceframework.org/project/TVyXZ/

Registered Reports

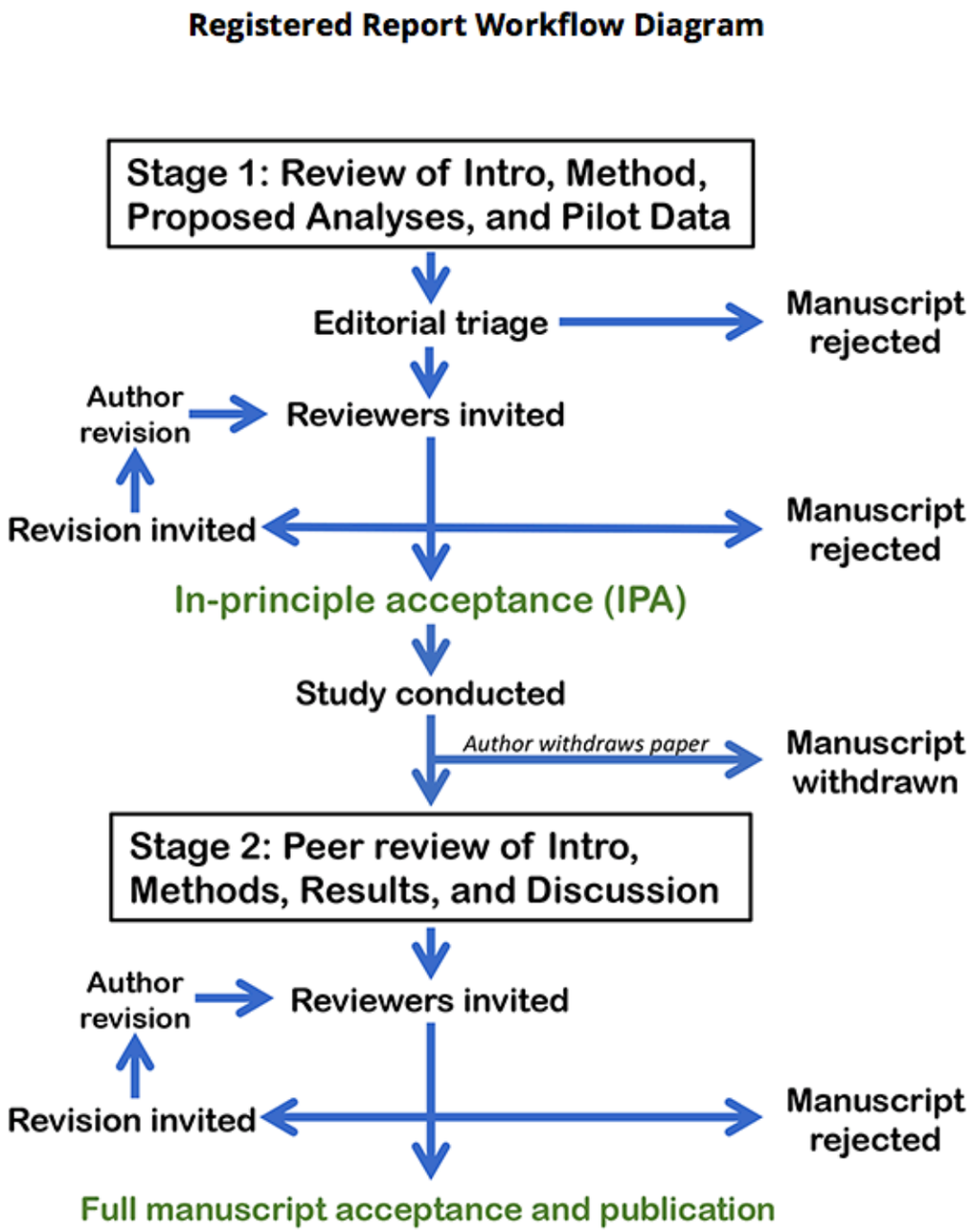
https://cos.io/rr/

Registered Reports

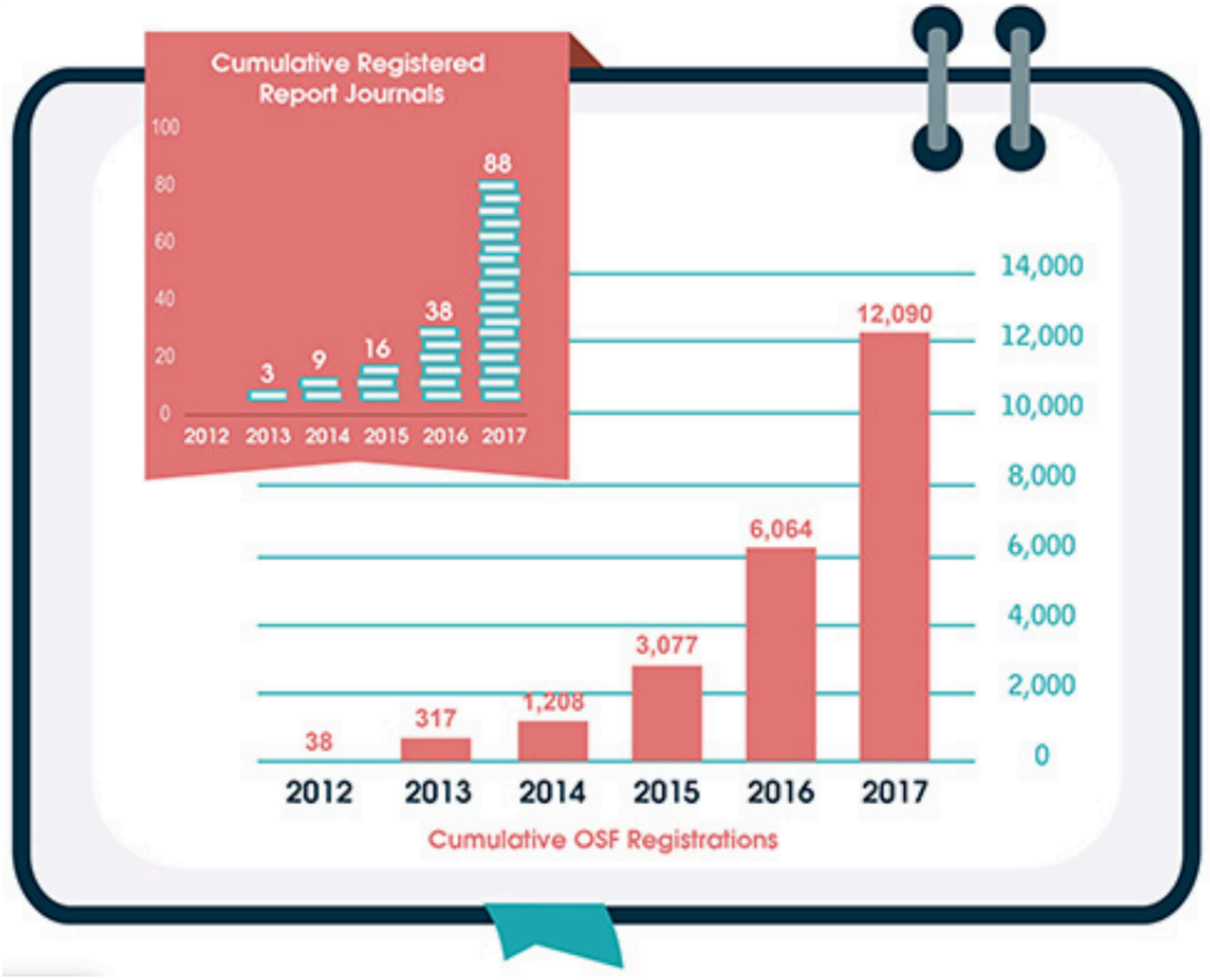


https://cos.io/rr/

Registered Reports



https://cos.io/rr/



Nosek, B. A. & Lindsay, D. S. Preregistration becoming the norm in psychological science. *APS Obs*. 31, (2018).

First analysis of ‘pre-registered’ studies shows sharp rise in null findings

Logging hypotheses and protocols before performing research seems to work as intended: to reduce publication bias for positive results.

30 October 2018

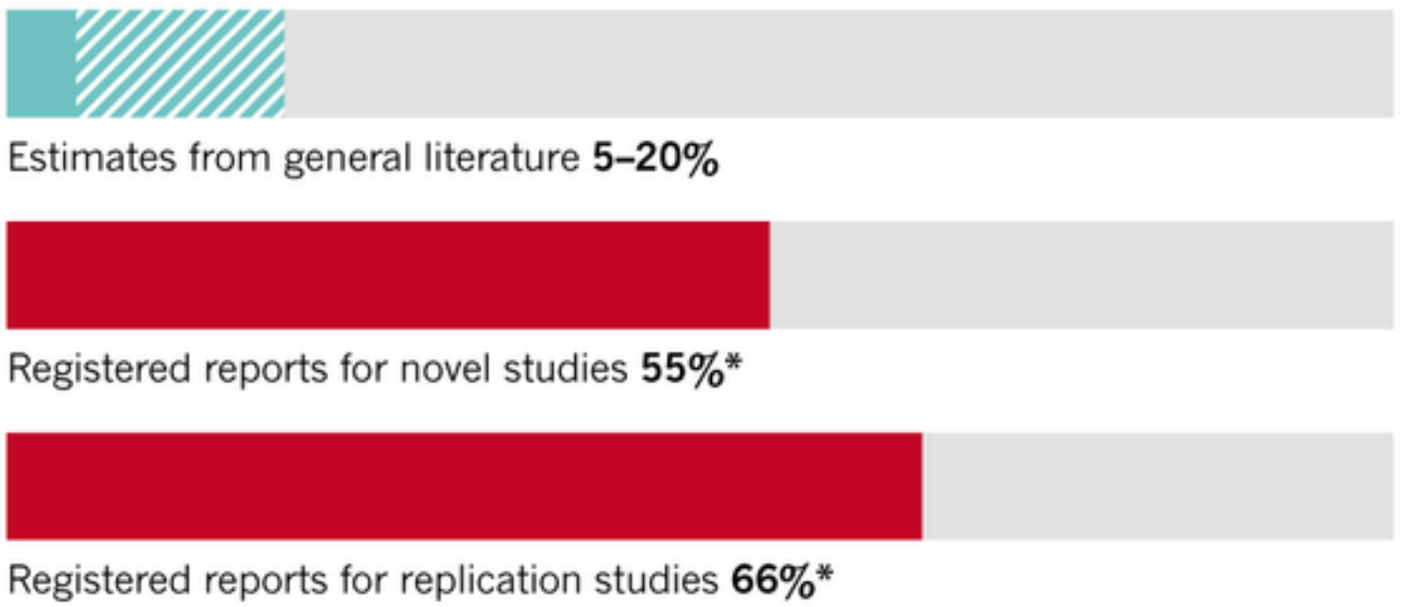
Matthew Warren

nature

REGISTERED REPORTS CUT PUBLICATION BIAS

Pre-registering research protocols in a ‘registered reports’ format could lead to less publication bias skewed towards positive results. Studies that pre-register their protocols publish more negative findings that don’t support their hypothesis, than those that don’t.

