Reproducible and Responsible Research at JEPS

Lea Jakob, Fabian Dablander, Martin Holst, Laura Van Heck, Helena Eidast

31st EFPSA Congress
Qakh, Azerbaijan

Outline

- Brief History of the Replication Crisis
- How to tackle the Replication Crisis
- How to publish in JEPS
- How you can help us promote JEPS and advance Psychological Science

Brief history of the Replication Crisis

Replication Crisis

- 2011 was a “year of horrors” (Wagenmakers, 2012)
  - Daryl Bem feeling the future paper (2011)
  - False-positive psychology paper (2011)
  - Questionable Research Practices paper (2012)

Replication Crisis

FEATURES

Text Size: A | A | A

53.1% of You Already Know What This Story's About. Or Do You? Need a Hint?

It's about Professor Daryl Bem and his cheerful case for ESP.

By Dan Kois Published Feb 27, 2011 Share

Journal of Personality and Social Psychology
2011, Vol. 100, No. 3, 407–425

0022-3514/11/$12.00 DOI: 10.1037/a0021524

Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect

Daryl J. Bem
Cornell University

The term *psi* denotes anomalous processes of information or energy transfer that are currently unexplained in terms of known physical or biological mechanisms. Two variants of psi are *precognition* (conscious cognitive awareness) and *premonition* (affective apprehension) of a future event that could not otherwise be anticipated through any known inferential process. Precognition and premonition are themselves special cases of a more general phenomenon: the anomalous retroactive influence of some future event on an individual's current responses, whether those responses are conscious or nonconscious, cognitive or affective. This article reports 9 experiments, involving more than 1,000 participants, that test for retroactive influence by "time-reversing" well-established psychological effects so that the individual's responses are obtained before the putatively causal stimulus events occur. Data are presented for 4 time-reversed effects: precognitive approach to erotic stimuli and precognitive avoidance of negative stimuli; retroactive priming; retroactive habituation; and retroactive facilitation of recall. The mean effect size (*d*) in psi performance across all 9 experiments was 0.22, and all but one of the experiments yielded statistically significant results. The individual-difference variable of stimulus seeking, a component of extraversion, was significantly correlated with psi performance in 5 of the experiments, with participants who scored above the midpoint on a scale of stimulus seeking achieving a mean effect size of 0.43. Skepticism about psi, issues of replication, and theories of psi are also discussed.

Keywords: psi, parapsychology, ESP, precognition, retrocausation

A recent meta-analysis of 90 experiments on precognition yielded overwhelming evidence in favor of an effect (Bem et al., 2015). Alan Turing, discussing research on psi related phenomena, famously stated that

> *"These disturbing phenomena seem to deny all our usual scientific ideas. How we should like to discredit them! Unfortunately, the statistical evidence, at least of telepathy, is overwhelming." (Turing, 1950, p. 453; cf. Wagenmakers et al., 2015)*

From Dablander (2015)

The culprit: Confirmation bias

- A (failed) replication of Bem’s results was not accepted for JPSP

- “This journal does not publish replication studies, whether successful or unsuccessful.” - JPSP Editor at the time

Daryl Bem giving advice ...

Planning Your Article

Which Article Should You Write?

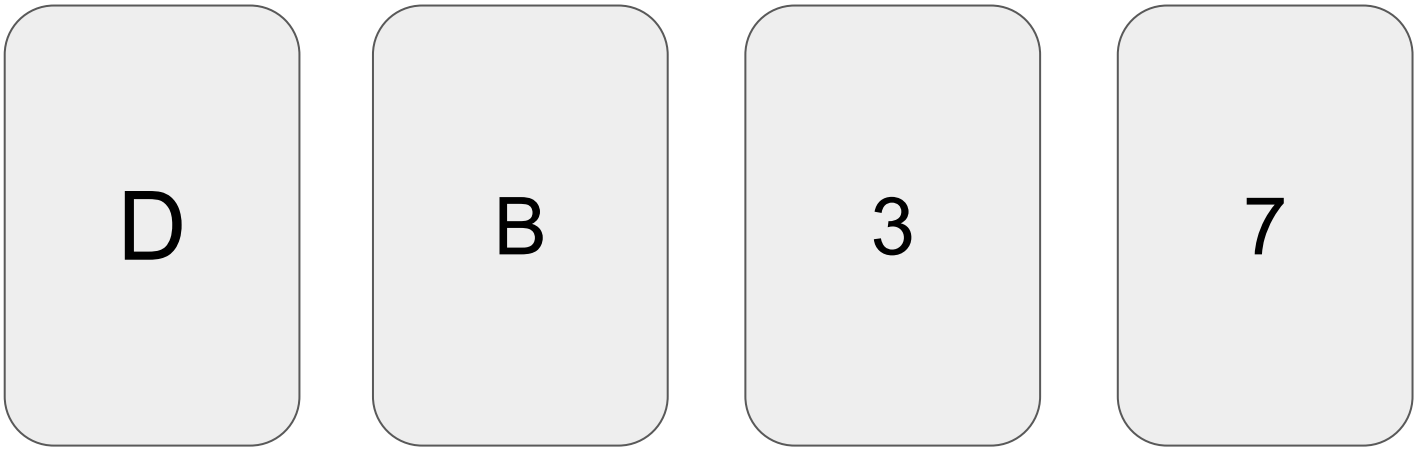
There are two possible articles you can write: (a) the article you planned to write when you designed your study or (b) the article that makes the most sense now that you have seen the results. They are rarely the same, and the correct answer is (b).

