LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Changing incentive structures to foster the actual sharing rate of open data

or:

Why everybody loves data sharing,
but nobody does it.

PD Dr. Felix Schönbrodt
Ludwig-Maximilians-Universität
München

RESEARCH TRANSPARENCY

OSC LMU Open Science Center

www.nicebread.de
www.researchtransparency.org
@nicebread303


#osc2018
#oscibar


Open Science

"the European Commission is now moving decisively from 'Open access' into the broader picture of '**Open science**'"

➔ Open Data is default (with opt-out possibility)

- German Research Foundation (DFG): Publicly funded research data belongs to the public

- G7 science ministers: „recognize open science practices during evaluation of funding proposals and outcomes; reward open science activities in career advancement"

http://ec.europa.eu/programmes/horizon2020/en/h2020-section/open-science-open-access
http://www8.cao.go.jp/cstp/english/others/20160517communique.pdf

The Mertonian norms of science

Communality
The findings of science belong to everyone, they are not private property.

Counternorm: Secrecy
Hiding procedures, materials, and results

Organized skepticism
All ideas must be tested and are subject to rigorous, structured community scrutiny.

Counternorm: Organized Dogmatism:
Old findings are not challenged, no independent verification takes place.

Disinterestedness
Scientists should be focused on finding the truth, not on their own success.

Universalism
The validity of a scientific claim does not depend on who is making it.

Anderson, M. S., & Martinson, B. C. (2007). doi.org/10.2307/1754865; Merton, 1947

The Mertonian norms of science

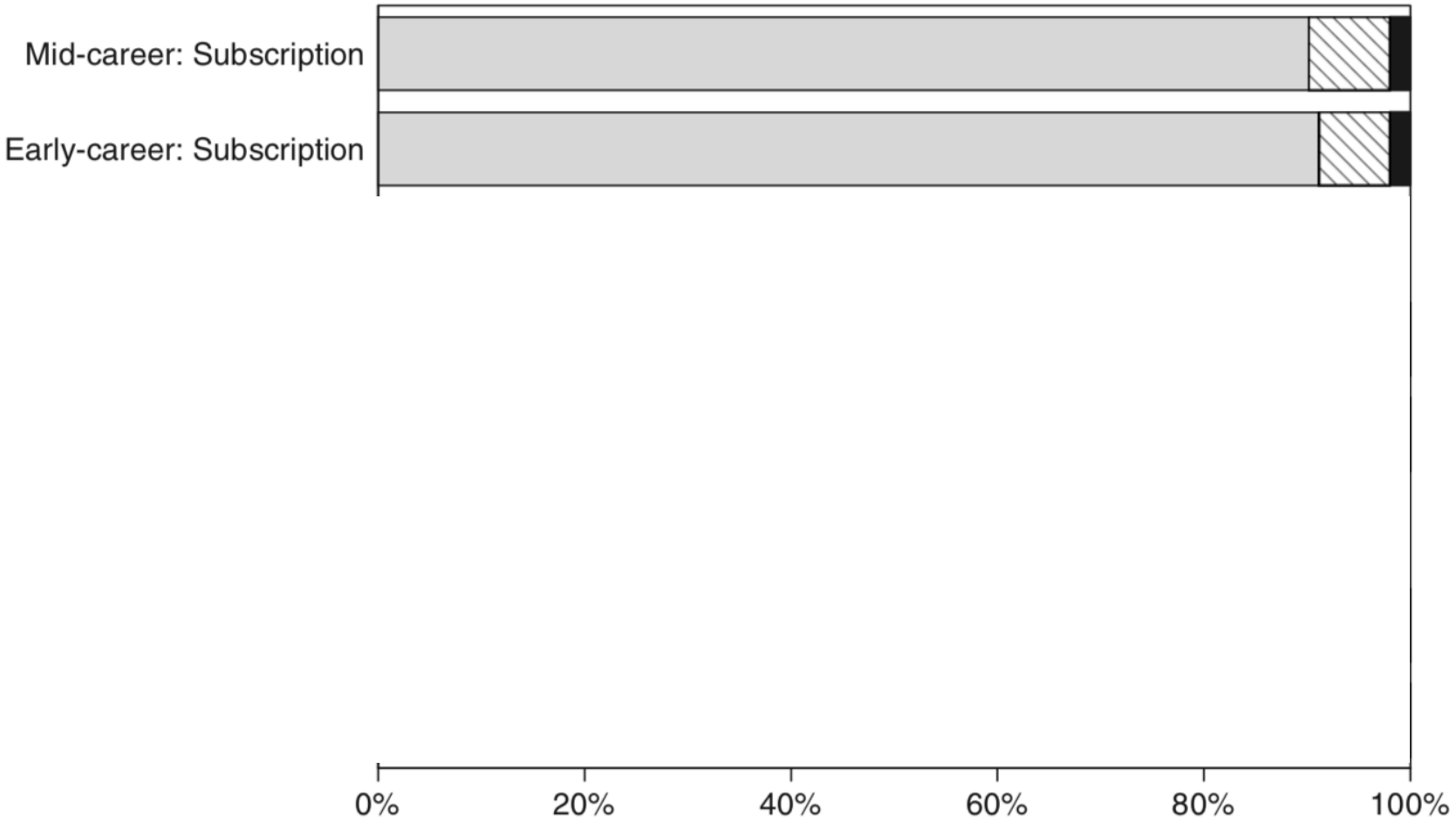
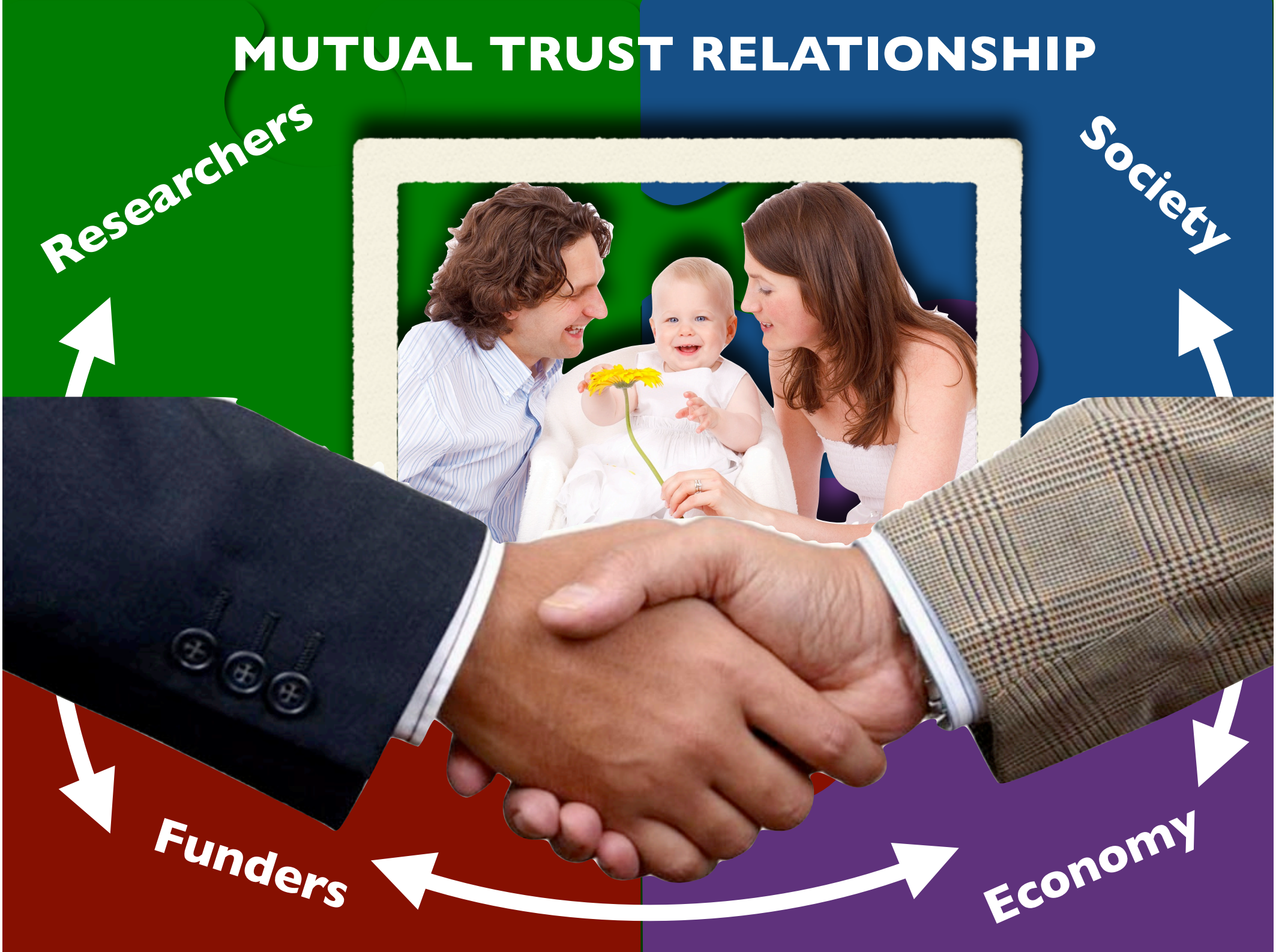