HYPOTHESES NOT SUPPORTED BY RESEARCH PAPERS (%)



*Sample size: 296 hypotheses across 113 studies in biomedicine and psychology

PLOS ONE

RESEARCH ARTICLE

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

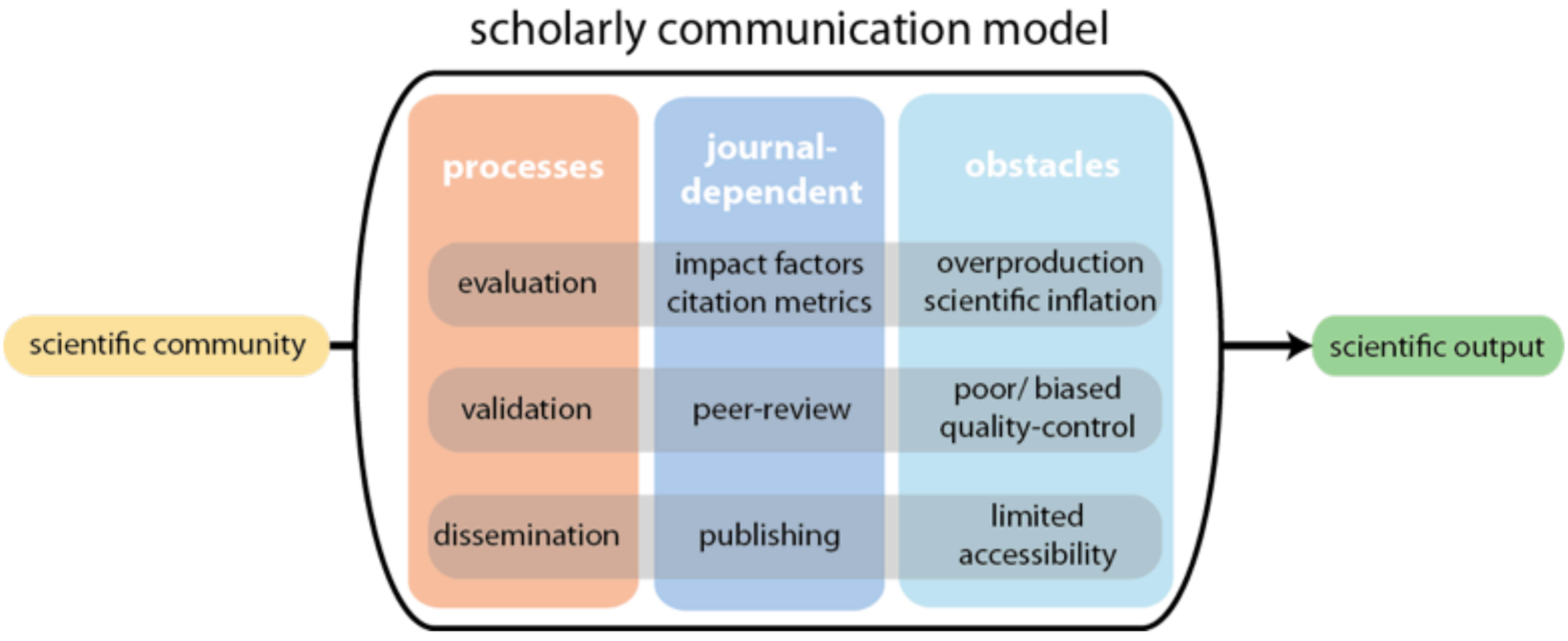
Robert M. Kaplan[1]*, Veronica L. Irvin[2]

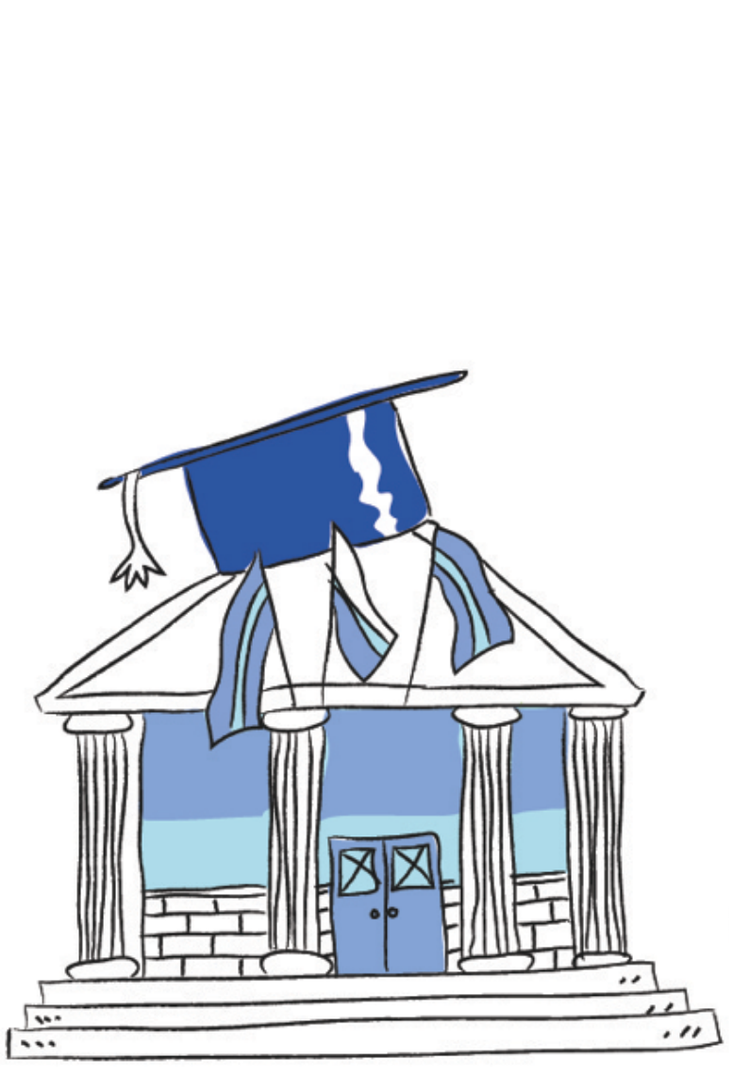
1 Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, Rockville, Maryland, United States of America, **2** Oregon State University, Corvallis, Oregon, United States of America

Number of studies finding medical interventions effective before preregistration introduced:
17/30 (55%)

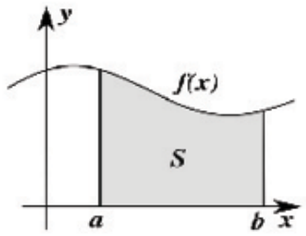
Afterwards:
2/25 (8%)

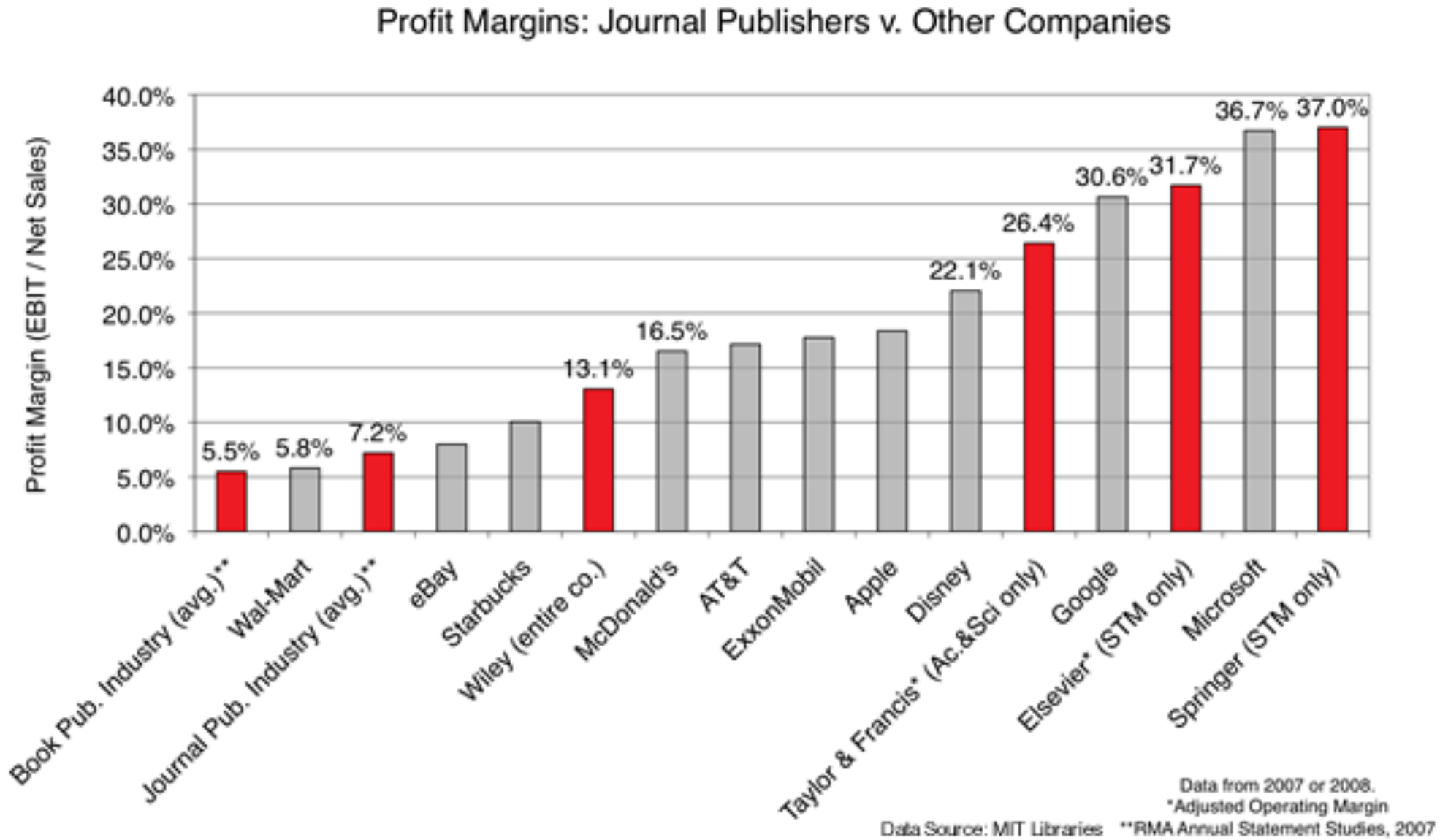
V. Systemic causes (research validation, evaluation and communication)

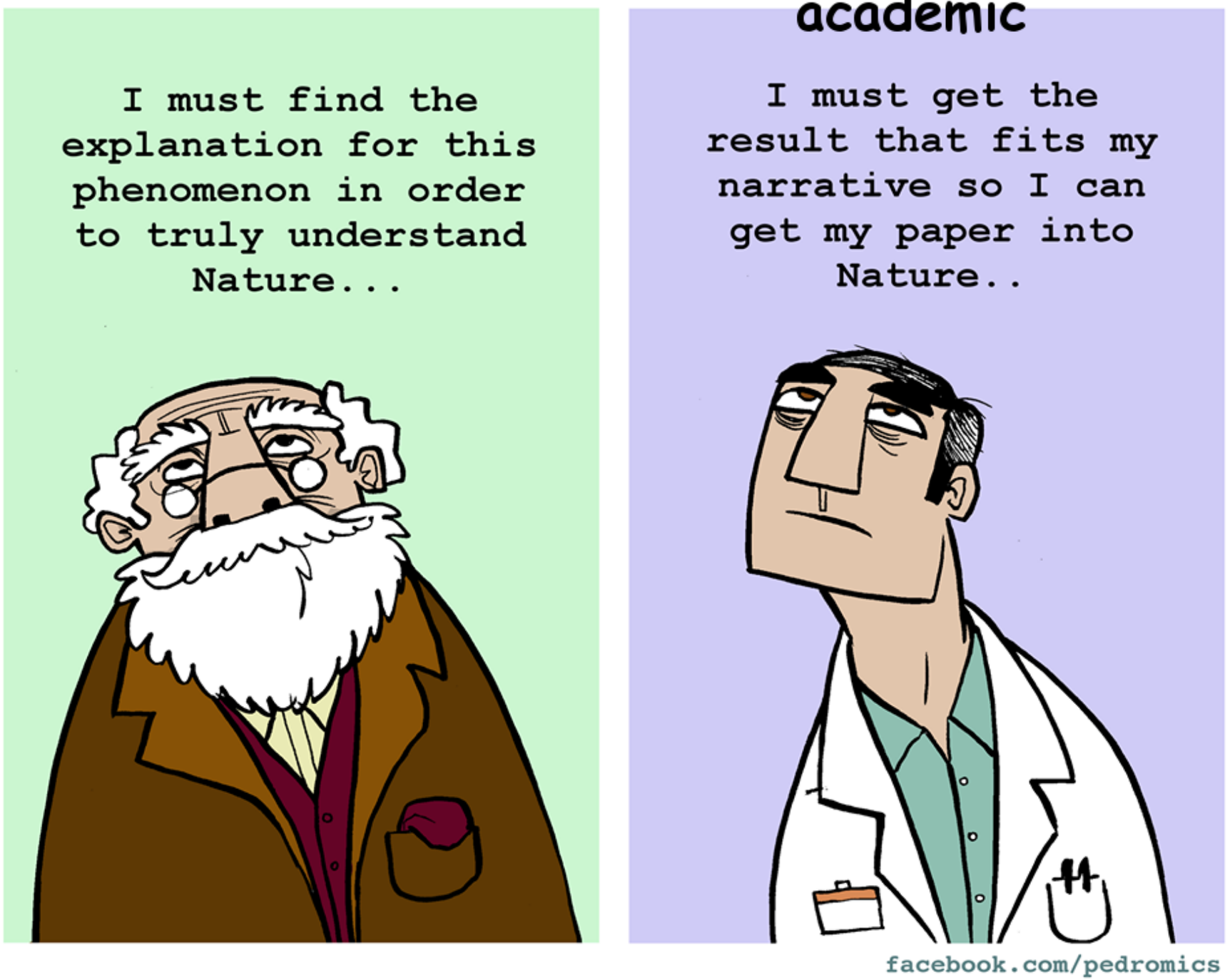

scholarly communication model
processes
journal-dependent
obstacles
evaluation
impact factors
citation metrics
overproduction
scientific inflation
validation
peer-review
poor/ biased
quality-control
dissemination
publishing
limited
accessibility
scientific community
scientific output




JOURNAL


y
f(x)
S
a
b
x


Profit Margins: Journal Publishers v. Other Companies
Profit Margin (EBIT / Net Sales)
40.0%
35.0%
30.0%
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
5.5%
5.8%
7.2%
13.1%
16.5%
22.1%
26.4%
30.6%
31.7%
36.7%
37.0%
Book Pub. Industry (avg.)**
Wal-Mart
Journal Pub. Industry (avg.)**
eBay
Starbucks
Wiley (entire co.)
McDonald's
AT&T
ExxonMobil
Apple
Disney
Taylor & Francis* (Ac.&Sci only)
Google
Elsevier* (STM only)
Microsoft
Springer (STM only)
Data from 2007 or 2008.
*Adjusted Operating Margin
Data Source: MIT Libraries
**RMA Annual Statement Studies, 2007



[Perakakis, P. Open scientists in the shoes of frustrated academics part I: Open-minded scepticism. EuroScientist (2017).]