From Bem (1987)

Publication bias

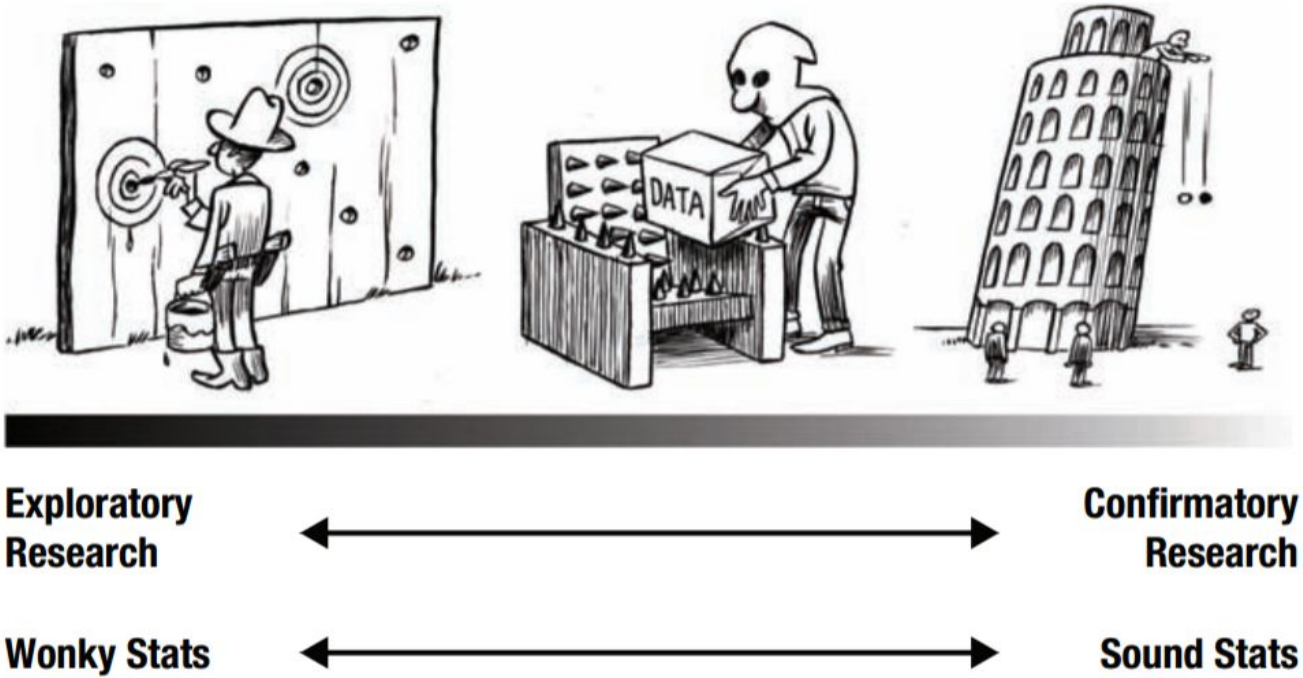
- **Publication bias** is an insidious form of confirmation bias
- Results **“in a literature that is about as representative of real science as porn movies are representative of real sex”** (Daniel Lakens)
- Can never falsify a hypothesis
- Is the greatest enemy of sound science

If D, then 3



“The first principle [of science] is that you must not fool yourself — and you are the easiest person to fool.”

- Richard Feynman



From Wagenmakers et al. (2012)

General Article

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

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

Psychological Science
XX(X) 1–8

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

Abstract

In this article, we accomplish two things. First, we show that despite empirical psychologists' nominal endorsement of a low rate of false-positive findings (≤ .05), flexibility in data collection, analysis, and reporting dramatically increases actual false-positive rates. In many cases, a researcher is more likely to falsely find evidence that an effect exists than to correctly find evidence that it does not. We present computer simulations and a pair of actual experiments that demonstrate how unacceptably easy it is to accumulate (and report) statistically significant evidence for a false hypothesis. Second, we suggest a simple, low-cost, and straightforwardly effective disclosure-based solution to this problem. The solution involves six concrete requirements for authors and four guidelines for reviewers, all of which impose a minimal burden on the publication process.

Table 1. Likelihood of Obtaining a False-Positive Result

Researcher degrees of freedom	Significance level		
	$p < .1$	$p < .05$	$p < .01$
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%

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)

From John et al. (2012)