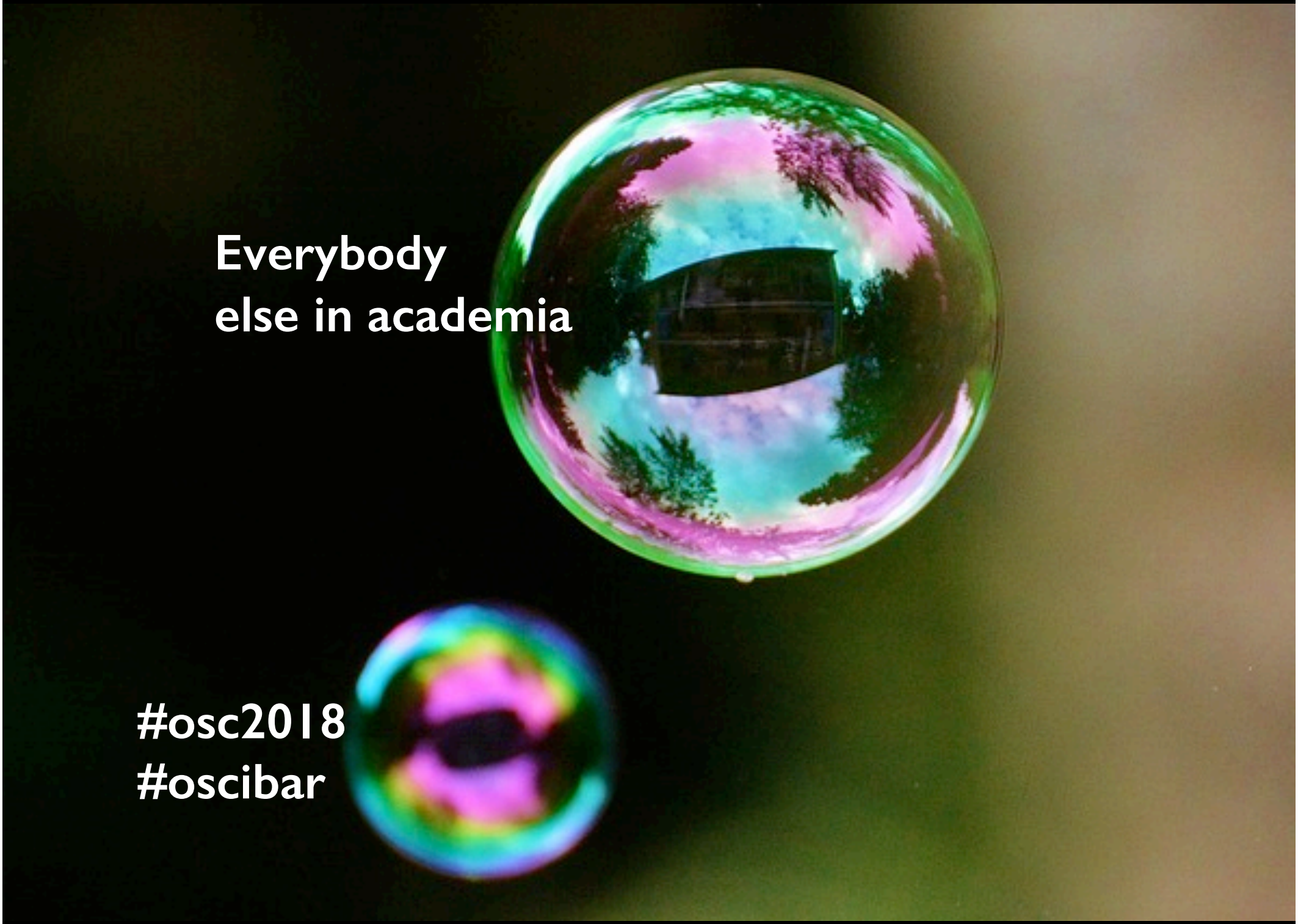
FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