Scientist	**Academic**
Open-minded	Content journal editors
Sceptical	No replications or negative results
Consider all data	Inaccessible articles, data, software code
Collaborate	Compete for scarce resources
No investment in the outcome	p-hacking

Perakakis, P. Open scientists in the shoes of frustrated academics part I: Open-minded scepticism. EuroScientist (2017).

Dissemination

Budapest Open Access Initiative

Home

BOAI10 Recommendations

Translations

FAQ

View signatures

Sign the the original BOAI

BOAI Forum

Resources

What you can do to help

Contact us

Budapest Open Access Initiative

In response to the growing demand to make research free and available to [illegible] with a computer and an internet connection, a diverse coalition has issued new guidelines that could usher in huge advances in the sciences, medicine, and health.

The recommendations were developed by leaders of the Open Access movement, which has worked for the past decade to provide the public with unrestricted, free access to scholarly research—much of which is publicly funded. Making the research publicly available to everyone—free of charge and without most copyright and licensing restrictions—will accelerate scientific research efforts and allow authors to reach a larger number of readers.

The recommendations are the result of a meeting organized by the Open Society Foundations to mark the tenth anniversary of Budapest Open Access Initiative, which first defined Open Access. The recommendations include the development of Open Access policies in institutions of higher education and in funding agencies, the open licensing of scholarly works, the development of infrastructure such as Open Access repositories and creating standards of professional conduct for Open Access publishing. The recommendations also establish a new goal of achieving Open Access as the default method for distributing new peer-reviewed research in every field and in every country within ten years' time.

Translations of the recommendations have already been made in several languages, with more to follow.

For more on the recommendations, please see the press release as well as a blog post by Peter Suber which provides additional background on the Open Access movement.

The two roads to Open Access

Green

Gold

To achieve open access to scholarly journal literature, we recommend two complementary strategies.

I. Self-Archiving: First, scholars need the tools and assistance to deposit their refereed journal articles in open electronic archives, a practice commonly called, self-archiving. When these archives conform to standards created by the Open Archives Initiative, then search engines and other tools can treat the separate archives as one. Users then need not know which archives exist or where they are located in order to find and make use of their contents.

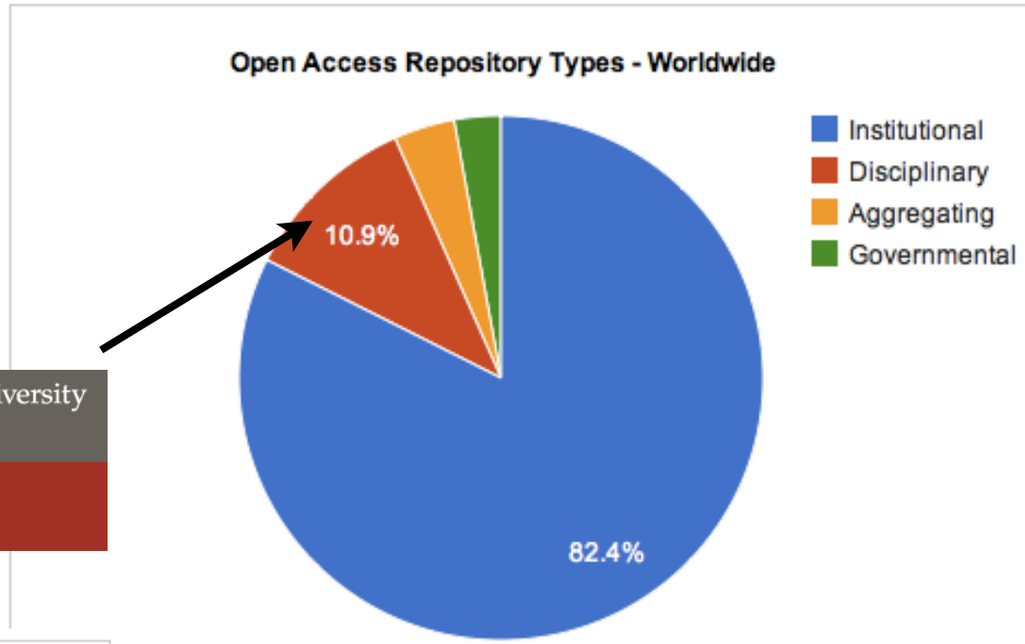
II. Open-access Journals: Second, scholars need the means to launch a new generation of journals committed to open access, and to help existing journals that elect to make the transition to open access. Because journal articles should be disseminated as widely as possible, these new journals will no longer invoke copyright to restrict access to and use of the material they publish. Instead they will use copyright and other tools to ensure permanent open access to all the articles they publish. Because price is a barrier to access, these new journals will not charge subscription or access fees, and will turn to other methods for covering their expenses. There are many alternative sources of funds for this purpose, including the foundations and governments that fund research, the universities and laboratories that employ researchers, endowments set up by discipline or institution, friends of the cause of open access, profits from the sale of add-ons to the basic texts, funds freed up by the demise or cancellation of journals charging traditional subscription or access fees, or even contributions from the researchers themselves. There is no need to favor one of these solutions over the others for all disciplines or nations, and no need to stop looking for other, creative alternatives.