RESEARCH ARTICLE

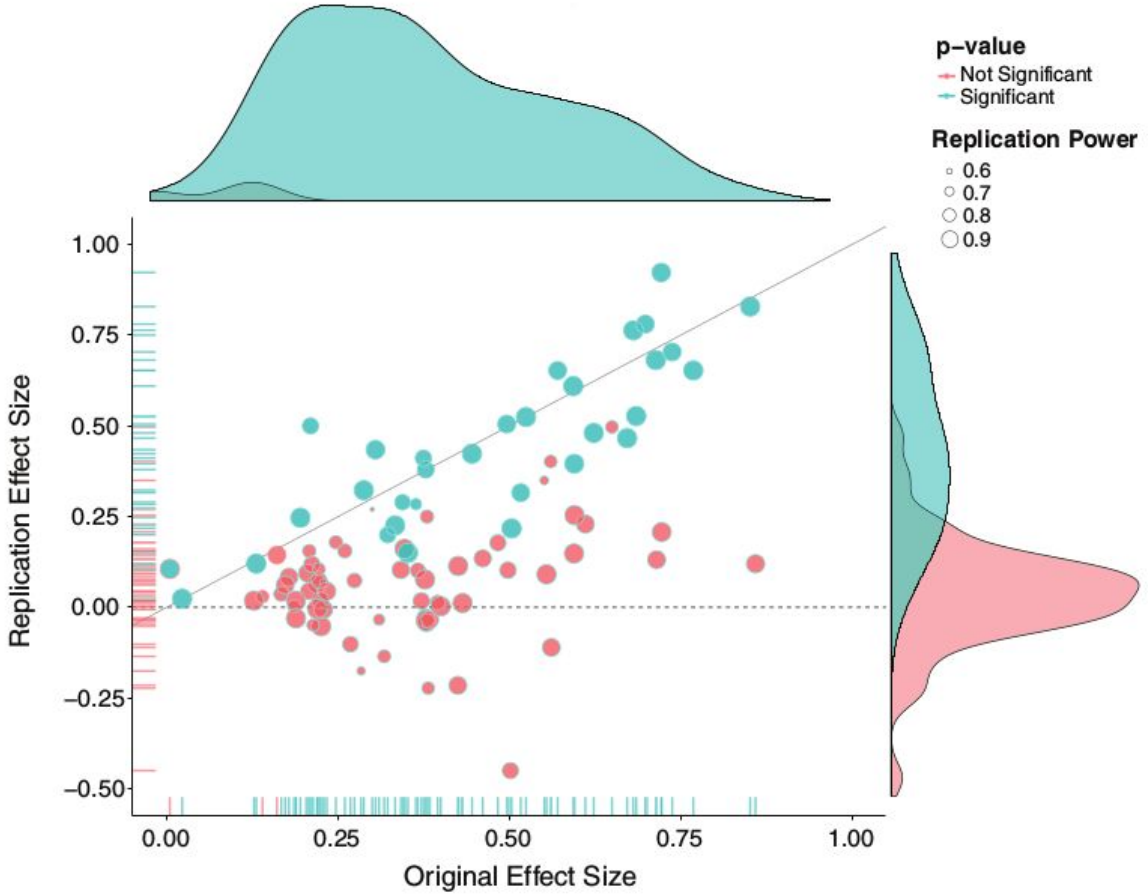
PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*†

Reproducibility is a defining feature of science, but the extent to which it characterizes current research is unknown. We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original materials when available. Replication effects were half the magnitude of original effects, representing a substantial decline. Ninety-seven percent of original studies had statistically significant results. Thirty-six percent of replications had statistically significant results; 47%

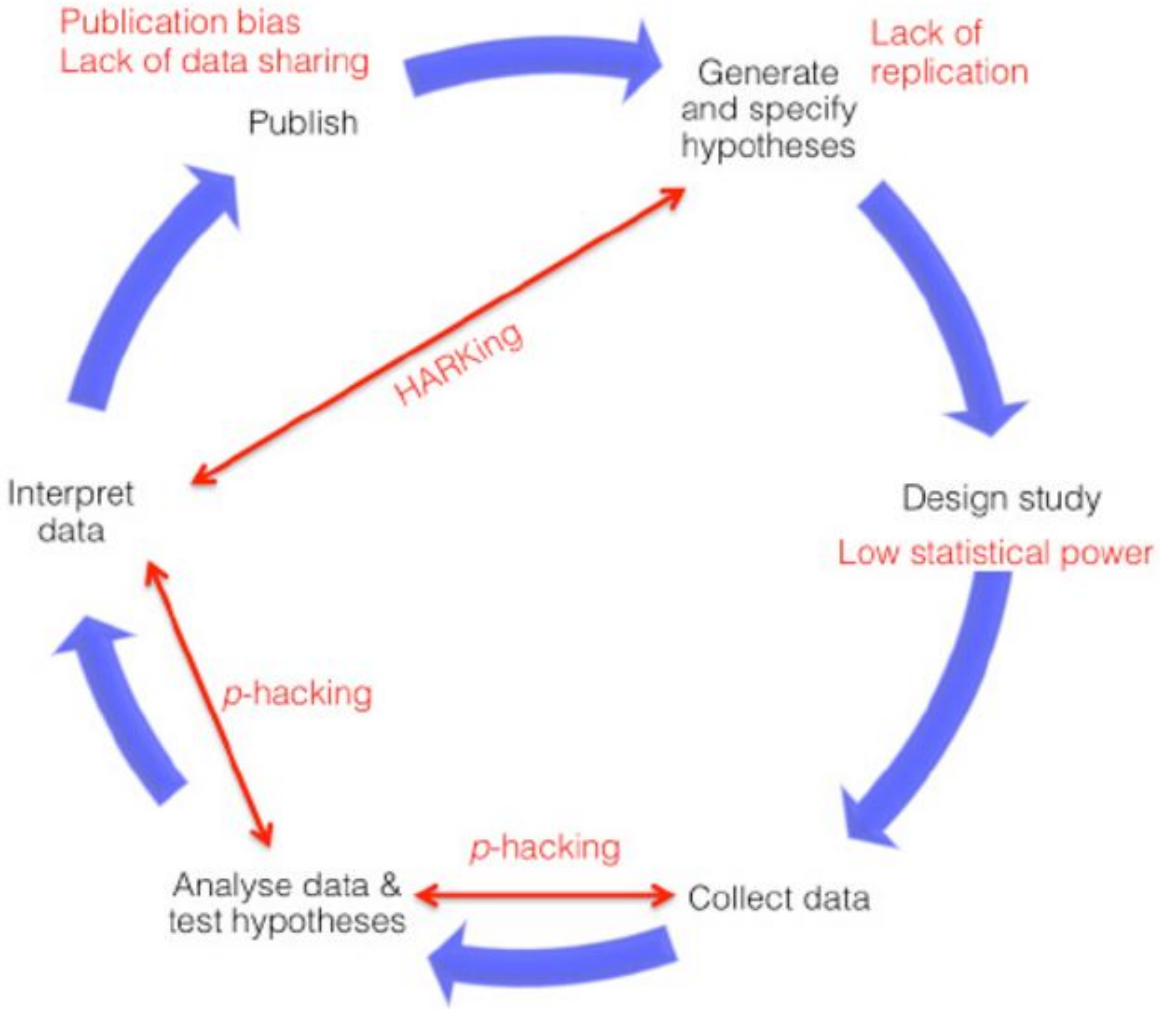
From Open Science Collaboration (2015)



