Open up for Open Science

Jonas Kubilius

Brain & Cognition / KU Leuven / LIPS / 2015-03-13

Introduction

in which everything goes wrong

	A	B	C	D	E	F	G
2			Real GDP growth				30 or less
3			Debt/GDP				
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or abov	30 or less
26			3.7	3	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	195l-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	-0.1	6.4
52							
53							

	A	B	C	D	E	F	G
			Real GDP growth				
2							
3			Debt/GDP				
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3.7	3	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	195l-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	=AVERAGE(F30:F44)	
52							
53							

The mistake(s)

Herndon et al. (2014) (pdf here):

- Positive growth actually (corrected from -0.1% to 0.3%)
- Unconventional analysis methods, some countries omitted (GDP growth up to 2.2%)
- Correlation does not imply causation

A perfect example because it exposes problems plaguing the traditional way of doing science:

- Serious analysis done by poor tools and not validated
- No peer review -- example of elitism and reputation-based science
- Not sharing analyses
- Not accepting mistakes
- Scientific claims exaggerated by politicians

Chapter 1

where somebody is in a much too comfortable position

The role of a journal
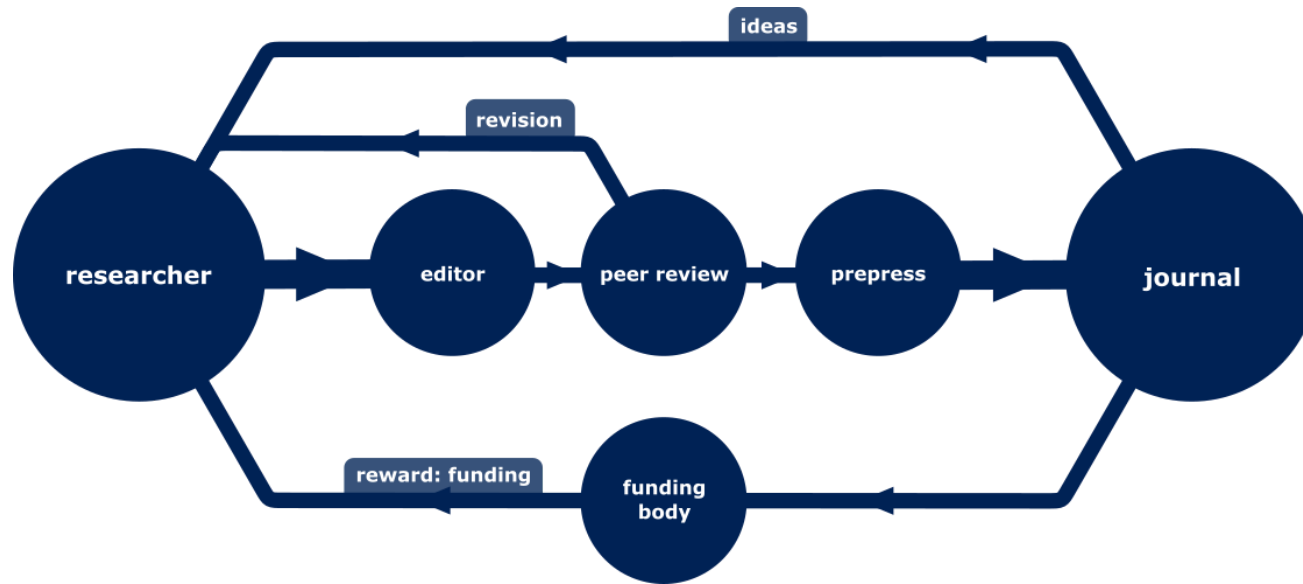
Establishing the quality of research:

- Tool for disseminating research
- Quality control
- Researcher reputation based in impact factor
- Formatting (prepress services)

Impact Factor:

- Average number of citations per article in two years.