Anderson, M. S., & Martinson, B. C. (2007). doi.org/10.2307/1754865

$N = 3{,}247$

MUTUAL TRUST RELATIONSHIP




Everybody
else in academia
#osc2018
#oscibar

The Mertonian norms of science

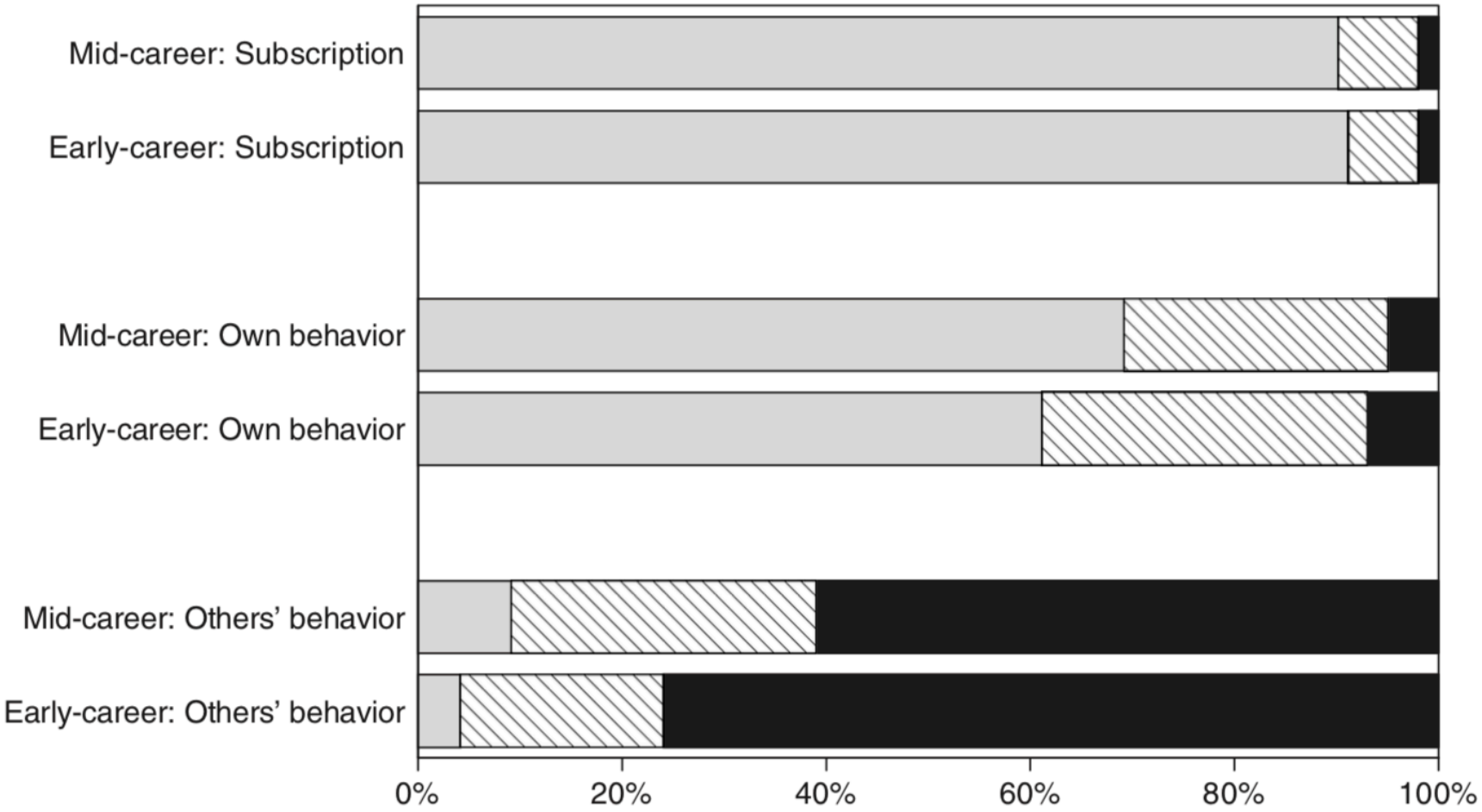


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

Secrecy

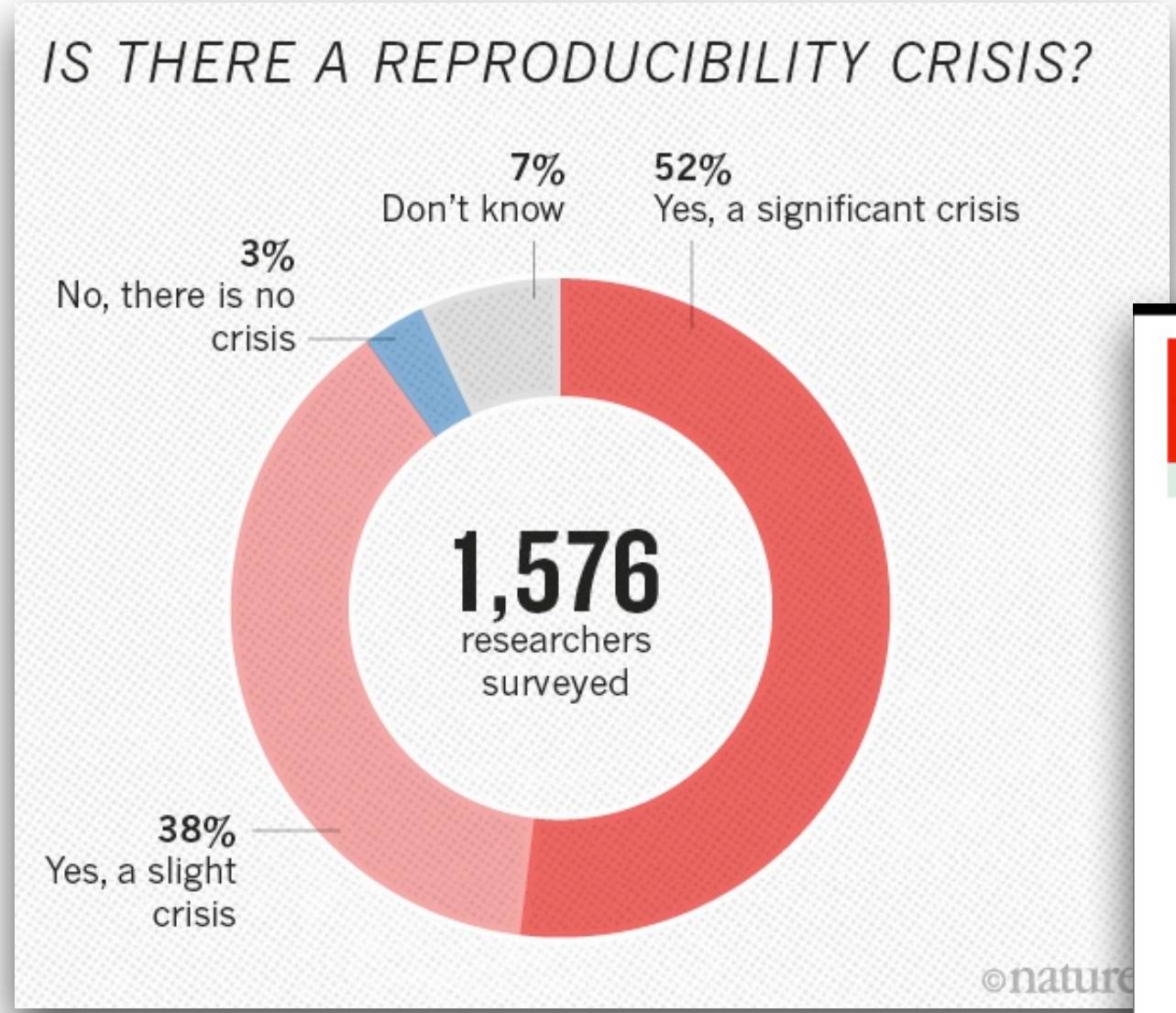
Organized dogmatism

…

Anderson, M. S., & Martinson, B. C. (2007). doi.org/10.2307/1754865

$N = 3{,}247$

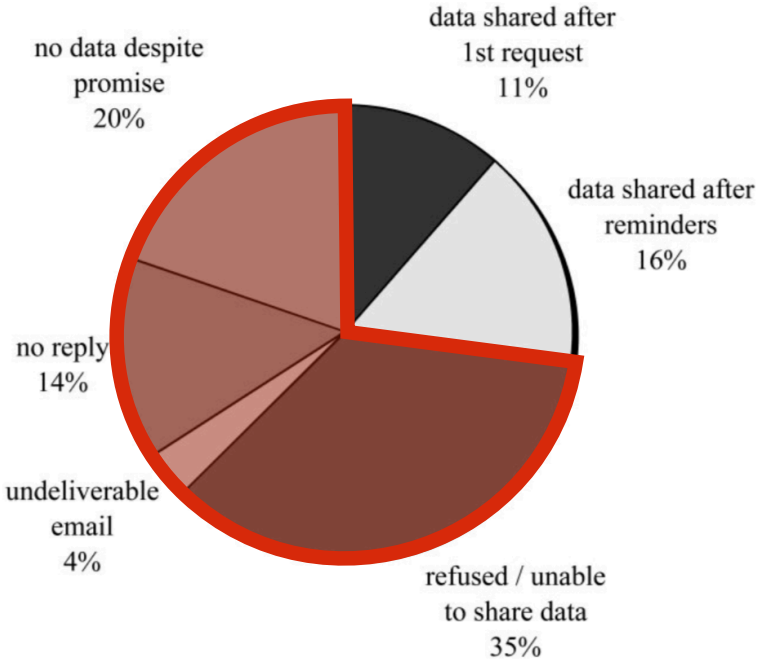
Eroding trust in science



90%: Yes

http://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

„Sharing upon request“ as a policy is dead



- **100%** of authors in these studies signed to share the data upon request
- Actual sharing rate (Wicherts et al., 2006):
  27% (out of 141 requests)
- Vanpaemel et al. (2015):
  38% (out of 394 requests)
- Stodden et al. (2018):
  44% (out of 204 requests) provided some „artifacts“, 26% could be reproduced

- Bus factor / long-term availability?
- Providing selective access (e.g., not to critics)?
- Data set providers should not be in charge for access control ➙ either fully open, or independent stewards grant access based on prespecified rules

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). http://doi.org/10.1037/0003-066X.61.7.726
Vanpaemel, W., Vermorgen, M., Deriemaecker, L., & Storms, G. (2015). http://doi.org/10.1525/collabra.13
Stodden, V., Seiler, J., & Ma, Z. (2018). http://doi.org/10.1073/pnas.1708290115

Why not sharing

Rewarding quantity, not quality

Actual (not desired) relevance in professorship hiring committees	Rank
Number of peer-reviewed publications	1
Fit of research profile to the hiring department	2
Quality of research talk	3
Number of publications	4
Volume of acquired third-party funding	5
Number of first authorships	6
…	…

N = 1453 psychology researchers, 66% were members of a professorship hiring committee.