From Open Science Collaboration (2015)

Fig. 3. 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.

Current state



From Munafo et al. (2017)

If you want to know more ...

- Solidify your understanding
  - Read our blog series on Registered Reports (here)
  - Read “Registered Reports for Student Research” (JEPS Editorial Team, 2015)
  - Read “An Agenda for Purely Confirmatory Research” (Wagenmakers et al., 2012)

- Update your statistical knowledge using the non-technical, free Coursera course “Improving your statistical Inference” by Daniel Lakens

- Read “The 7 Deadly Sins of Psychology” by Chris Chambers

- Hop onto twitter and follow open science people (start here)

Some promising solutions

Some promising solutions

- Registered Reports
- Open Access
- Publish everything
- Open Data
- Altmetrics

Registered Reports (RR)

- **Problem**: Everything we just talked about
- **Solution**: Specifying methods and analyses before actually collecting data
- JEPS accepts RR and guarantees publication when carried out

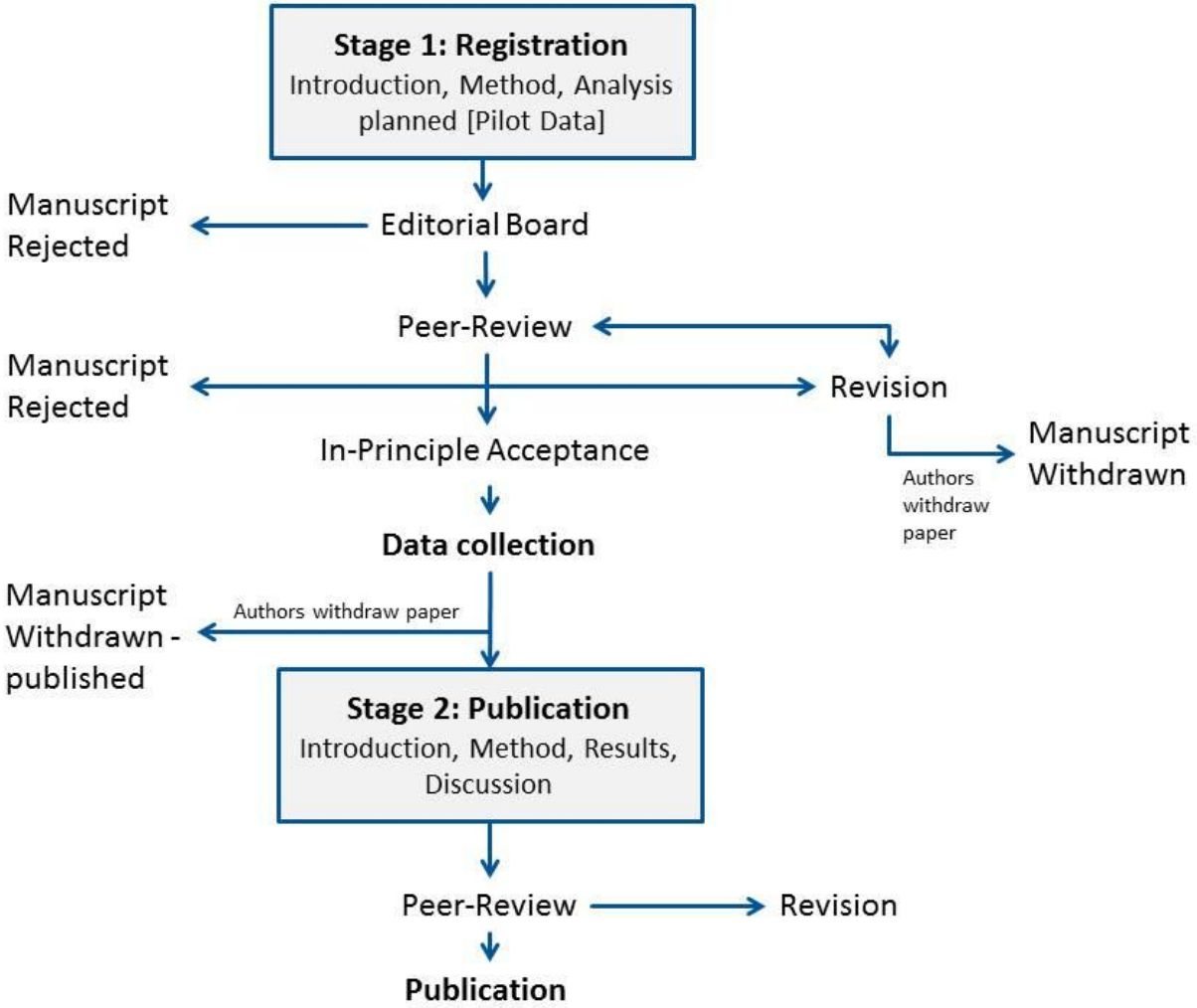
Letter from the Editors

Registered Reports for Student Research

Authors: Maedbh King, Fabian Dablander, Lea Jakob, Maria L. F. Agan, Felicitas Huber, Jonas M. B. Haslbeck, Katharina F. Brecht

Abstract

The pre-registration of research via registered reports is a recent development in the field of psychology. The aim of pre-registration is to encourage research that presents sound hypotheses and methodology (Chambers, 2014) in order to counter undesirable but prevalent research practices such as cherry-picking and p-hacking. In this Letter from the Editors, we wish to echo calls for registered reports and outline how we, the Editors at the Journal of European Psychology Students (JEPS), plan to introduce registered reports for student research. We address the issues necessitating the introduction of registered reports and outline the approach needed for implementing this initiative in a student journal.