A neat business model
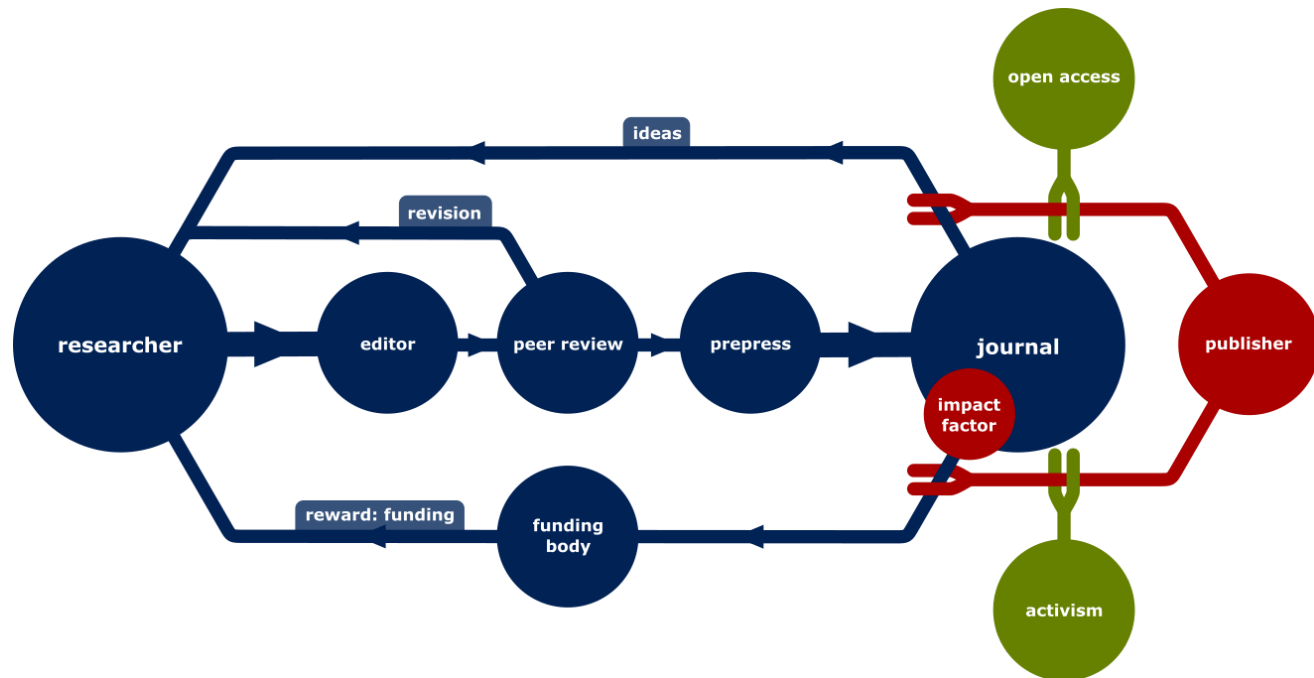
Academic publishing in practice

Cortex

Available online 10 February 2015

In Press, Uncorrected Proof — Note to users

ELSEVIER

Cortex

Special issue: Review

Brain-decoding fMRI reveals how wholes relate to the sum of parts

Jonas Kubilius[a, b], Annelies Baeck[a, b], Johan Wagemans[b], Hans P. Op de Beeck[a],

+ Show more

Choose an option to locate/access this article:

Check if you have access through your login credentials or your institution

Check access

📄 Purchase $35.95

Get Full Text Elsewhere

Get rights and content

Profits

How much do universities pay for it?

- Cambridge (*Elsevier* alone): £1,161,571
- Harvard (total budget): $3.75m

Profits

How much do universities pay for it?

- Cambridge (*Elsevier* alone): £1,161,571
- Harvard (total budget): $3.75m

Profit margins:

- Springer: 34%
- Elsevier: 36%
- Wiley: 40%

Profits

How much do universities pay for it?

- Cambridge (*Elsevier* alone): £1,161,571
- Harvard (total budget): $3.75m

Profit margins:

- Springer: 34%
- Elsevier: 36%
- Wiley: 40%

David Hughes   Follow
@usernameerror

Roses are red
Violets are blue
The rest of this poem
Will cost $31.52

#ElsevierValentines

1:36 PM - 10 Feb 2015

393 RETWEETS   **246** FAVORITES

Yann LeCun

Where has the academic publishing world come to?
#openaccess

Audrey Watters originally shared:
More pressure on academia to be more open -- a response to Aaron
Swartz's indictment.