Abele-Brehm, A. E., & Bühner, M. (2016). Wer soll die Professur bekommen? Psychologische Rundschau, 67(4), 250–261. http://doi.org/10.1026/0033-3042/a000335

Early career researchers are stuck

> What would be a good balance between Open Science and having a career in academia? […] Being open IMHO is a competitive disadvantage. Can you only afford open science when you are tenured?

> Why should I share my hard-won data with my rivals that presumably compete with me for the next post-doc position?

> My contract is limited to two years – although it would be nice to publish the data, I have no time to do it. I rather have to churn out another publication.

➞ felt contradiction between „good research“/„open research“ and „having a career in science“

Quantity, not quality

Actual (not desired) relevance in professorship hiring committees	Rank
Number of peer-reviewed publications	1
Fit of research profile to the hiring department	2
Quality of research talk	3
Number of publications	4
Volume of acquired third-party funding	5
Number of first authorships	6
…	…
Quality rating of the three best publications	17
…	…
Indicators of research transparency	**41 (of 41)**

N = 1453 psychology researchers, 66% were members of a professorship hiring committee.

Abele-Brehm, A. E., & Bühner, M. (2016). http://doi.org/10.1026/0033-3042/a000335

Farnham *et al. Genome Biology* (2017) 18:221
DOI 10.1186/s13059-017-1351-7

Genome Biology

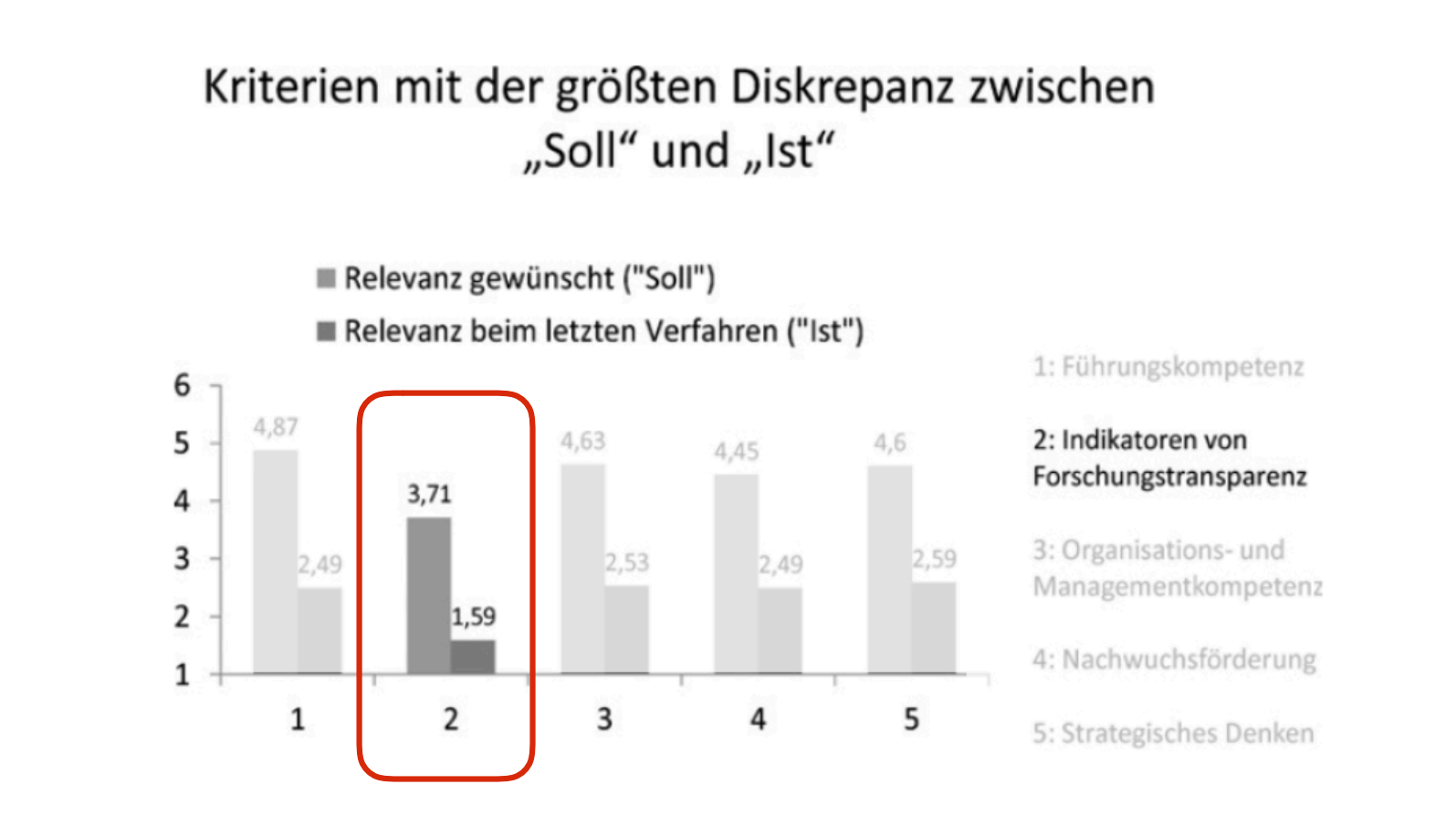
COMMENT Open Access

Early career researchers want Open Science

Andrea Farnham[1], Christoph Kurz[2,3*], Mehmet Ali Öztürk[4], Monica Solbiati[5], Oona Myllyntaus[6], Jordy Meekes[7], Tra My Pham[8], Clara Paz[9], Magda Langiewicz[10], Sophie Andrews[11], Liisa Kanninen[6], Chantal Agbemabiese[12], Arzu Tugce Guler[13], Jeffrey Durieux[14], Sarah Jasim[15], Olivia Viessmann[11], Stefano Frattini[16], Danagul Yembergenova[17], Carla Marin Benito[9], Marion Porte[18], Anaïs Grangeray-Vilmint[19], Rafael Prieto Curiel[8], Carin Rehncrona[20], Tareq Malas[21], Flavia Esposito[9] and Kristina Hettne[21]

It is the responsibility of senior researchers, funders, and policy makers to resolve this social dilemma for young researchers.

Quantity, not quality



Highest discrepancies between desired relevance and actual relevance

N = 1453 psychology researchers, 66% were members of a professorship hiring committee.