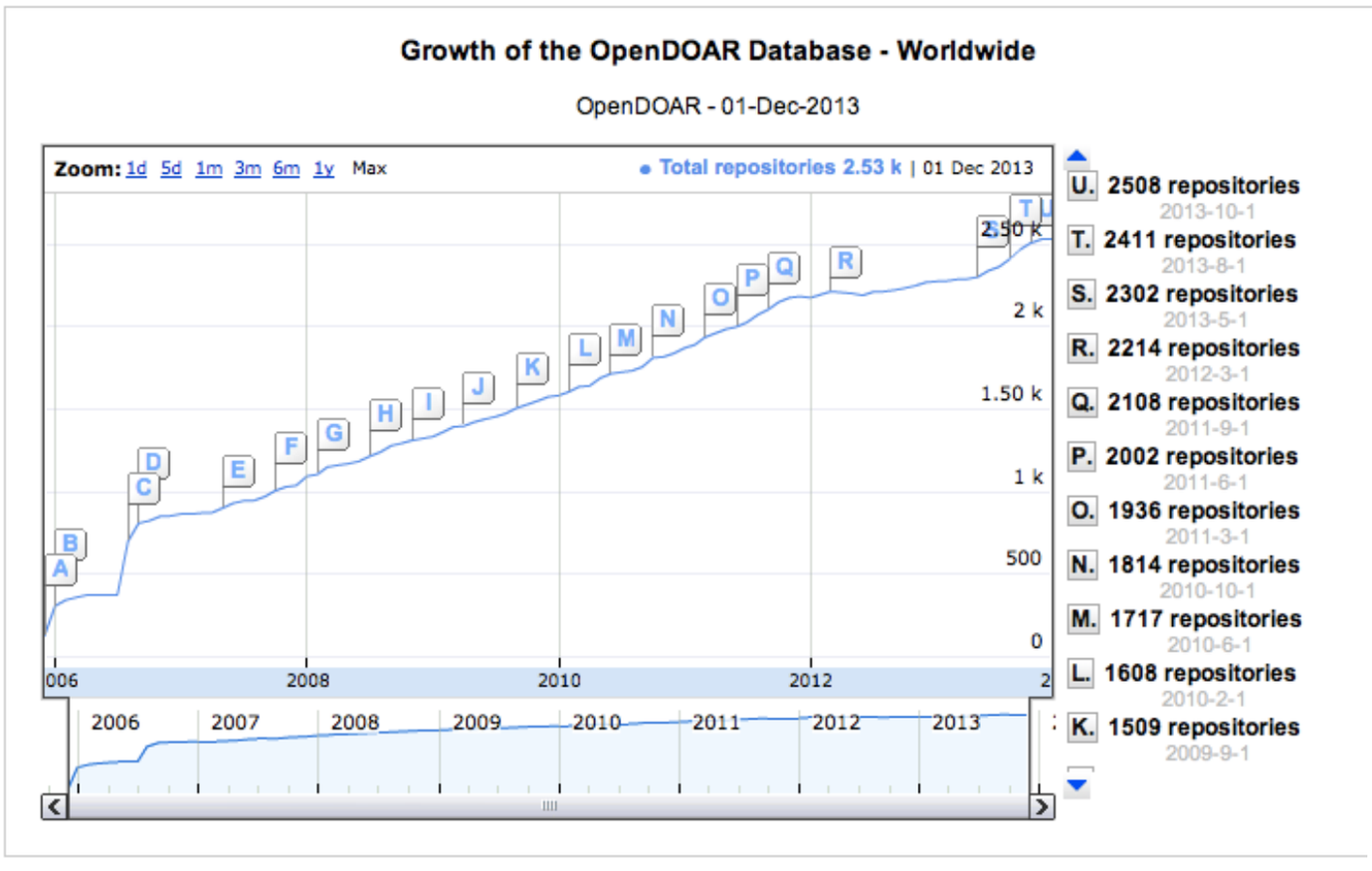
green repositories

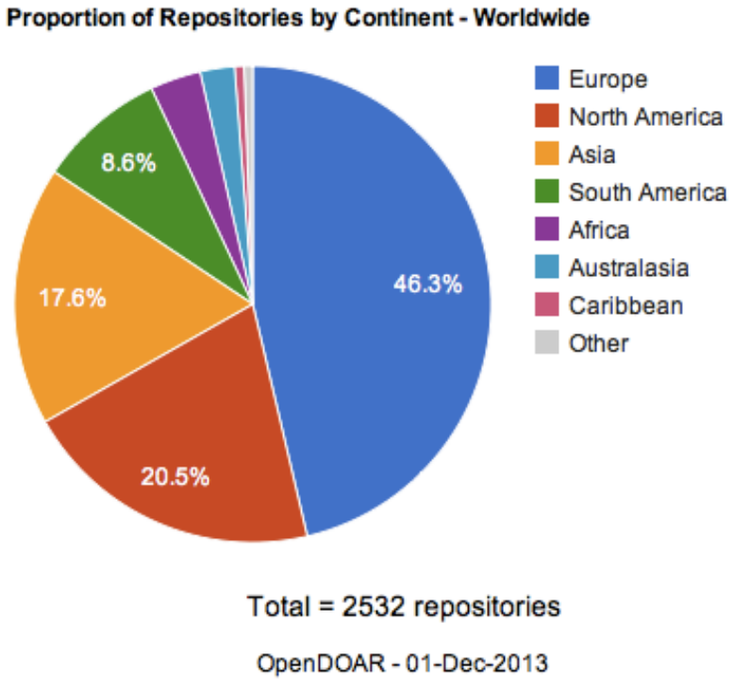
OpenDOAR

Cornell University Library

arXiv.org





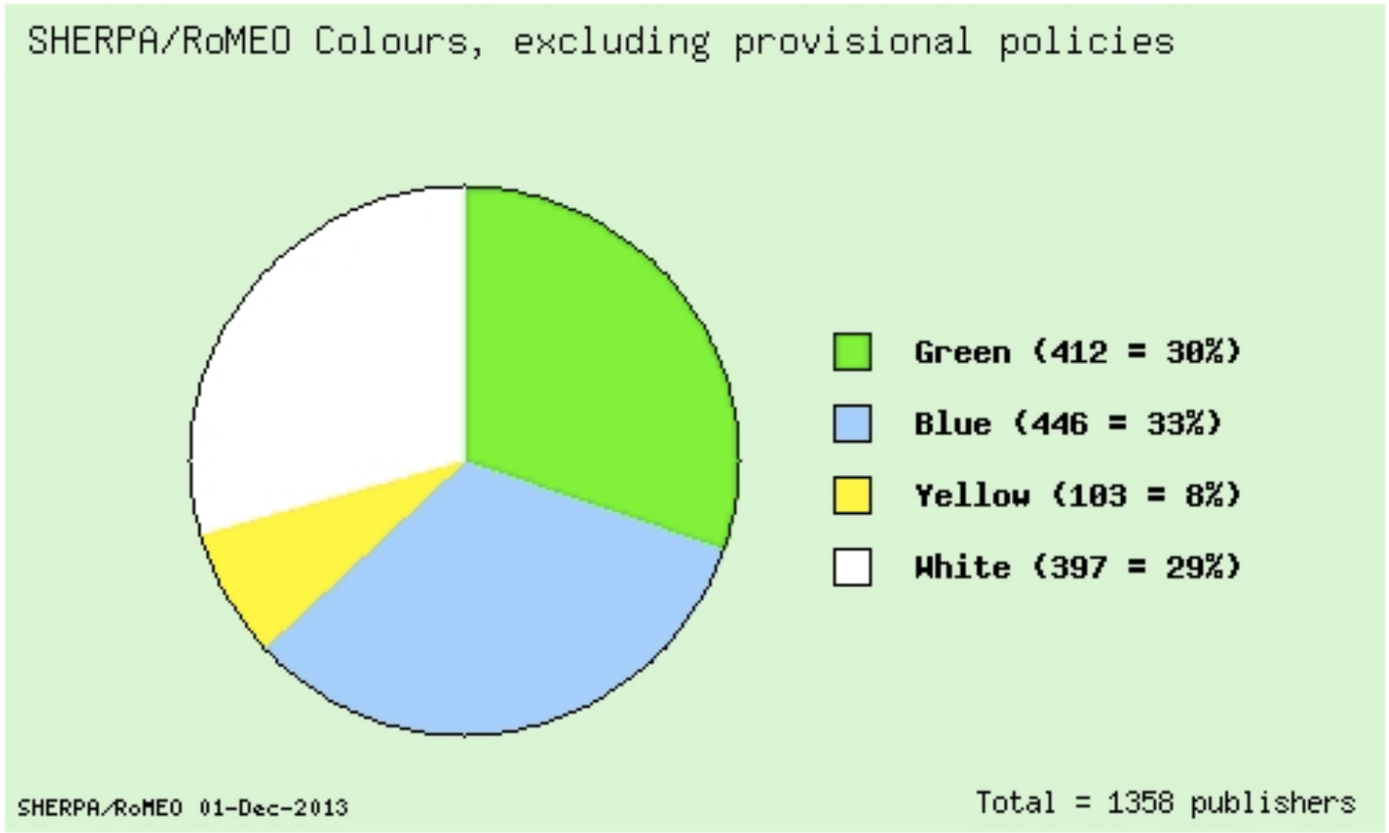


http://www.opendoar.org/

SHERPA/RoMEO

Search - Publisher copyright policies & self-archiving

Summary: **71%** of publishers on this list formally **allow** some form of self-archiving.

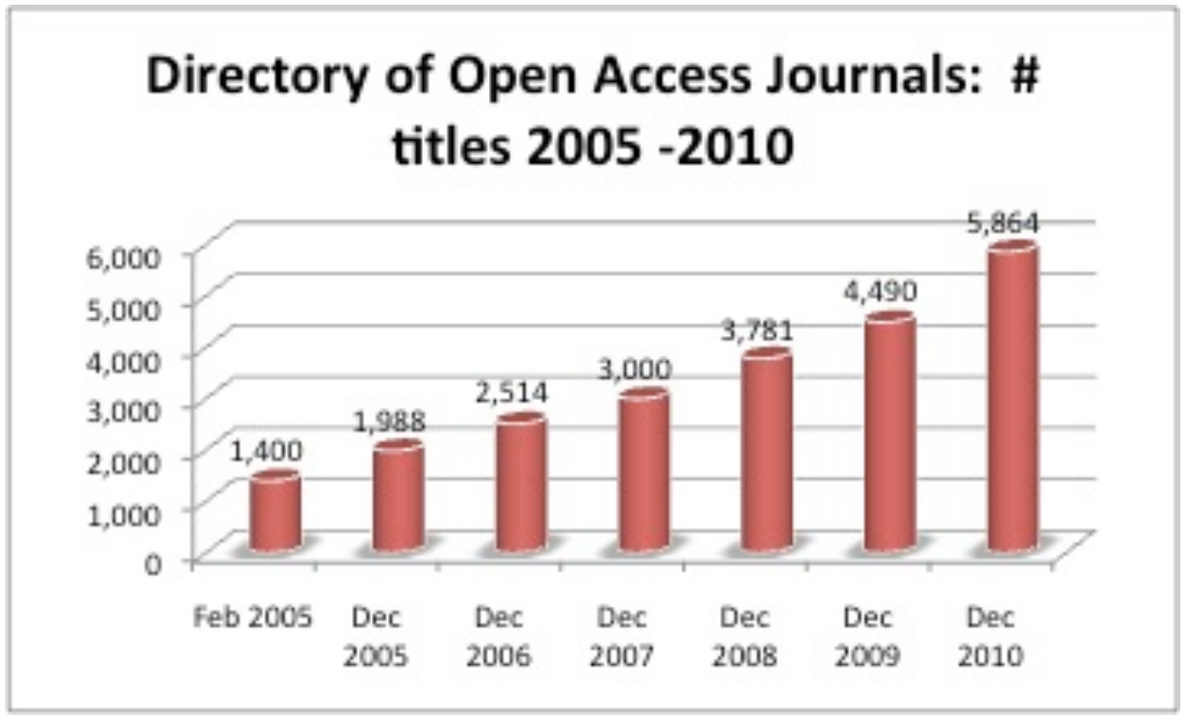


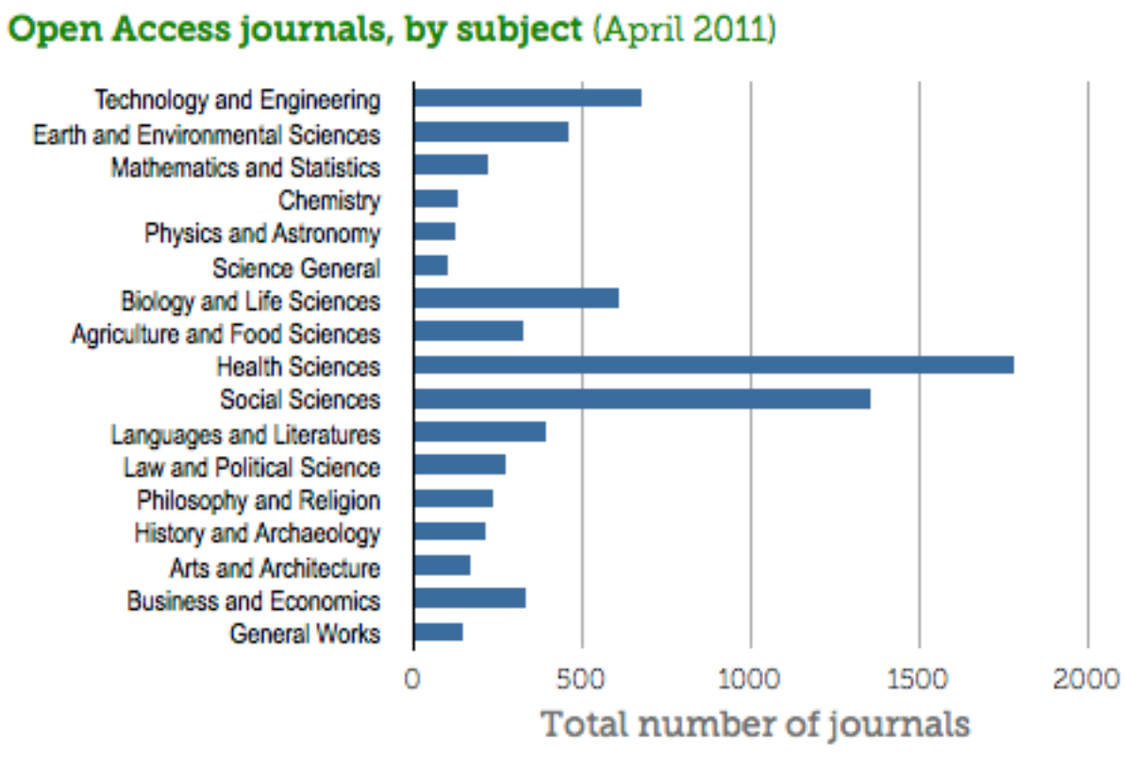
RoMEO colour	Archiving policy
GREEN	Can archive pre-print and post-print
BLUE	Can archive post-print
YELLOW	Can archive pre-print
WHITE	Archiving not formally supported

http://www.sherpa.ac.uk/romeo/

and golden journals

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS





http://www.doaj.org/

The two open-access strategies: **Gold** and **Green**

Open-Access Publishing (OApub) (BOAI-2)

1. Create or Convert 23,000 open-access journals (1000 exist currently)
2. Find funding support for open-access publication costs ($500-$1500+)
3. Persuade the authors of the annual 2,500,000 articles to publish in new open-access journals instead of the existing toll-access journals

Open-Access Self-Archiving (OAarch) (BOAI-1)

1. Persuade the authors of the annual 2,500,000 articles they publish in the existing toll-access journals to also self-archive them in their institutional open-access archives.

Why gold then???

How publishers think...

Whatever one may think about the relative merits of Green and Gold OA (a matter that my colleagues on the Kitchen and myself have discussed numerous times) or the economic implications of embargoes of various lengths, what is clear is that Green OA has no promise of delivering augmented revenues to the publisher, but Gold OA opens up a new customer, the author him or herself, who in many instances pays for the article to be OA. Gold OA, in other words, represents a business opportunity, whereas Green OA represents a business problem.

Thus we have the emergence of a relatively new market, where publishers fight to collect fees from this new class of customers: authors. How to compete is another matter. Most traditional publishers rely on the strength of their brands to bring the articles in. This is most obvious in cascading peer review, where the established publication represents the wide end of the marketing funnel and the Gold OA venues sit at the narrow end. (It's worth remembering that this model works for purely toll-access publications as well, as the enormous success of *Nature*'s line-extension proves.) Other publishers focus on metrics of different kinds and boast of their Web-friendly tools for submission, discovery, and dissemination. As one would expect, wherever there is competition, the matter of pricing comes up. And here the established publisher may have a problem.

About Joseph Esposito

I am a management consultant working primarily in the world of digital media, software, and publishing. My clients include both for-profits and not-for-profits. A good deal of my activity concerns research publishing, especially when the matter at issue has to do with the migration to digital services from a print background. Prior to setting up my consulting business, I served as CEO of three companies (Encyclopaedia Britannica, Tribal Voice, and SRI Consulting), all of which I led to successful exits. Typically I work on strategy issues, advising CEOs and Boards of Directors on direction; I also have managed a number of sticky turnarounds. Among other things, I have been the recipient of grants from the Mellon, MacArthur, and Hewlett Foundations, all concerning research into new aspects of publishing.

View all posts by Joseph Esposito »

Money is power!!!