33 GB of Scientific Papers - and a Manifesto
- Uploaded to Pirate Bay

readwriteweb.com

3 komentarai apie pradinį įrašą

+3   ➔4

Open Access

So let's have it open.

KU Leuven takes Open Access seriously:

- Lirias
- Resources

Doesn't work well yet

- Green open access: self-archiving of preprint or postprint (free)
- Golden open access: publisher's version freely available (€€€)

Check conditions at the SHERPA RoMEO website.

Meaning: pay twice (once for publishing, once for subscription)

Chapter 2

where you learn how to turn on a computer

Reproducibility crisis

- Ionidis (2005): *"Why Most Published Research Findings Are False."*
- Reproducibility rates in drug development:
    - 6% (Amgen / Begley & Ellis, *Nature*, 2012)
    - 20-25% (Bayer HealtCare / Prinz et al., *Nature Rewievs Drug Discovery*, 2011)
- *"Many Labs" Replication Project:* 10 out of 13 Social Psychology studies replicated successfully (Klein et al., 2012)
- *Reproducibility Project: Psychology:* 11 out of 29 Experimental Psychology studies replicated successfully

(See more at Nature's Special on Challenges in Irreproducible Research and Science's Special Section on Data Replication & Reproducibility)

Reproducibility Project

"Reproduced Quadrant"

r = .62

Gilbert (2014)
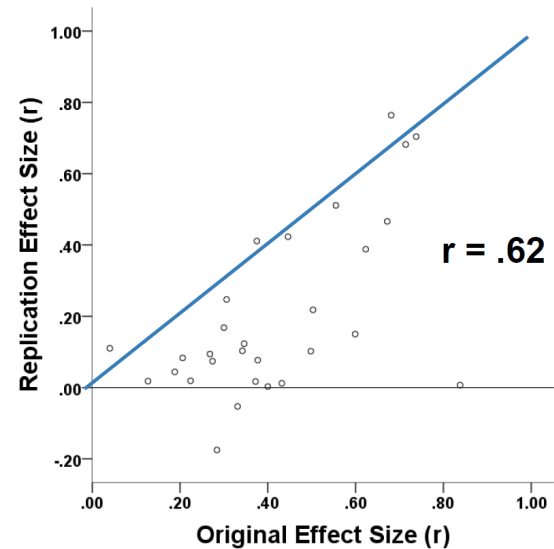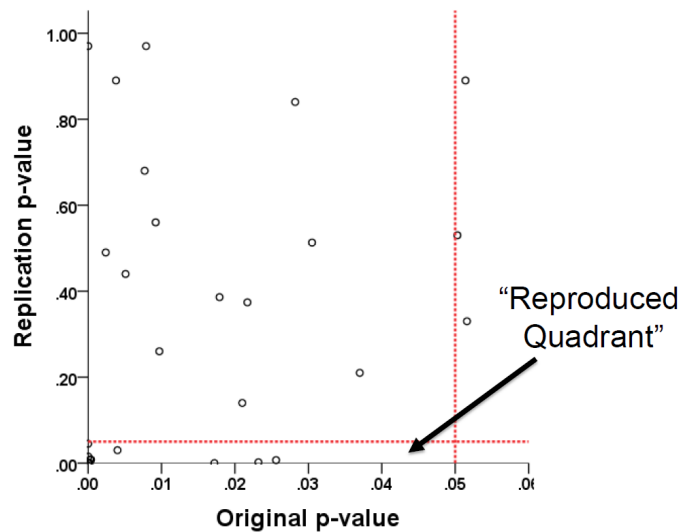
What would Newton say?