Abele-Brehm, A. E., & Bühner, M. (2016). http://doi.org/10.1026/0033-3042/a000335

The typical researcher's narrative about data sharing / open science

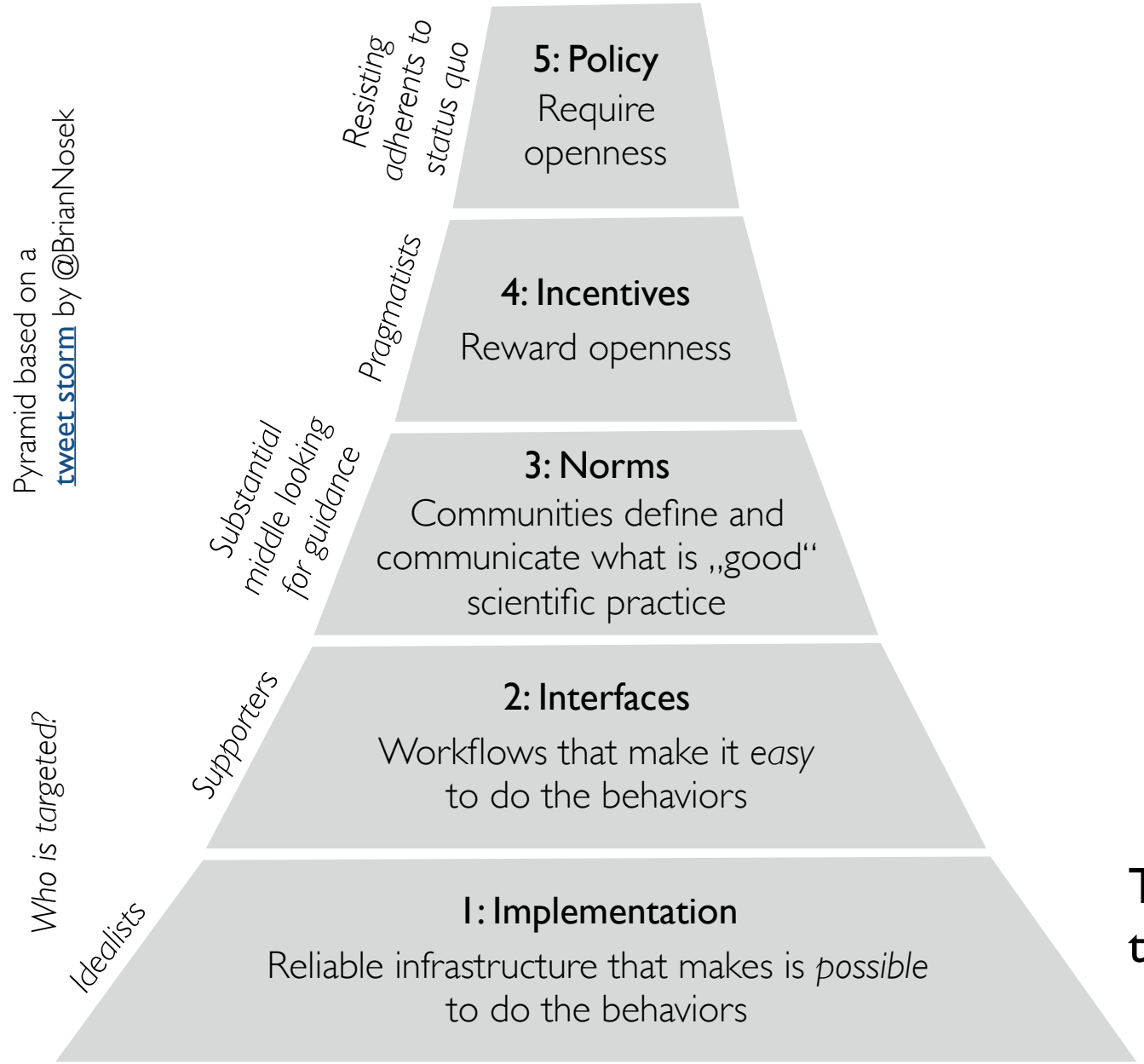
maybe slightly exaggerated (but maybe not)

- Nobody does it – why should I?
- A lot of work, which is not rewarded.
- RDM is BORING
- Strategic trade-off: More papers on CV, or documenting old stuff? In order to get tenure/more grant money, I'd rather optimize the former.
- Please: No bureaucratic over-regulation. Protect academic freedom!

Going forward: How to increase the actual sharing rate

https://commons.wikimedia.org/wiki/File:Soap_Bubble_-_foliage_background_-_iridescent_colours_-_Traquair_040801.jpg

How to achieve cultural change



Barriers

No proper recognition for sharing (27%)

Sharing data is not a common practice in my field (68%)

Preparing data is too time-consuming (55%)

There is no suitable repository to share my data (12%)

I never learned to share data online (54%)

n = 600 researchers

Houtkoop, B. L., et al (2018) http://doi.org/10.1177/2515245917751886

How to achieve cultural change

5: Policy
Require openness

4: Incentives
Reward openness

3: Norms
Communities define and communicate what is „good“ scientific practice

2: Interfaces
Workflows that make it *easy* to do the behaviors

1: Implementation
Reliable infrastructure that makes is possible to do the behaviors

3. Features of the DC-DS-XML Syntax

3.1 URIs in DC-DS-XML

The Abstract Model uses Uniform Resource Identifiers (URIs) [RFC3986] to refer bot
schemes).

In DC-DS-XML, URIs are encoded as URI references, used as XML attribute values. A
different XML attributes in detail. The purpose of this section is to make some gener

The URI may be represented in full. The following example shows a URI as the valu

XML Example 1: URI as attribute value

```
<?xml version="1.0" encoding="UTF-8" ?>
<dcds:descriptionSet
  xmlns:dcds="http://purl.org/dc/xmlns/2008/09/01/dc-ds-xml/">
  <dcds:description>
                                  <!-- Property URI -->
    <dcds:statement dcds:propertyURI="http://purl.org/dc/terms/title">
      <dcds:literalValueString>DCMI Home Page</dcds:literalValueString>
    </dcds:statement>
  </dcds:description>
</dcds:descriptionSet>
```

The representation of the URI may be abbreviated through the use of an XML entity

XML Example 2: URI as attribute value (with XML entity reference)