"The Finch Report is a successful case of lobbying by publishers to protect the interests of publishing at the expense of the interests of research and the public that funds research," argues University of Southampton cognitive scientist Stevan Harnad. "The Finch Report proposes doing precisely what the US Research Works

We therefore recommend that:

i. a clear policy direction should be set towards support for publication in open access or hybrid journals, funded by APCs, as the main vehicle for the publication of research, especially when it is publicly funded;

Key actions: overall policy and funding arrangements

i. Make a clear commitment to support the costs of an innovative and sustainable research communications system , with a clear preference for publication in open access or hybrid journals. (*Government, Research Councils, Funding Councils, universities*)

Validation

The peer review drugs don’t work

A process at the heart of science is based on faith rather than evidence, says Richard Smith, and vested interests keep it in place

Peer review is supposed to be the quality assurance system for science, weeding out the scientifically unreliable and reassuring readers of journals that they can trust what they are reading. In reality, however, it is ineffective, largely a lottery, anti-innovatory, slow, expensive, wasteful of scientific time, inefficient, easily abused, prone to bias, unable to detect fraud and irrelevant.

Perhaps the biggest argument against the peer review of completed studies is that it simply isn’t needed. With the World Wide Web everything can be published, and the world can decide what’s important and what isn’t. This proposition strikes terror into many hearts, but with so much poor-quality science published what do we have to lose?

Richard Smith, former *British Medical Journal* editor

May 28, 2015

Effects of Editorial Peer Review
A Systematic Review

Tom Jefferson, MD
Philip Alderson, MBChB
Elizabeth Wager, MA
Frank Davidoff, MD

Conclusions Editorial peer review, although widely used, is largely untested and its effects are uncertain.

JAMA. 2002;287:2784-2786

www.jama.com

Cochrane

Trusted evidence.
Informed decisions.
Better health.

Editorial peer review for improving the quality of reports of biomedical studies

Published:
18 April 2007

Authors:
Jefferson T, Rudin M, Brodney Folse S, Davidoff F

Authors' conclusions:

At present, little empirical evidence is available to support the use of editorial peer review as a mechanism to ensure quality of biomedical research.

Published online 5 October 2011 | *Nature* **478**, 26-28 (2011) |
doi:10.1038/478026a

nature International weekly journal of science

News Feature

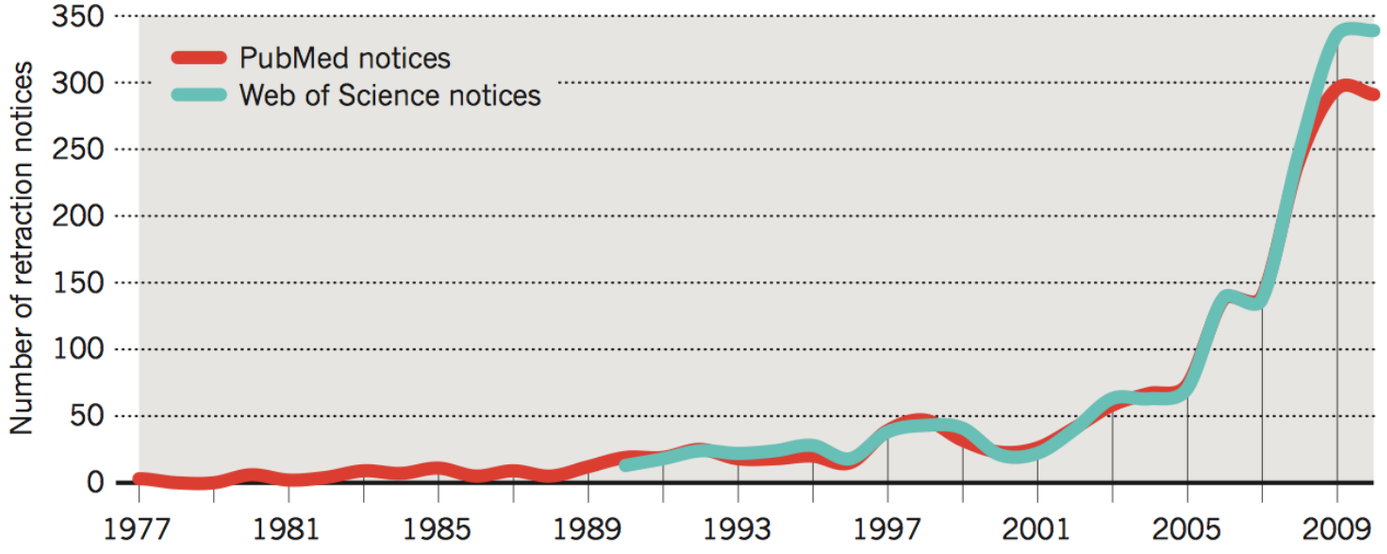
Science publishing: The trouble with retractions

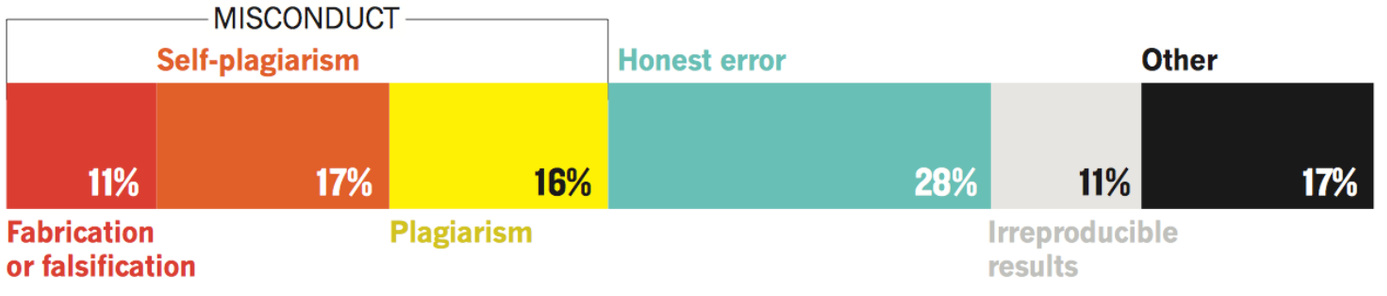
A surge in withdrawn papers is highlighting weaknesses in the system for handling them.

Richard Van Noorden

RISE OF THE RETRACTIONS

In the past decade, the number of retraction notices has shot up 10-fold (**top**), even as the literature has expanded by only 44%. It is likely that only about half of all retractions are for researcher misconduct (**middle**). Higher-impact journals have logged more retraction notices over the past decade, but much of the increase during 2006–10 came from lower-impact journals (**bottom**).

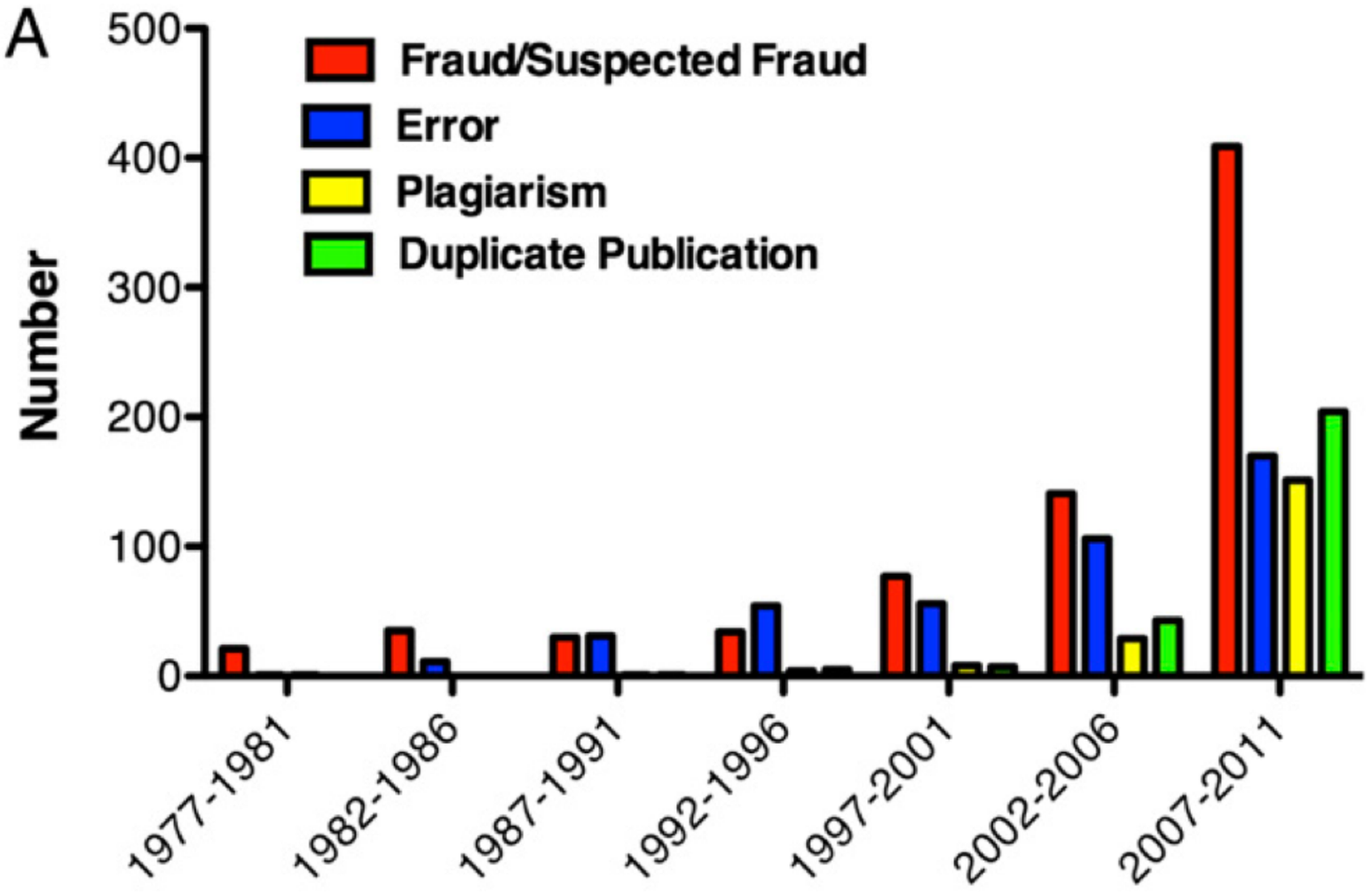




PNAS

Misconduct accounts for the majority of retracted scientific publications

Ferric C. Fang[a,b,1], R. Grant Steen[c,1], and Arturo Casadevall[d,1,2]



nature International weekly journal of science

Drug development: Raise standards for preclinical cancer research

C. Glenn Begley & Lee M. Ellis

85 per cent of preclinical studies could not be replicated

Building a stronger system

What reasons underlie the publication of erroneous, selective or irreproducible data? The academic system and peer-review process tolerates and perhaps even inadvertently encourages such conduct[5]. To obtain funding, a job, promotion or tenure, researchers need a strong publication record, often including a first-authored high-impact publication. Journal editors, reviewers and grant-review committees often look for a scientific finding that is simple, clear and complete — a 'perfect' story. It is therefore tempting for investigators to submit selected data sets for publication, or even to massage data to fit the underlying hypothesis.

Nature **483**, 531–533 (29 March 2012) | doi:10.1038/483531a
Published online 28 March 2012

Jointly published by Akadémiai Kiadó, Budapest
and Springer, Dordrecht

Scientometrics, Vol. 81, No. 2 (2009) 549–565
DOI: 10.1007/s11192-008-2141-5

Rejecting and resisting Nobel class discoveries: accounts by Nobel Laureates

JUAN MIGUEL CAMPANARIO

Departamento de Física, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain

I review and discuss instances in which 19 future Nobel Laureates encountered resistance on the part of the scientific community towards their discoveries, and instances in which 24 future Nobel Laureates encountered resistance on the part of scientific journal editors or referees to manuscripts that dealt with discoveries that later would earn them the Nobel Prize.