Reproducibility: What's up?
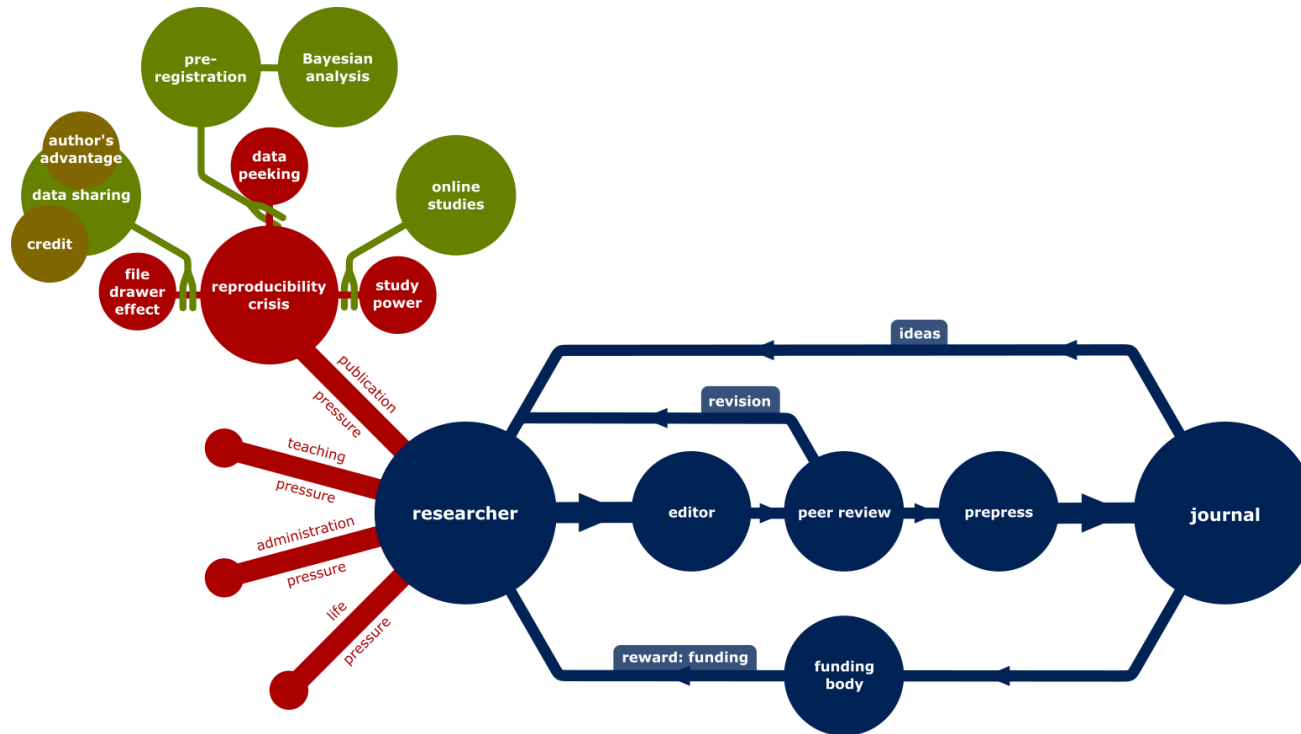
Publication pressure:

- Underpowered studies -- so use more participants!
- Data peeking -- so pre-register your studies
- File-drawer effect -- so publish your findings anyway + journals should change (see PsychFileDrawer)
- Nobody cares