Stage 1: Registration
Introduction, Method, Analysis planned [Pilot Data]
Manuscript Rejected
Editorial Board
Peer-Review
Manuscript Rejected
Revision
Manuscript Withdrawn
Authors withdraw paper
In-Principle Acceptance
Data collection
Manuscript Withdrawn - published
Authors withdraw paper
Stage 2: Publication
Introduction, Method, Results, Discussion
Peer-Review
Revision
Publication

Open Access (OA)

- **Problem**: Not all literature is available to everyone
- **Solution**: Green OA, Gold OA
- JEPS is a gold open access journal and we usually waive the publication fee

Publishing Everything

- **Problem**: Publication bias
  - Results **“in a literature that is about as representative of real science as porn movies are representative of real sex”** (Daniel Lakens)

- **Solution**: Publishing everything (regardless of *p*-value)

- JEPS does guarantee publishing after RR, regardless of the actual results

Open Data

- **Problem**: Most research is not reproducible

- **Solution**: Uploading all data to an online database (+ analysis scripts), accessible by everybody
  - Ensures transparency
  - Makes reproducibility and replication possible

- JEPS requires researchers to upload their data to Open Science Framework

Altmetrics

- **Problem**: Impact factor as a measure for the quality of a journal
- **Solution**: Altmetrics
- JEPS makes use of *amount of views, downloads* and *mentions* on Twitter

How to publish in JEPS

JEPS Policy

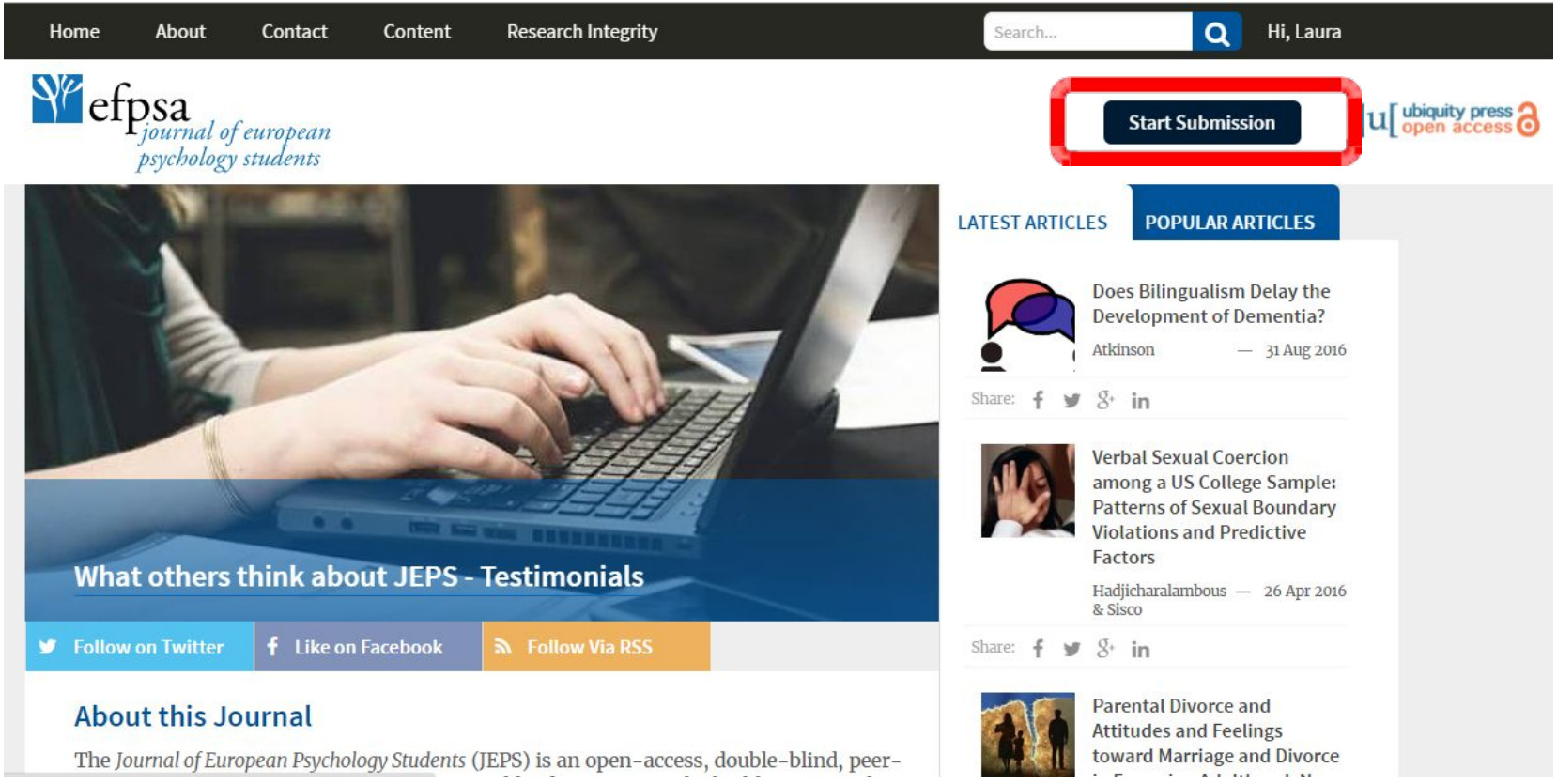
- Double-blind peer-review, Open Access, Registered Reports, etc.