Evaluation

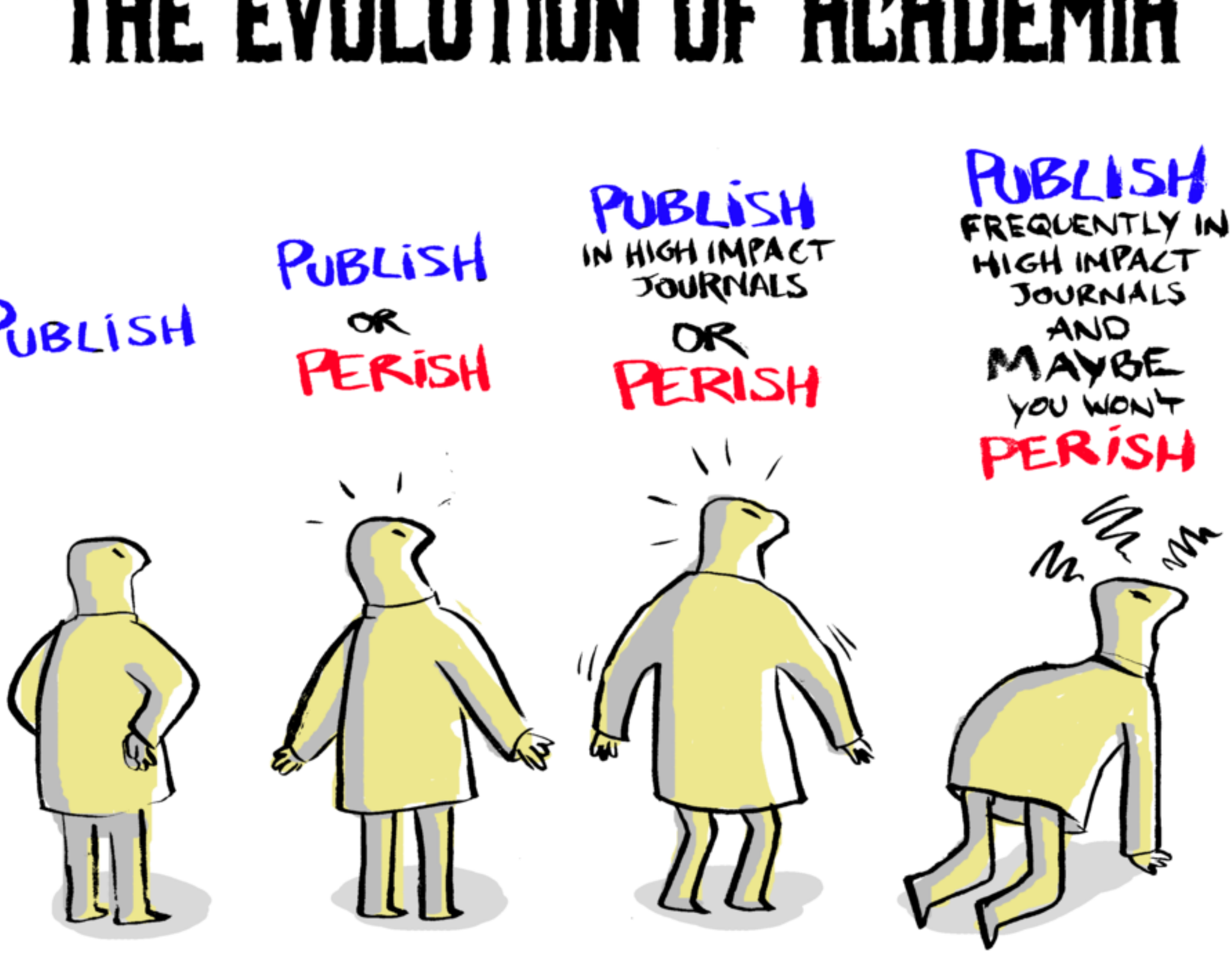
THE EVOLUTION OF ACADEMIA

PUBLISH

PUBLISH OR PERISH

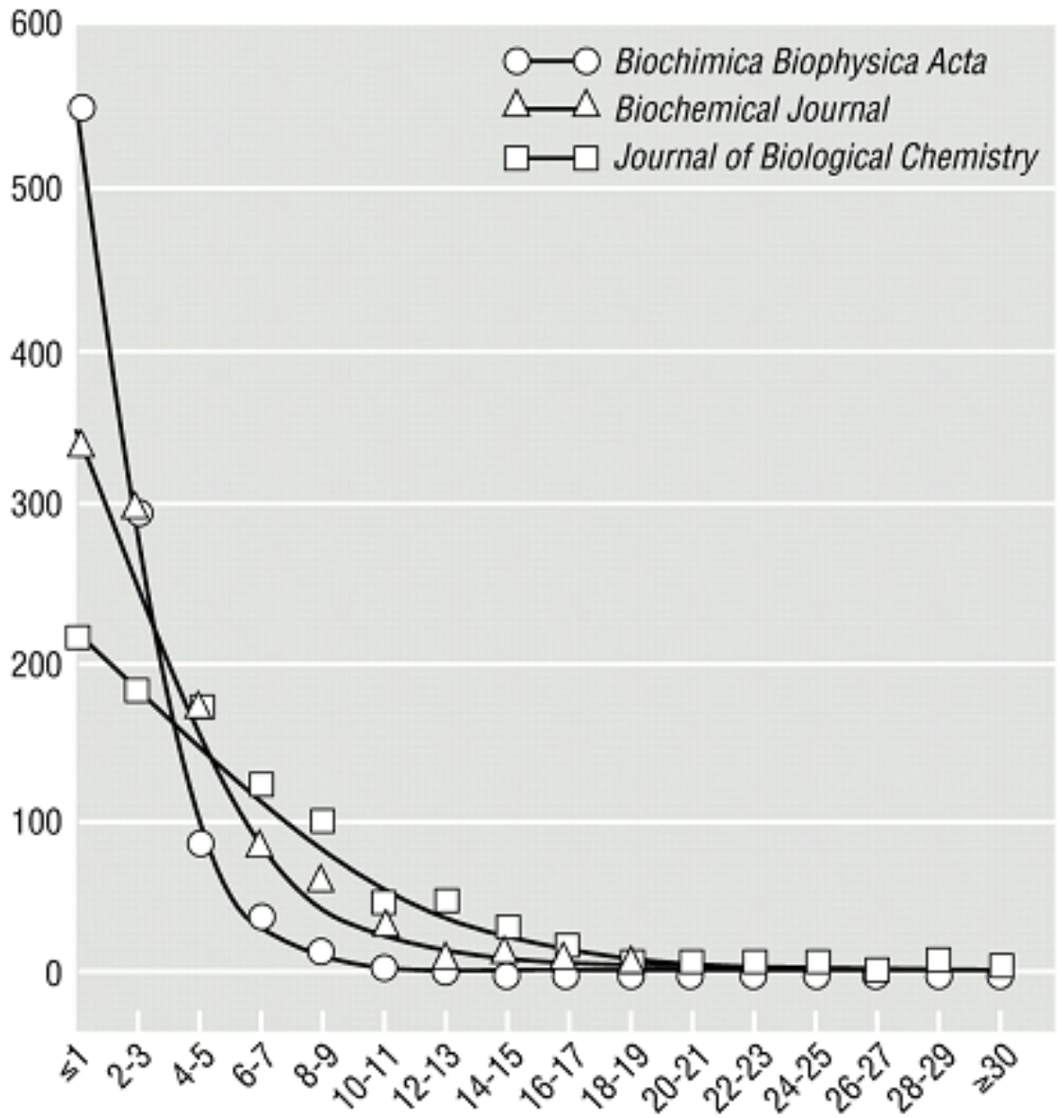
PUBLISH IN HIGH IMPACT JOURNALS OR PERISH

PUBLISH FREQUENTLY IN HIGH IMPACT JOURNALS AND MAYBE YOU WON'T PERISH



facebook.com/pedromics

The IF as a measure of article quality



University rankings

Indicators and Weights for ARWU

Criteria	Indicator	Code	Weight
Quality of Education	Alumni of an institution winning Nobel Prizes and Fields Medals	Alumni	10%
Quality of Faculty	Staff of an institution winning Nobel Prizes and Fields Medals	Award	20%
	Highly cited researchers in 21 broad subject categories	HiCi	20%
Research Output	Papers published in Nature and Science*	N&S	20%
	Papers indexed in Science Citation Index-expanded and Social Science Citation Index	PUB	20%
Per Capita Performance	Per capita academic performance of an institution	PCP	10%
Total			100%

* For institutions specialized in humanities and social sciences such as London School of Economics, N&S is not considered, and the weight of N&S is relocated to other indicators.