We don't understand p-values
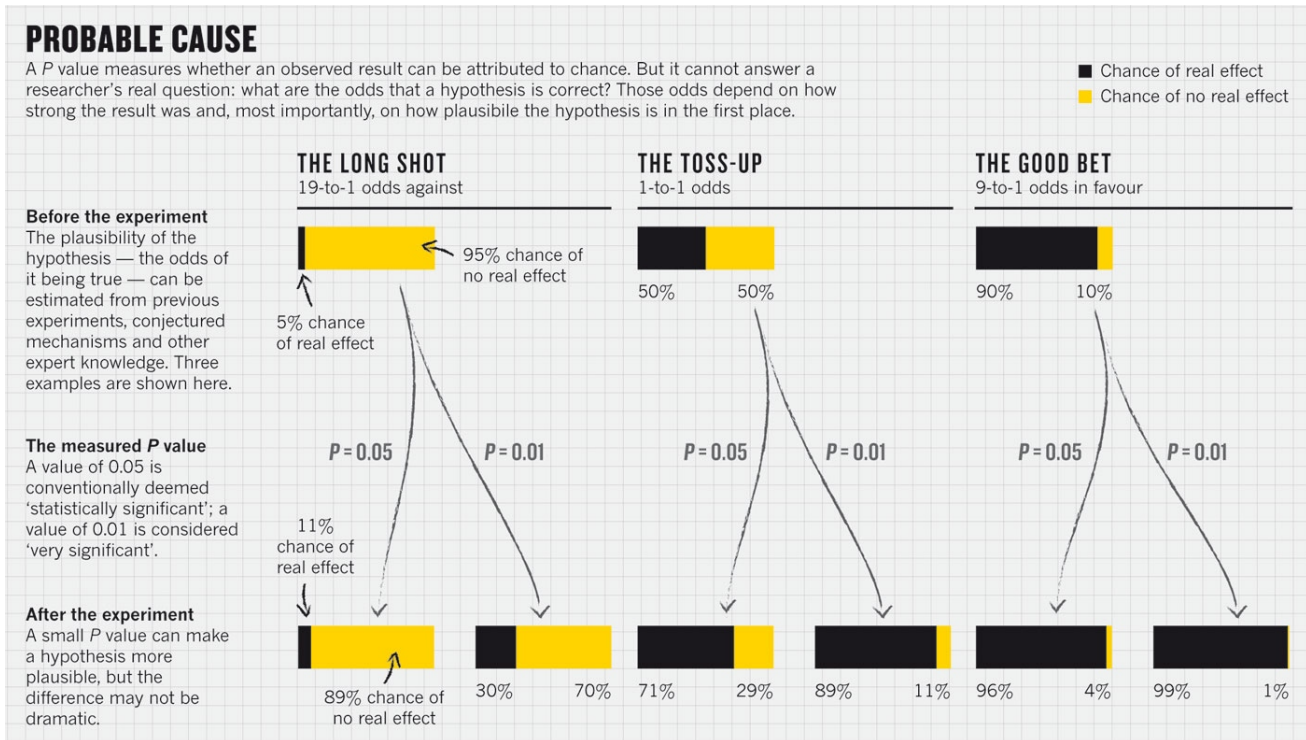
Larger samples

Why

- Larger and more representative samples
- Cheap
- Fast
- More brains means a higher probability to find the right expertise (see Nielsen, 2011)

Larger samples

Why

- Larger and more representative samples
- Cheap
- Fast
- More brains means a higher probability to find the right expertise (see Nielsen, 2011)

Why not

- Need to control what people are actually doing
- Browsers have poor performance

Larger samples

Why

- Larger and more representative samples
- Cheap
- Fast
- More brains means a higher probability to find the right expertise (see Nielsen, 2011)

Why not

- Need to control what people are actually doing
- Browsers have poor performance

Available tools

- jsPsych
- Amazon Mechanical Turk (use with psiTurk), CrowdSource, ...
- Or create a virtual reality with Oculus Rift

Data sharing

Why

- More transparency
- Better data archiving standards
- More data makes Big Data, i.e., we can discover more

Data sharing

Why

- More transparency
- Better data archiving standards
- More data makes Big Data, i.e., we can discover more

Why not

- Data acquisition is costly and precious
- People want to know who is using their data
- No time to play these games of sharing
- Who is going to use all of that data anyways? The problem of filtering again.

Data sharing

Why

- More transparency
- Better data archiving standards
- More data makes Big Data, i.e., we can discover more

Why not

- Data acquisition is costly and precious
- People want to know who is using their data
- No time to play these games of sharing
- Who is going to use all of that data anyways? The problem of filtering again.

Available tools

GitHub, figshare, Zenodo, BitBucket, Dropbox …

Licencing

- **Copyright (c):** all rights reserved
- **Copyleft:** all rights reversed
    - **Permissive** (anybody can use your stuff for anything they like – even sell it):
        - Software: BSD 3-clause or MIT
        - Else: Creative Commons Attribution (cc-by)
    - **Viral** (reciprocal – anything based on your work will be of the same same license type, ensuring perpetual access to it):
        - Software: GNU General Public License (recommended for scientific software so that it always remains open)
        - Else: Creative Commons ShareAlike (cc-by-sa) (think twice before using)
- **Public domain:** no rights reserved Creative Commons Zero (cc0) if you need to want to put it in the public domain sooner than it would get there legally

Free and Open Source Software

Copylefted software

Why

- Transparent
- Accessible
- Persistent

Why not

- Often less powerful
- Less polished

Available tools

- E-Prime, MATLAB, Presentation -> Python
- SPSS -> Python, R
- Photoshop -> GIMP
- Illustrator -> Inkscape
- ...