- Our goal is to help psychology students
  - Waive the publishing fees
  - Provide extensive feedback authors
  - Promote best practices

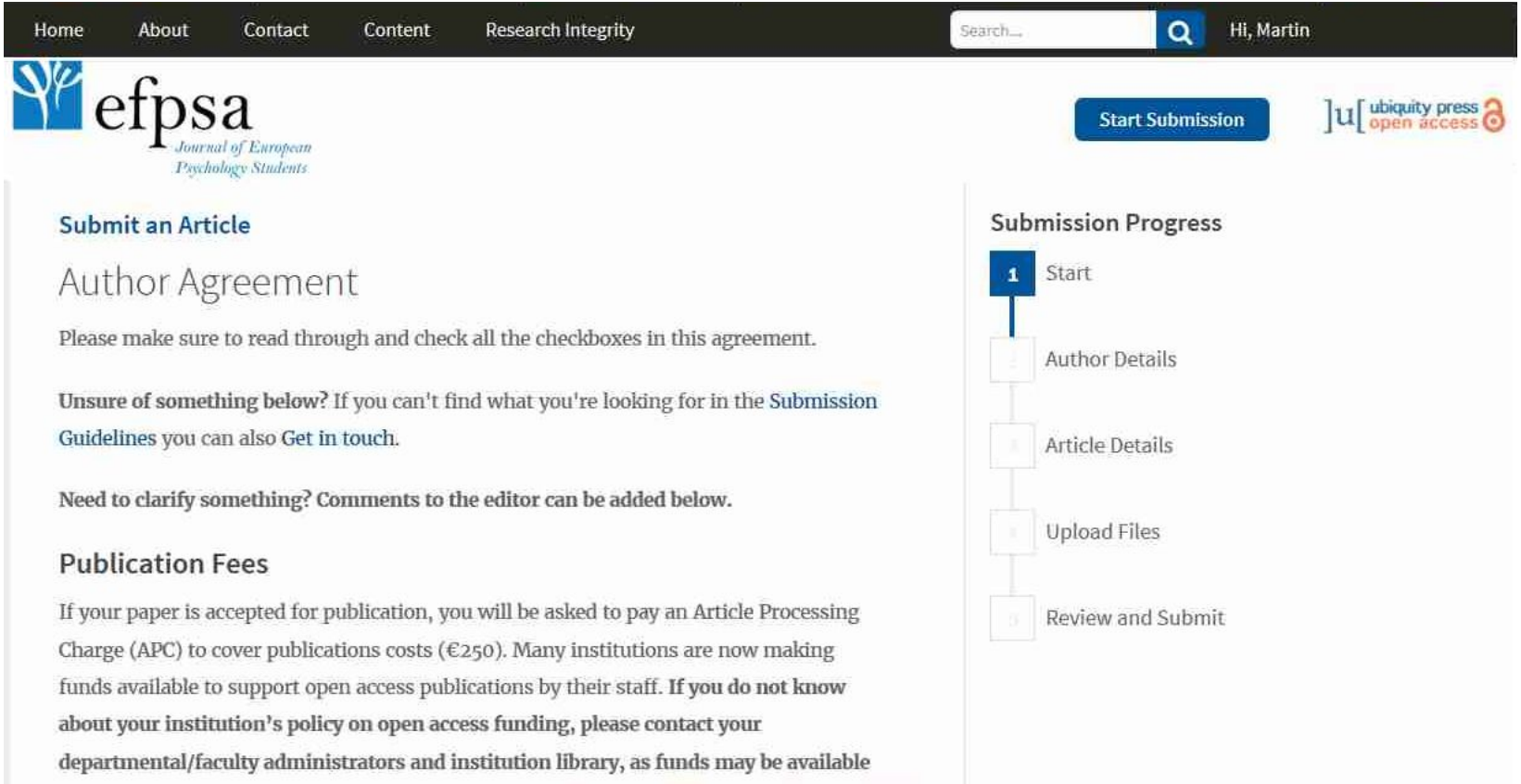
What can you send to us?