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE dcds:descriptionSet [
  <!ENTITY dcterms 'http://purl.org/dc/terms/'>
```

Dublin Core Metadata Scheme

http://dublincore.org/documents/2008/09/01/dc-ds-xml/

5: Policy
Require openness

4: Incentives
Reward openness

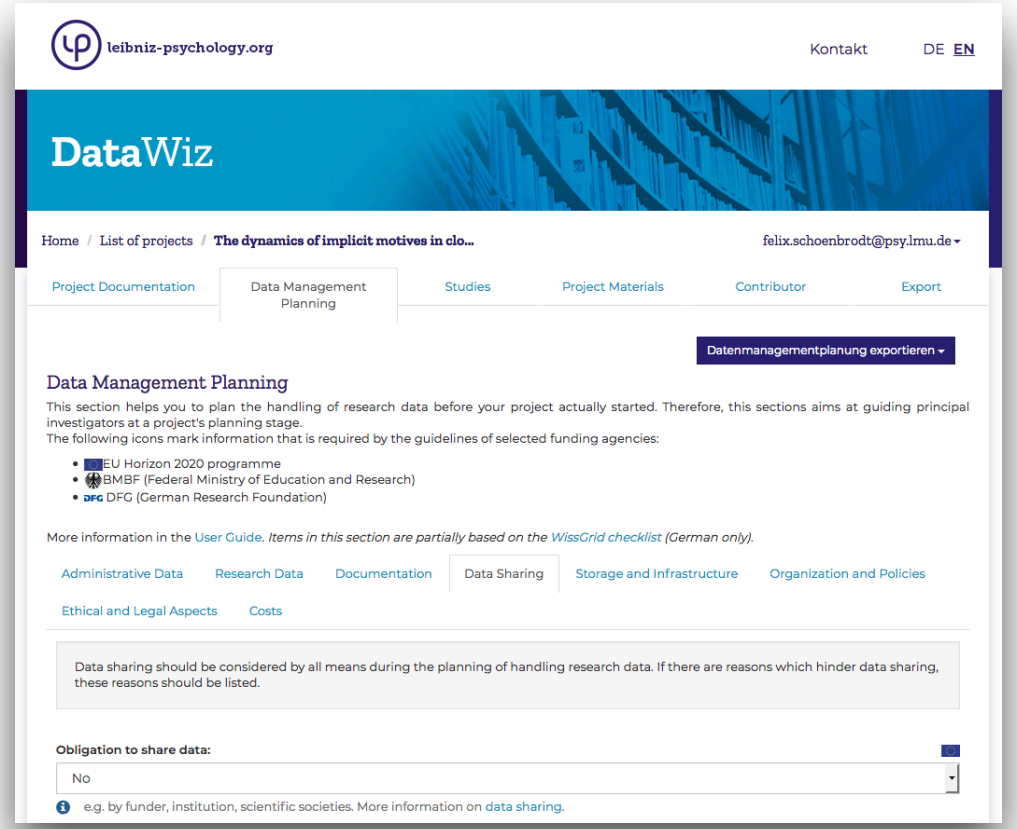
3: Norms
Communities define and communicate what is „good“ scientific practice

2: Interfaces
Workflows that make it *easy* to do the behaviors

1: Implementation
liable infrastructure that makes is *possible* to do the behaviors

OpenAIRE

FOSTER



https://datawiz.leibniz-psychology.org

GO FAIR

GO TRAIN
- Education/training
- Certification

➞ Software solutions + supporting persons (data stewards) at the local level

PEER REVIEWERS' OPENNESS INITIATIVE

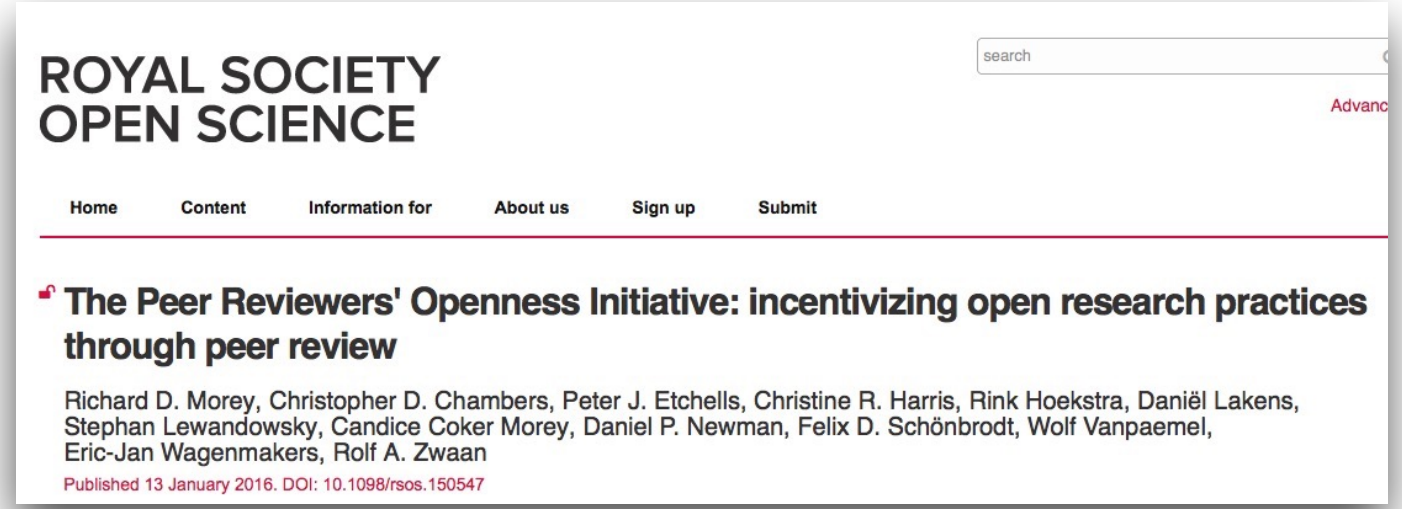
https://opennessinitiative.org/

Sign up!

5: Policy
Require openness

4: Incentives
Reward openness

3: Norms
Communities define and communicate what is „good" scientific practice


ROYAL SOCIETY OPEN SCIENCE

search

Advanc

Home Content Information for About us Sign up Submit

The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review

Richard D. Morey, Christopher D. Chambers, Peter J. Etchells, Christine R. Harris, Rink Hoekstra, Daniël Lakens, Stephan Lewandowsky, Candice Coker Morey, Daniel P. Newman, Felix D. Schönbrodt, Wolf Vanpaemel, Eric-Jan Wagenmakers, Rolf A. Zwaan

Published 13 January 2016. DOI: 10.1098/rsos.150547

„We suggest that beginning January 1, 2017, **reviewers make open practices a pre-condition for more comprehensive review**.
This is already in reviewers' power; to drive the change, all that is needed is for reviewers to collectively agree that the time for change has come."

- More and more journals change from an opt-in to an opt-out (+public justification) policy
- Educate students:
  This is how science is done - these are the norms of good scientific practice and integrity.

Open Science Badges

Extra cost for journals	**Very few** (add badges to workflow)
Extra cost for reviewers	**Few** (verify availability) to **some** (reproduce)
Extra cost for researchers	**Some**

Open Science Badges

Extra cost for journals	**Very few** (add badges to workflow)
Extra cost for reviewers	**Few** (verify availability) to **some** (reproduce)
Extra cost for researchers	**Some**

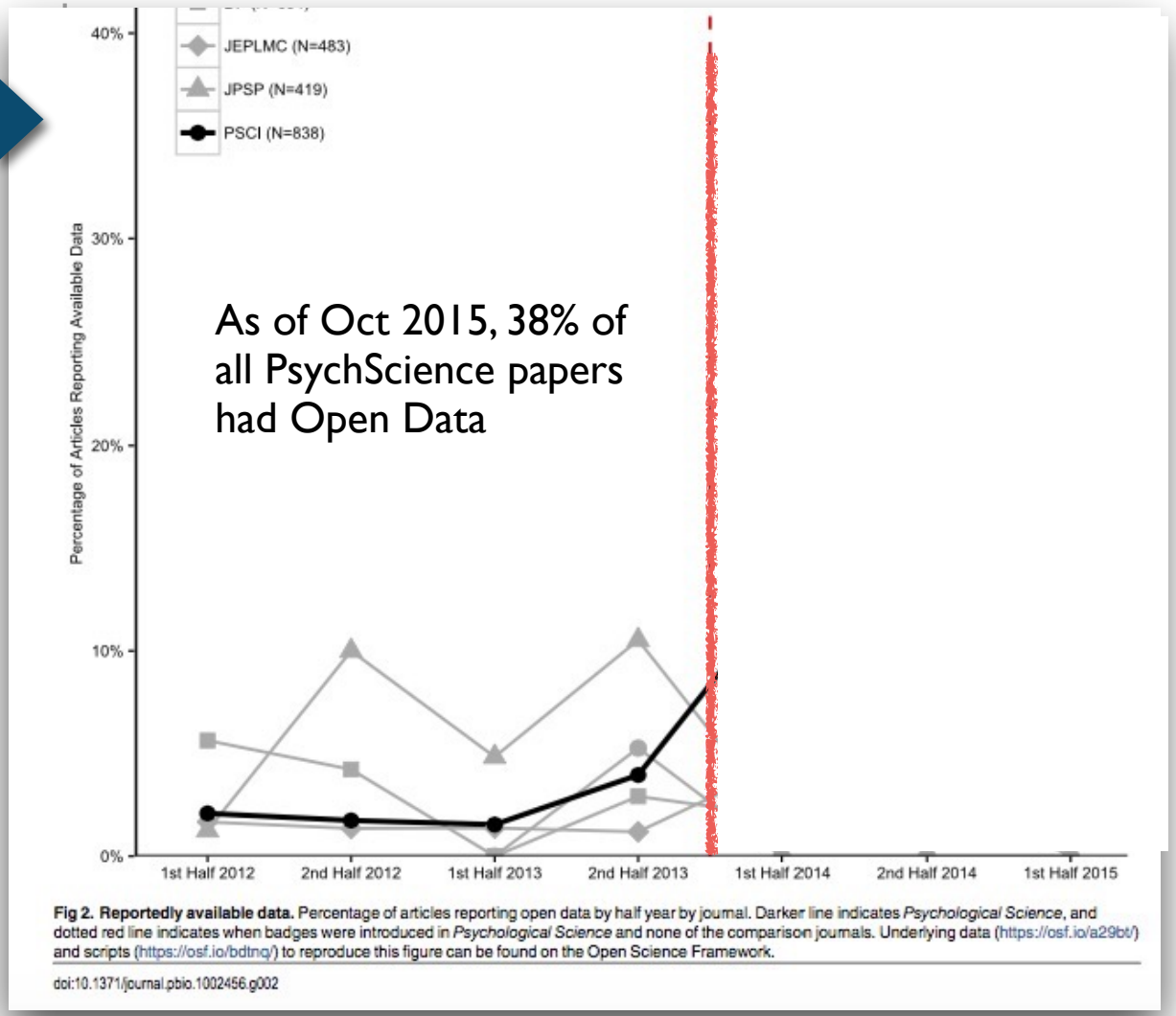
5: Policy
Require openness

4: Incentives
Reward openness

3: Norms
Communities define and communicate what is „good" scientific practice

2: Interfaces
Workflows that make it *easy* to do the behaviors