Chapter 3

where we get stuck in human nature

Quality

- Need nice outputs (check eLife Lens, PeerJ Paper Now)
- Online repositories (arXiv, bioRxiv) need recommendation systems (just like in Amazon, Netflix, Spotify), e.g., Pubget, Google Scholar
- Change reward system (Kubilius et al., 2013)

Peer review
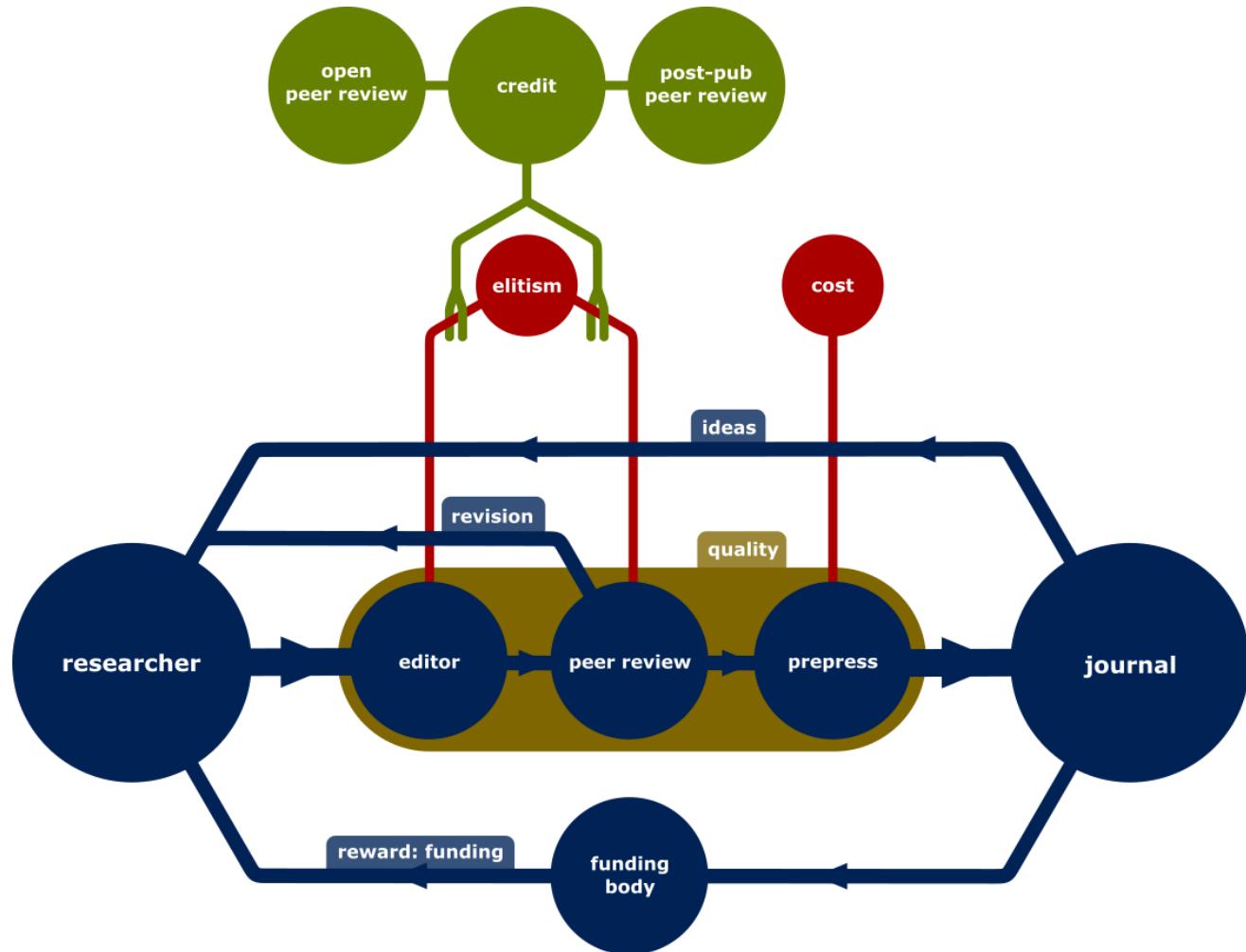
Peer review

Issues

- How much we get paid for doing reviews (~5 hours of work (Ware, 2011)):
    - €20
    - €45 (~minimal wage in Belgium)
    - €60
    - €100?
- Major bottleneck (with revision might take a year)
- Elitism
- Why only two people decide the fate of your paper?

Peer review

Issues

- How much we get paid for doing reviews (~5 hours of work (Ware, 2011)):
  - €20
  - €45 (~minimal wage in Belgium)
  - €60
  - €100?
- Major bottleneck (with revision might take a year)
- Elitism
- Why only two people decide the fate of your paper?

Solutions

- Double-blind review – nope
- Open peer review (e.g, PeerJ)
- Giving credit to reviewers (e.g., Publons)
- Post-publication peer review

Academia's 1%

You've got no chances: Only 3% of PhDs become professors (30% of PhDs get a postdoc, and 10% of postdocs become a professor)

Academia's 1%

You've got no chances: <u>Only 3% of PhDs become professors</u> (30% of PhDs get a postdoc, and 10% of postdocs become a professor)

The academic elite has much more chances:

- Top 25% US institutions produce 71-86% tenure-track faculty