RESEARCH ARTICLE

High Impact = High Statistical Standards? Not Necessarily So

Patrizio E. Tressoldi, David Giofré, Francesco Sella, Geoff Cumming

Published: February 13, 2013 • http://dx.doi.org/10.1371/journal.pone.0056180

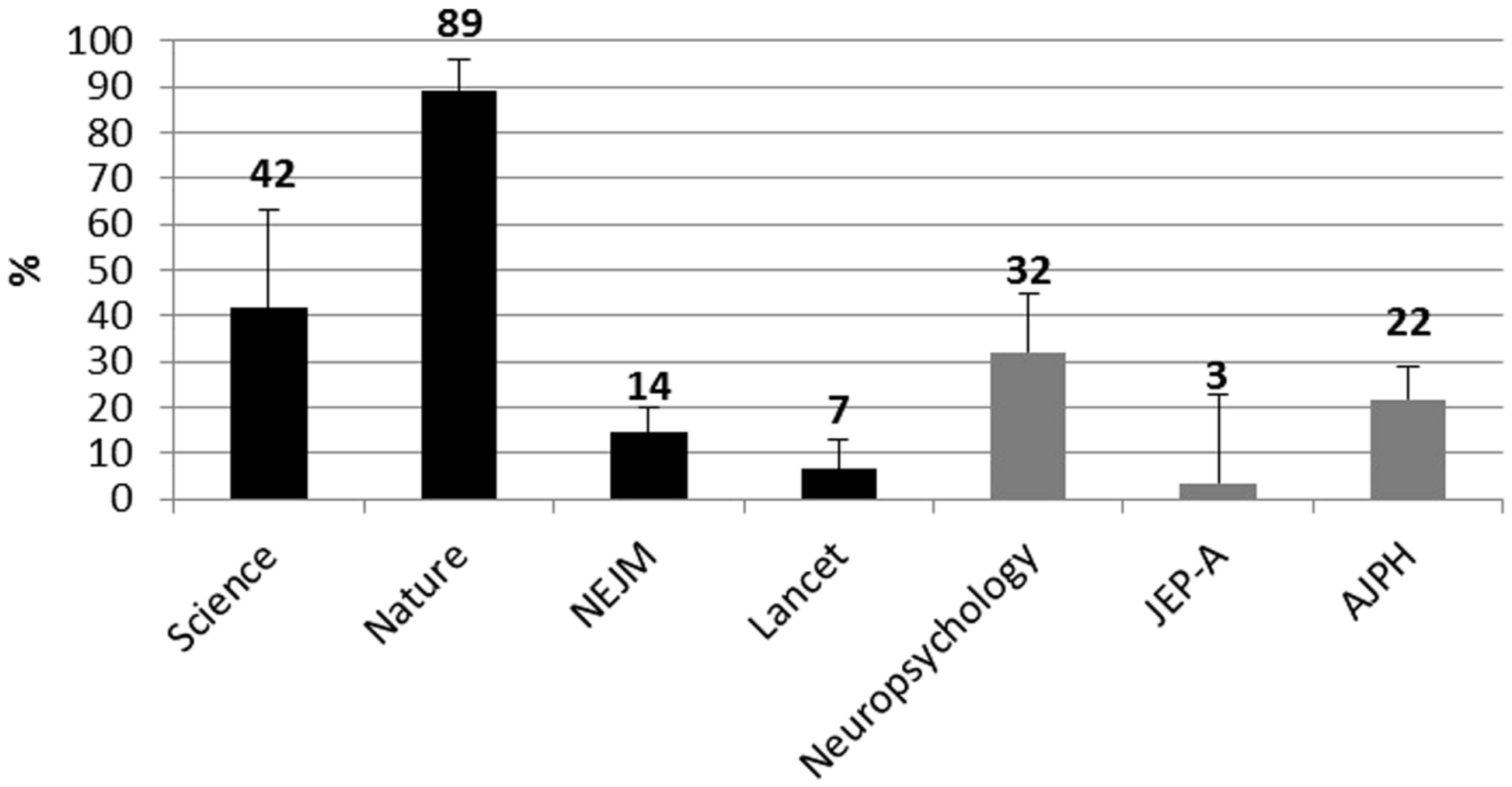
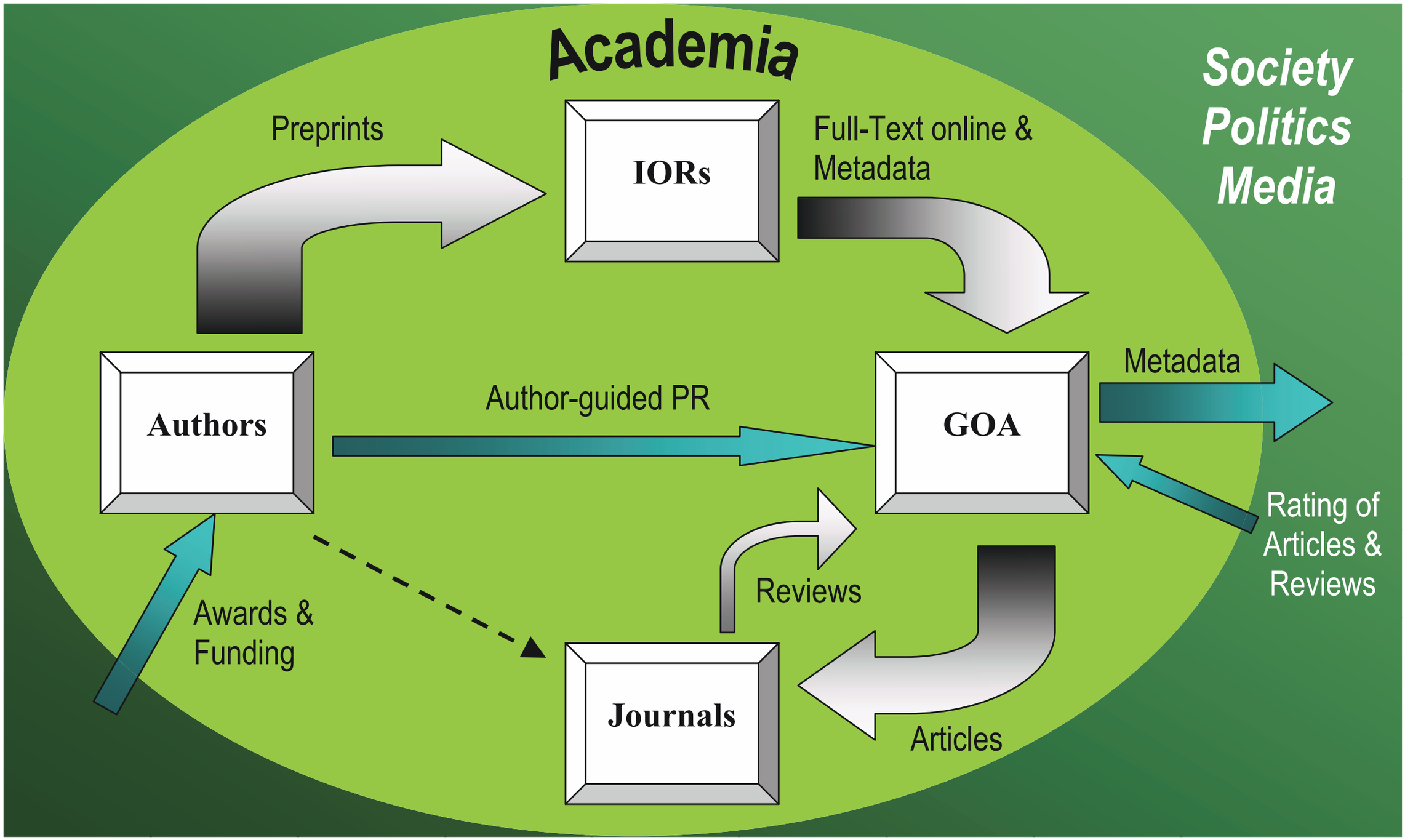
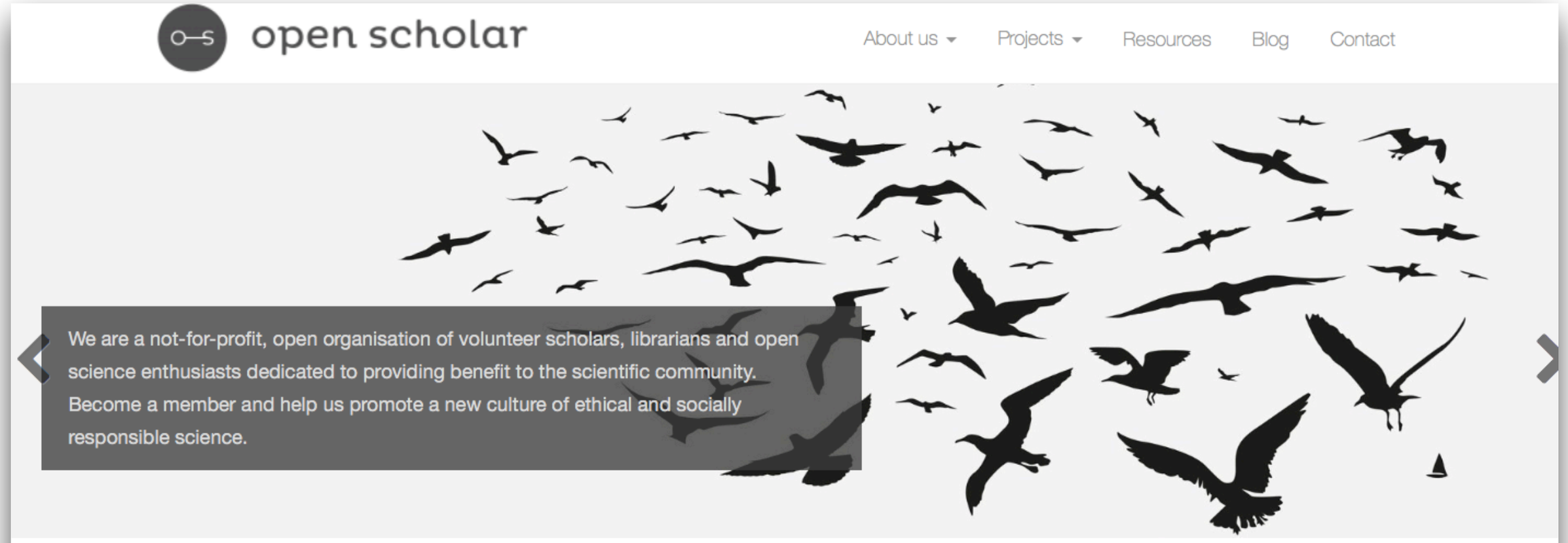


Figure 6. Percentages of selected articles in each journal that used NHST without CI, ES or Model and Power estimation.

VI. Solutions (peer review, preprints, repositories, etc)



Perakakis, P., Taylor, M., Mazza, M. & Trachana, V. Natural selection of academic papers. *Scientometrics* 85, 553–559 (2010).



http://www.openscholar.org.uk

Priority

3 Who?

Prioritize

This article has reached scientific standards 6 /8 Who? Vote

This article still needs revisions 2 /8 Who? Vote

Add to my self-journal

Follow this article

Critiques	50 + 55
Comments	13 + 13
Views	4991
Curators	0

Download the pdf

Supplemental files

All critiques & comments

Cite as Share

Article outline

Abstract

Essay and Opinion

Novel processes and metrics for a scientific evaluation rooted in the principles of science

VERSION 1 Released on 26 January 2017 under Creative Commons Attribution 4.0 International License

Michaël Bon[1], Michael Taylor[2], Gary S McDowell[3,4]

Authors' affiliations

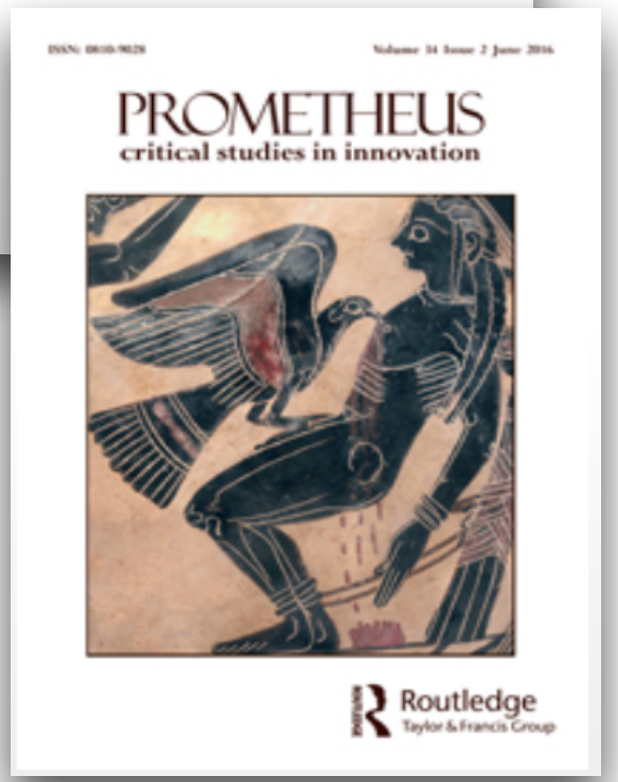
1. SJS - The Self-Journals of Science
2. Department of Physics – Aristotle University of Thessaloniki
3. The Future of Research, Inc.
4. ManyLabs (www.manylabs.org)

http://www.sjscience.org

Academic self-publishing: a not-so-distant future

Pandelis Perakakis & Michael Taylor

Pages 257-263 | Published online: 21 May 2014



Independent Peer Review Manifesto

I endorse the Independent Peer Review Manifesto

[Signature]

Add My Name

http://www.openscholar.org.uk/independent-peer-review-manifesto/

IIIA
Institut d'Investigació en
Intel·ligència Artificial

DIGITAL.CSIC
OPEN SCIENCE

e-IEO

OpenAIRE
Open Access Infrastructure for Research in Europe

open scholar

ARVO CONSULTORES
digital repositories

OPENPEERREVIEW
module
are you an open reviewer?

http://www.openscholar.org.uk/institutional-repositories-start-to-offer-peer-review-services/

http://proyectos.bibliotecas.csic.es/digitalcsic/oprm/index_eng.html

http://proyectos.bibliotecas.csic.es/digitalcsic/workshop_oa_2014/index_eng.html

Title:	**Bibliometrics: The Leiden Manifesto for research metrics**
Authors:	Hicks, Diana; Wouters, Paul; Waltman, Ludo; Rijcke, Sarah de; Rafols, Ismael
Issue Date:	2015
Publisher:	Nature Publishing Group
Citation:	Nature 520(7548): 429-431 (2015)
Abstract:	Use these ten principles to guide research evaluation, urge Diana Hicks, Paul Wouters and colleagues.
Description:	Comment.
Publisher version (URL):	http://dx.doi.org/10.1038/520429a
URI:	http://hdl.handle.net/10261/132304
ISSN:	0028-0836
DOI:	10.1038/520429a
E-ISSN:	1476-4687
Appears in Collections:	(INGENIO) Artículos

Files in This Item:

File	Description	Size	Format	
Leidenmanifesto.pdf		384,04 kB	Adobe PDF	View/Open

Enter your name.

Author *

Last name, e.g. Smith

First name(s) + "Jr", e.g. Donald Jr

Enter your affiliation.

Affiliation *

The reviewer must indicate her affiliation

Select the language of your review.

Language *

No aplicable

Copyright, use and reproduction of your review.

Copyright, use and reproduction. *

Acceso Abierto / Open Access

Link to terms of use and reproduction of your review.

Link to terms of use and reproduction. *

http://creativecommons.org/licenses/by/4.0/

By default, all reviews and comments have a CC-BY license

Resource type

Type *

Revisión

New resource types

Choose option 1 if you consider that the work is scientifically acceptable. Choose option 0 if you consider that the authors should revise the work taking into account your evaluation. The scientific standard refers to various relevant parameters such as methodology, clarity of presentation, use of language, inclusion of key references, soun etc. Choose "No Aplicable" if you do not wish to score these criteria or they are not applicable to the work under review

Scientific standards *

1

Qualitative and quantitative peer review

Rate in a scale from 0-100 the importance of this work for its academic field. [0-100]

Importance of this work for its academic field *

No aplicable

Rate in a scale from 0-100 how interesting this work is for other academic fields. [0-100]

General interest *

No aplicable

Rate in a scale from 0-100 the importance of this work for society in general (social value: how relevant this work is for the problems society is currently facing). [0-100]

Social value *

No aplicable

Please provide below your detailed review about the work, including all necessary information to help its authors improve their contribution. Try to be constructive in your c you need to submit formatted text with figure, equations, etc., you will later have the opportunity to attach one or more additional files.