1: Implementation
infrastructure that makes is *possible* to do the behaviors



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

Kidwell et al. (2016). http://doi.org/10.1371/journal.pbio.1002456

Funders: Add „Public data sets“ as a section to CV templates

Extra cost for funders	**None** (add a few sentences to guidelines)
Extra cost for reviewers	**None** (take information into consideration)
Extra cost for researchers	**~5 min.**

DFG

1.1 Publications list as part of the academic curriculum vitae:

- must be included for each applicant,
- need not directly relate to the proposed project,
- must include up to ten of the most important publications for each applicant,

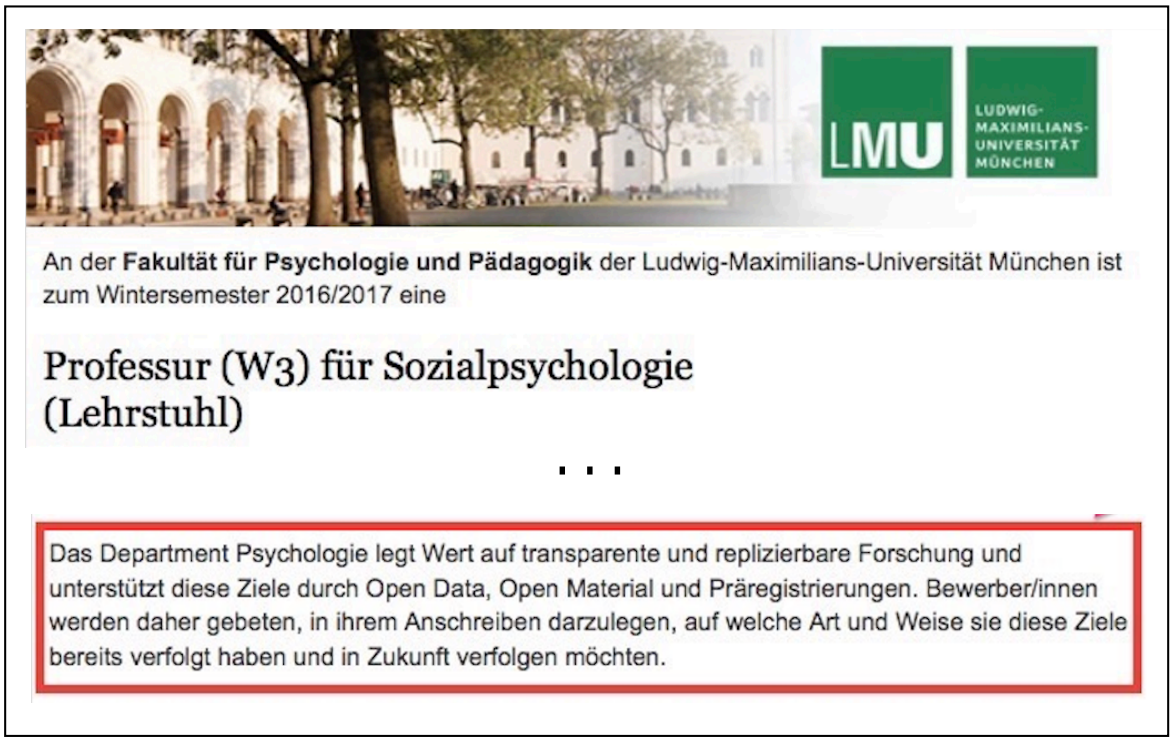
Suggestion:

„Publication list must include a section with up to 5 of the most impactful public data sets that an applicant provides, with a one-sentence statement about each data set's specific impact.“

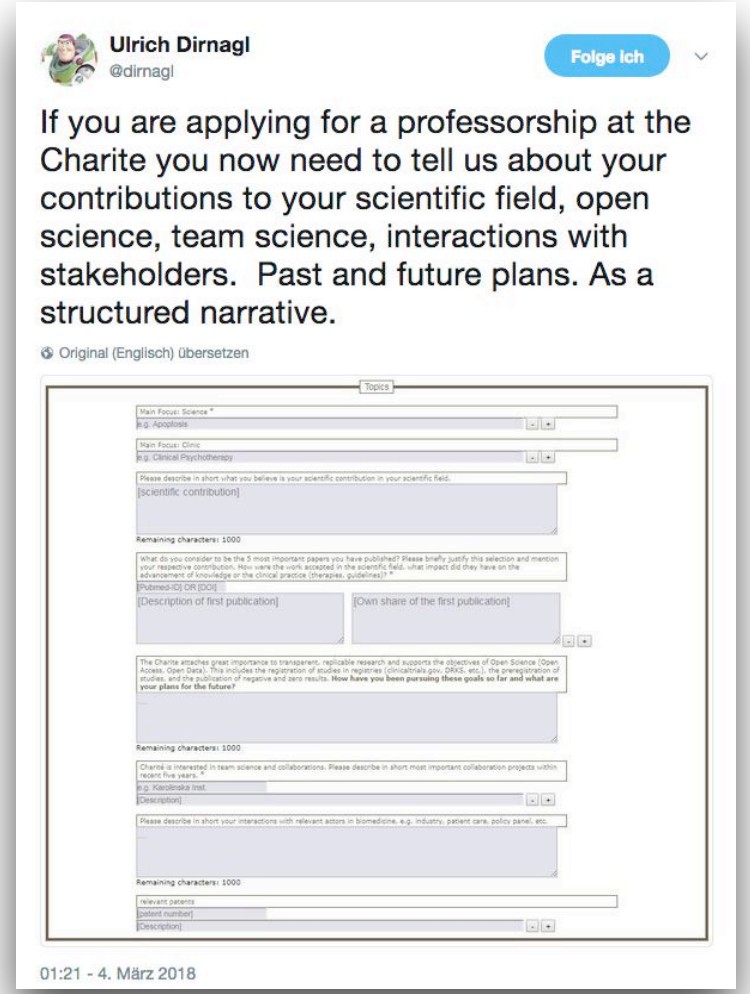
http://www.dfg.de/formulare/1_91/1_91_en.pdf

Hiring committees: Make „open science“ a desirable or essential job characteristic

Extra cost for committees	**None** (add a paragraph to job description)
Extra cost for reviewers	**None** (take information into consideration)
Extra cost for applicants	**a few minutes**



Since 2015: All professorship job descriptions use this requirement



See more such prof job ads at: https://osf.io/7jbnt/

Hiring committees: Make „open science“ a desirable or essential job characteristic

Extra cost for committees	**None** (add a paragraph to job description)
Extra cost for reviewers	**None** (take information into consideration)
Extra cost for applicants	**a few minutes**

For staff roles involving **at least some research**, signatories (employers) self-certify to meet ONE of the levels below. Signatories may wish to apply different levels of commitment for different grades or type of appointment. **Typical categories could be (a) PhD students/ research assistants, (b) Post-Doc, or (c) faculty (i.e., associate and full professors).**

	Level 0	Level I	Level II	Level III
	Individual or organisation makes no commitment to mention open research practices in published hiring policies or advertised research job descriptions.	Individual or organisation makes no commitment to mention open research practices in advertised research-related job descriptions but does include them as desirable characteristics in published hiring policy. All else being equal, candidates with greater open science track records may be preferred over candidates with no or lesser open science track record.	