- Only 9-14% get positions in more prestigious institutions than their PhD

Academia's 1%

You've got no chances: Only 3% of PhDs become professors (30% of PhDs get a postdoc, and 10% of postdocs become a professor)

The academic elite has much more chances:

- Top 25% US institutions produce 71-86% tenure-track faculty
- Only 9-14% get positions in more prestigious institutions than their PhD

And needless to say: Don't be a woman. A report on women researchers found that

> "Both Latinas and Black women report regularly being mistaken as janitors." *New Study Reveals "Double Jeopardy" Faced by Women of Color in STEM* / UC Hastings Newsroom

A beggar's life

Even if successful, your life will be based on Impact Factors:

Steve Pettifer
@srp

Follow

Stop crowing about your new Impact Factors. They're calculated like this. (tho at least here the sum is transparent)
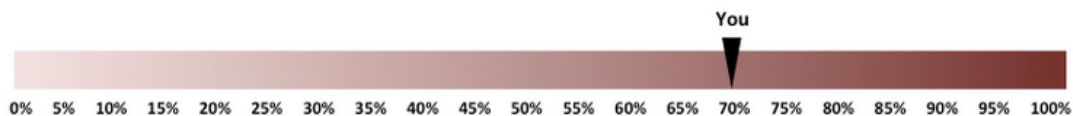
12:57 PM - 30 Jul 2014
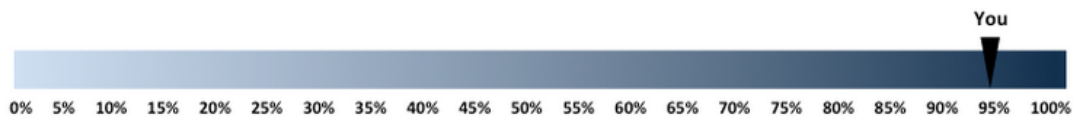
76 RETWEETS **54** FAVORITES

I'm fine, folks

Your probability to become a PI is:

75%

You vs. PIs (you outscore this percent of people who became a PI)

You

| 0% | 5% | 10% | 15% | 20% | 25% | 30% | 35% | 40% | 45% | 50% | 55% | 60% | 65% | 70% | 75% | 80% | 85% | 90% | 95% | 100% |

You vs. Non-PIs (you outscore this percent of people who did not become a PI)

You

| 0% | 5% | 10% | 15% | 20% | 25% | 30% | 35% | 40% | 45% | 50% | 55% | 60% | 65% | 70% | 75% | 80% | 85% | 90% | 95% | 100% |