- Literature Review
- Research Article
- Registered Report

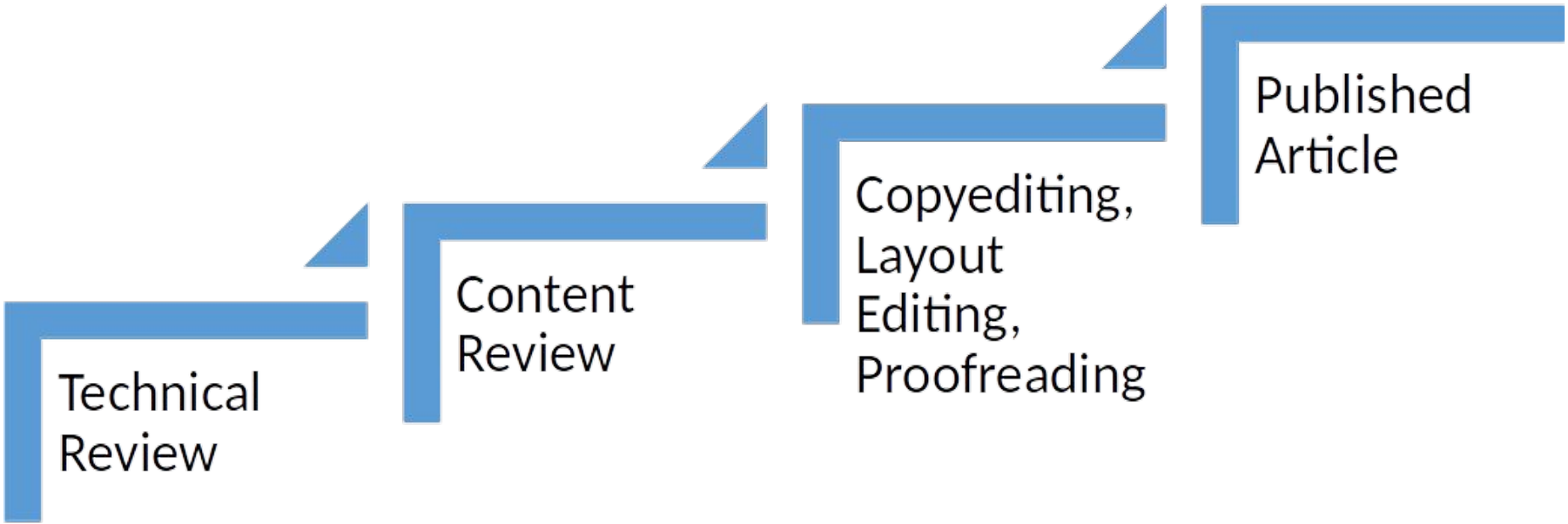
jeps.efpsa.org



jeps.efpsa.org

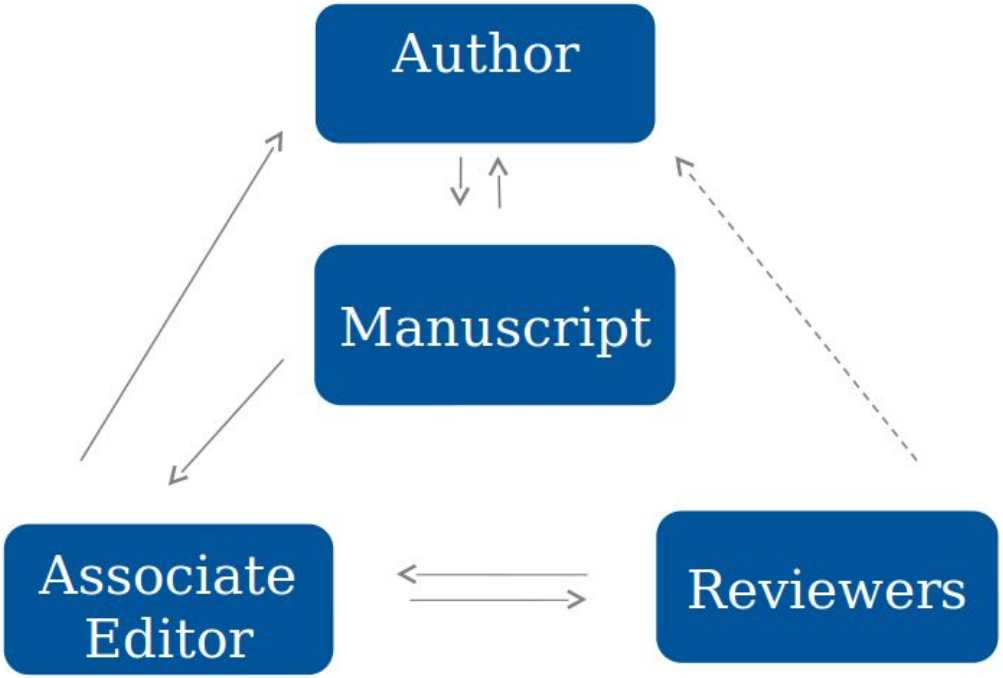


The Publication Process

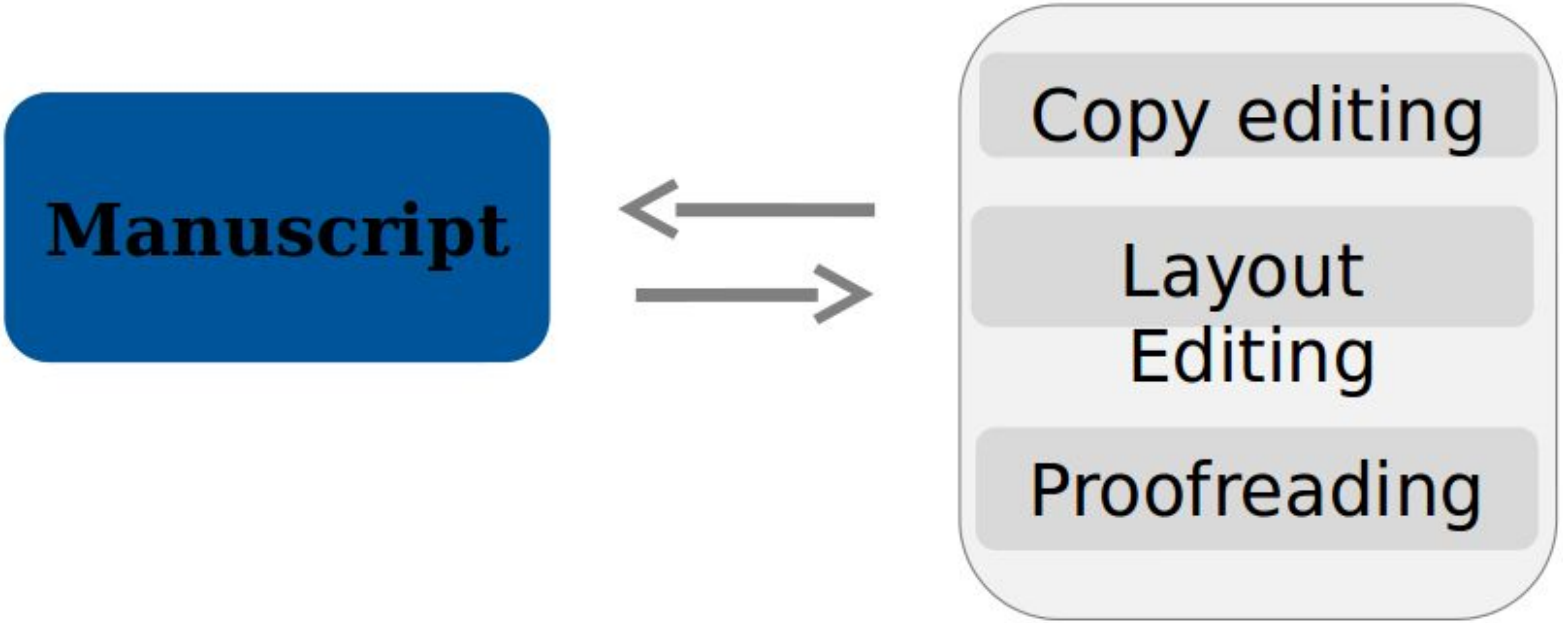


Technical Review

Content Review



Copyediting, Layout, Editing, Proofreading



Publication!

efpsa journal of european psychology students

Lange, F. et al (2014). Turn It All You Want: Still No Effect of Sugar Consumption on Ego Depletion. *Journal of European Psychology Students*, *5*(3), 1-8, DOI: http://dx.doi.org/10.5334/jeps.cc

RESEARCH ARTICLE

Turn It All You Want: Still No Effect of Sugar Consumption on Ego Depletion

Florian Lange[*,†], Caroline Seer[†,‡], Mariça Rapior[*], Jan Rose[*] and Frank Eggert[*]

After having completed an initial self-control task, individuals typically show less self-controlled behavior on a consecutive task. In addition, this so-called ego-depletion effect is assumed to be alleviated by the consumption of sugar-containing drinks. However, a recent replication study indicates that this effect has been substantially overestimated. In contrast to mainstream ego-depletion research, initial and consecutive self-control tasks were identical in that study. Here we evaluate the generalizability of these results by testing 70 participants on a dual-task paradigm involving dissimilar tasks. Between self-control tasks, participants consumed a drink containing either sugar or an artificial sweetener. Results suggest that sugar consumption does not counteract ego depletion even when dissimilar self-control tasks are used.

Keywords: self-control; ego depletion; sugar consumption; replication; strength model

Following initial demonstrations of ego depletion by Baumeister and colleagues (Baumeister, Bratlavsky, Muraven, & Tice, 1998; Muraven, Baumeister, & Tice, 1999) a large number of empirical studies have generated evidence in support of a folk psychological notion: the human capacity to override habitual responses, impulses, thoughts and emotions is limited. According to glucose and that supplementing glucose, but not artificial sweetener, counteracts the ego-depletion effect. This latter finding has been replicated several times (DeWall, Baumeister, Gailliot, & Maner, 2008; Gailliot, Peruche, Plant, & Baumeister, 2009; Masicampo & Baumeister, 2008; Wang & Dvorak, 2010). The meta-analysis by Hagger et al. (2010) yielded a large moderating effect ($d = 0.75$) of

Video outlining these steps

(https://www.youtube.com/watch?v=I6thhDmnAbY)

How you can help us promote JEPS and advance Psychological Science

You can ...

… be a JEPS ambassador

… promote us on EFPSA Day

… print our posters and post them everywhere

… submit your thesis and tell everyone about the experience

… write an article about us for a newsletter

Social Media

Facebook Group

Question and Answers

Thanks for your attention.