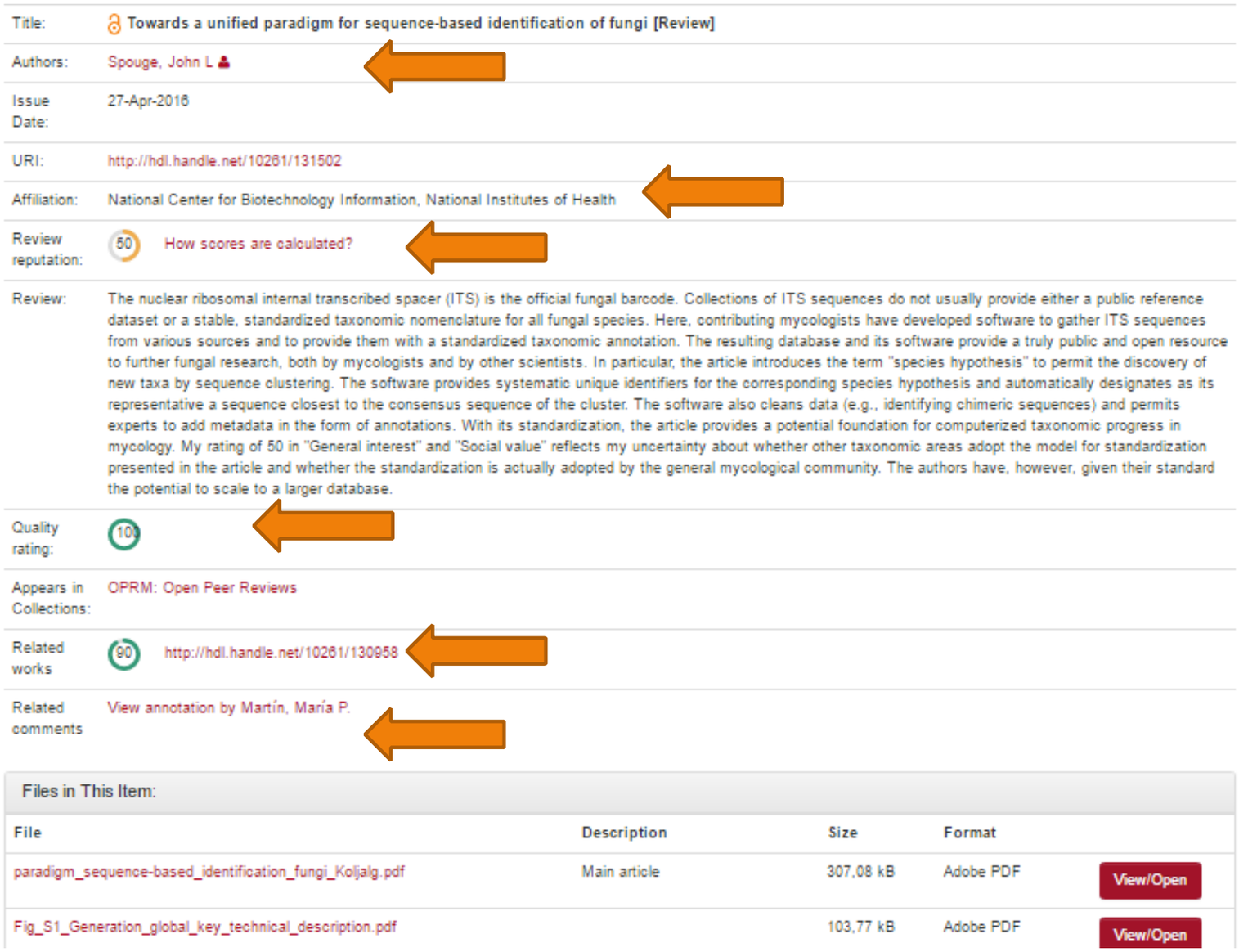
Review text *

Rate in a scale from 0-100 the article globally.

Overall quality assessment *

0

Open reviews and comments generate new items in the repositories



Title:	Towards a unified paradigm for sequence-based identification of fungi [Review]
Authors:	Spouge, John L
Issue Date:	27-Apr-2016
URI:	http://hdl.handle.net/10261/131502
Affiliation:	National Center for Biotechnology Information, National Institutes of Health
Review reputation:	50 How scores are calculated?
Review:	The nuclear ribosomal internal transcribed spacer (ITS) is the official fungal barcode. Collections of ITS sequences do not usually provide either a public reference dataset or a stable, standardized taxonomic nomenclature for all fungal species. Here, contributing mycologists have developed software to gather ITS sequences from various sources and to provide them with a standardized taxonomic annotation. The resulting database and its software provide a truly public and open resource to further fungal research, both by mycologists and by other scientists. In particular, the article introduces the term "species hypothesis" to permit the discovery of new taxa by sequence clustering. The software provides systematic unique identifiers for the corresponding species hypothesis and automatically designates as its representative a sequence closest to the consensus sequence of the cluster. The software also cleans data (e.g., identifying chimeric sequences) and permits experts to add metadata in the form of annotations. With its standardization, the article provides a potential foundation for computerized taxonomic progress in mycology. My rating of 50 in "General interest" and "Social value" reflects my uncertainty about whether other taxonomic areas adopt the model for standardization presented in the article and whether the standardization is actually adopted by the general mycological community. The authors have, however, given their standard the potential to scale to a larger database.
Quality rating:	100
Appears in Collections:	OPRM: Open Peer Reviews
Related works	90 http://hdl.handle.net/10261/130958
Related comments	View annotation by Martín, María P.

Files in This Item:

File	Description	Size	Format	
paradigm_sequence-based_identification_fungi_Koljalg.pdf	Main article	307,08 kB	Adobe PDF	View/Open
Fig_S1_Generation_global_key_technical_description.pdf		103,77 kB	Adobe PDF	View/Open

Open reviews records contain:

- Name of the reviewer and affiliation
- Links to the reviewed work
- Links to items with related open comments
- Individual quality rating of the reviewed work
- Weighed review reputation metrics

https://digital.csic.es/handle/10261/131502

Records of the reviewed works link to their open reviews, individual quality ratings and overall reputation metrics

Cold-water corals in the Cap de Creus canyon, northwestern Mediterranean: spatial distribution, density and anthropogenic impact

View/Open

m397p037.pdf (861.8Kb)

Identifiers

URI: http://hdl.handle.net/10508/7818
DOI: 10.3354/meps08314

Export

RIS

Share

Statistics

1

Cited 73 times in Scopus

View Usage Statistics

Metadata

Show full item record

Authors

Orejas, C. (Covadonga) IEO; Gori, A. (Andrea); Lo-Iacono, C. (Claudio); Puig, P. (Pere); Gili, J.M. (Josep Maria); Dale, M.R.T. (Mark R.T.)

Editor's version

http://www.int-res.com/abstracts/meps/v397/p37-51/

Date

2009

Type

article

Publication reputation

Reputation value

85

Related reviews

90 View review by Cristobo, J. (Javier)

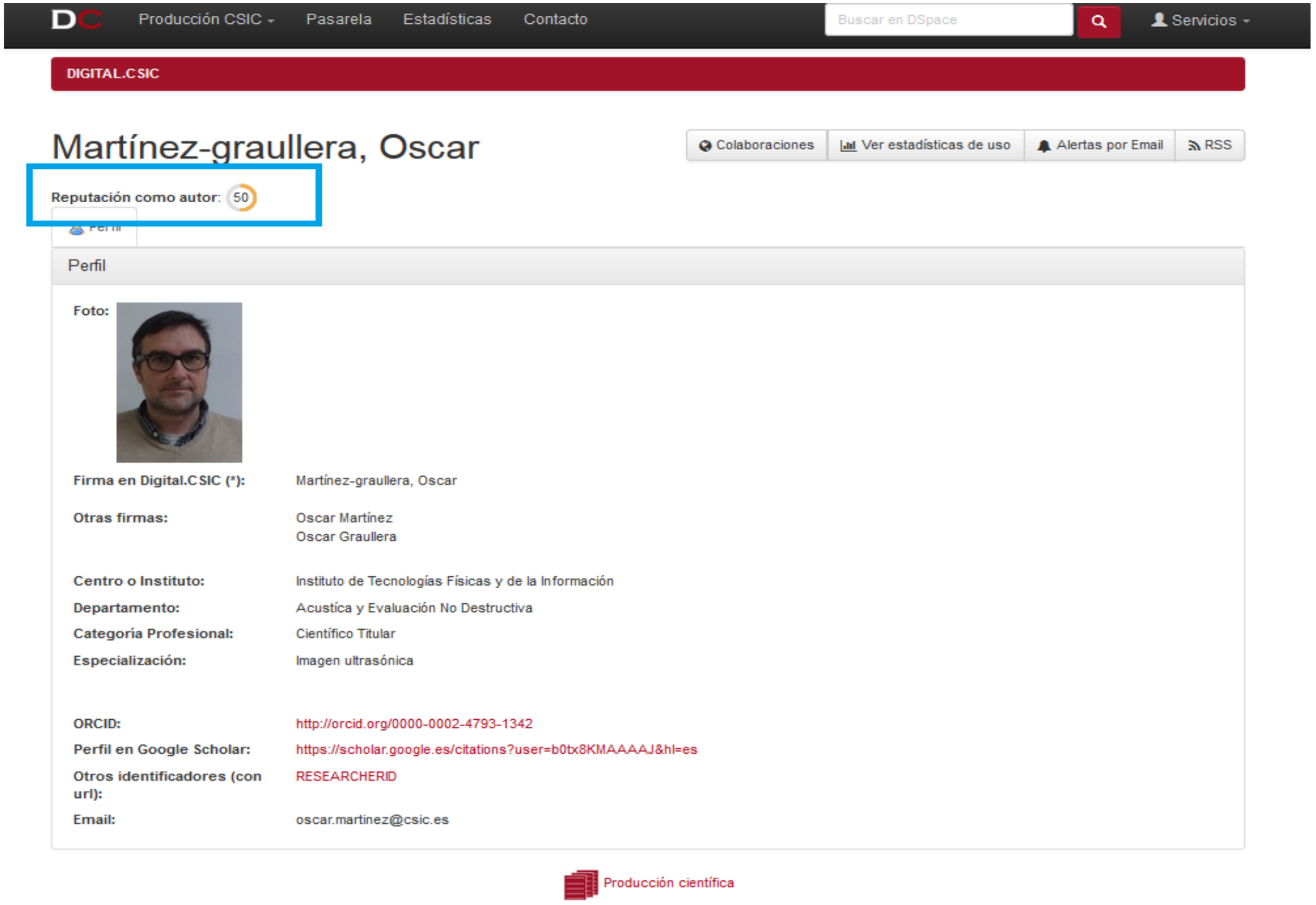
Quality ratings

80 View review by Serrano, A. (Alberto)

Keywords

Cold water corals
Submarine canyons

Author/reviewer reputation metrics show in their personal page



Author/reviewer reputation metrics show in their personal page

Orejas, C. (Covadonga)

Biography

Covadonga Orejas is graduate in
in Natural Sciences from the Univ
staff of the IEO (currently at COB
and other benthic organisms sinc
sustainable management of the C
projects conducted in 12 different
(44 SCI), and 7 book chapters. Sh
scientific journals, and several fur

Download in PDF

Download in CVN fragment

Contact Information

Fields of Specialization

Biology and ecology of cold-water corals (CWC) and other benthic org

Degrees

Graduate in Biology from the Universidad Complutense de Madrid (Sp
(Germany)

Departments

C.O. BALEARES

Last updated marzo 15, 2016

Author's reputation

79

Jerez, S. (Salvador)

Biography

Salvador Jerez, born in La Gomera (Spain)
of La Laguna (Spain). Researcher at the IE
Island Oceanographic Centre since 1989. E
with 20 years participating in 19 Spanish an
included in JCR/SC, and 40 National and In
greater amberjack in the European Project

Download in PDF

Download in CVN fragment

Contact Information

Fields of Specialization

Expert on Greater amberjack (Seriola dumerili) aquaculture specializing in reproc

Degrees

Graduate and PhD in Biology from the University of La Laguna (Spain)

Departments

C.O. CANARIAS

Last updated marzo 15, 2016

Reputation as a reviewer

86

And new exciting projects on their way!…

VII. Discussion