Your score is greater than that of 70% of PIs in our data, while compared to non-PIs, you outscore 95% of them (i.e., false discovery rate of 5%)

Here are some of the statistics we computed from your publication history:

Feature	Value
Number of Publications	6
Number of Publications as First	4
Mean Impact Factor	3.5
Max Impact Factor	7.1
Max IF As First	7.1

PIPredictor

Chapter 4

where we wrap up
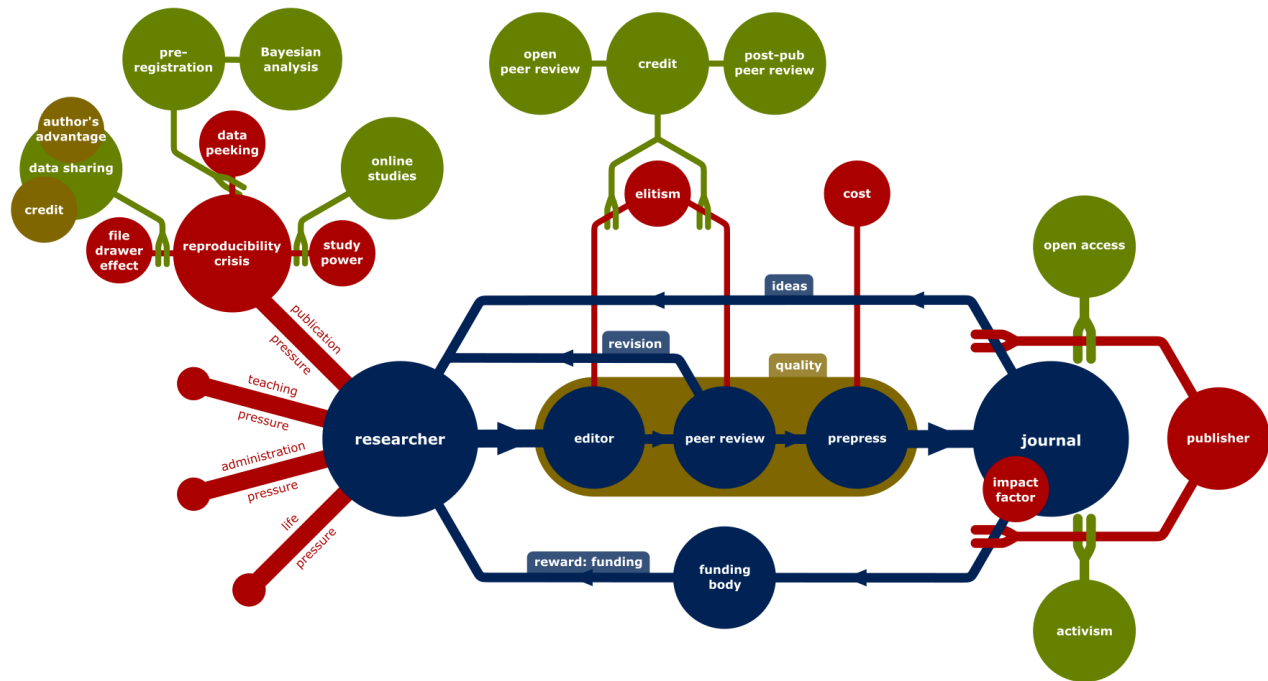
But why bother?

Open Science / Open Culture lead to transparency, accountability, clear thought.

The End

where we thank each other for patience

Extra resources

- Belluz, J. (2015). Taxpayers spend $140 billion funding science each year — but can't access many of the results. Vox.
- Nielsen, M. (2011). Reinventing Discovery: The New Era of Networked Science. Princeton, N.J.: Princeton University Press.
- Kubilius, J. (2014). Sharing code. I-Perception, 5(1), 75–78.
- Nosek, B. A., & Bar-Anan, Y. (2012). Scientific Utopia: I. Opening Scientific Communication. Psychological Inquiry, 23(3), 217–243.
- Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. Perspectives on Psychological Science, 7(6), 615–631.
- Viskontas, I., & Mooney, C. (2014). Scenes from the Postdocalypse. Mother Jones.