Individual or organisation commits to including proven track record of open research practices as desirable characteristics (but not necessarily as essential characteristics) in all advertised research-related job descriptions. All else being equal, candidates with greater open science track records are preferred over candidates with no or lesser open science track record.	Individual or organisation commits to including proven track record of open practices in all advertised research-related job descriptions as essential characteristics. Only candidates with an open science track record are interviewed and/or appointed. All else being equal, candidates with greater open science track records are preferred over candidates with lesser open science track record.

https://docs.google.com/document/d/1ty43Syw0Flkh8ncjW8MZArIkvYe8hLwwhLIlwbtSk_Y/edit?usp=drive_web&ouid=108982640291853577145

Hiring committees: Require an annotated CV with limited items (e.g., <= 10)

Extra cost for committees	**None** (add a paragraph to job description)
Extra cost for reviewers	**None** (take information into consideration)
Extra cost for applicants	**~ 30 min.**

No journal; JIF is irrelevant or misleading

Paper-level citation metrics

Basic information for judging evidential value

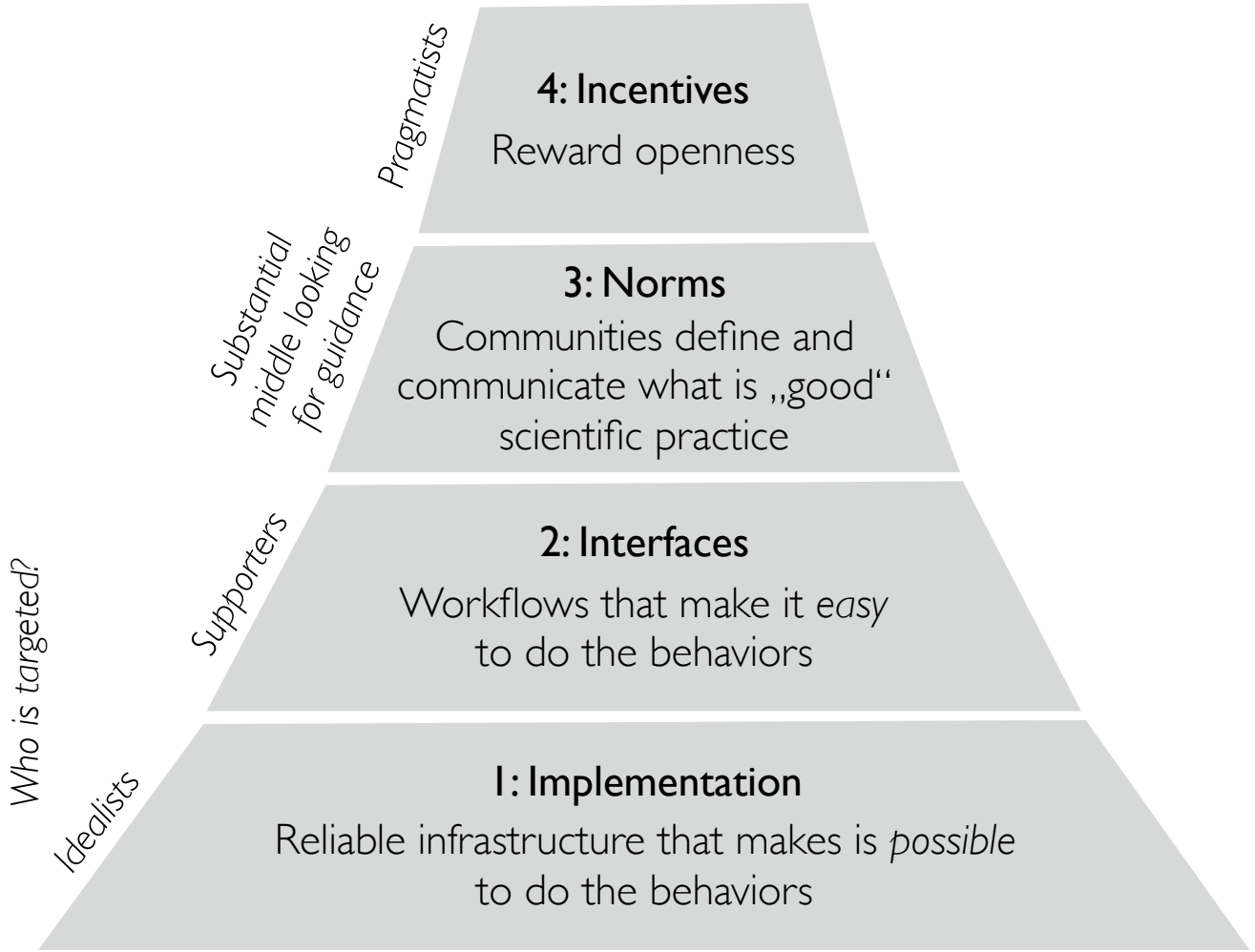
Open science indicators: Judging replicability

Data: own collection or reuse?

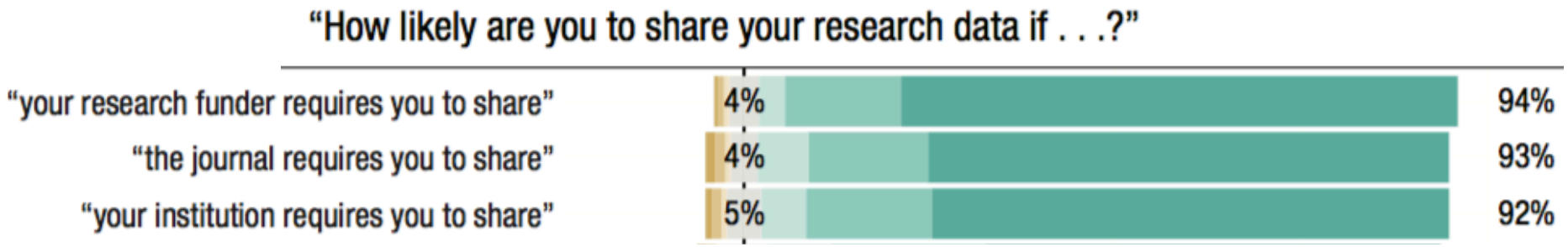
Authors & title	Year	Cit-ations	Sample size per study	*p*-value per study	Open Science indicators	Data set	Applicants contribution
Doe, John & Smith, Peter	2001	47	$n_1 = 21$ $n_2 = 30$ $n_3 = 19$	$p_1 = .048$ $p_2 = .050$ $p_3 = .023$	☐Open Data ☐Open Material ☐Preregistered	☑Own data collection →URL *NA* ☐Archival data	• Analyzed data • Wrote manuscript
Doe, John	2016	26	$n_1 = 180$ $n_2 = 158$	$p_1 = .012$ $p_2 = .001$	☑ Open Data ☑Open Material ☑Preregistered	☑ Own data collection →URL osf.io/as1cd ☐ Archival data	• Designed study • Wrote manuscript

Dougherty, M. R., Slevc, L. R., & Grand, J. (2018, February 2). Retrieved from psyarxiv.com/48qux

"How likely are you to share your research data if . . .?"



Pyramid based on a tweet storm by @BrianNosek



Houtkoop, B. L., Chambers, C., Macleod, M., Bishop, D. V. M., Nichols, T. E., & Wagenmakers, E. J. (2018). Data Sharing in Psychology: A Survey on Barriers and Preconditions. Advances in Methods and Practices in Psychological Science, 63(1), 251524591775188–16. http://doi.org/10.1177/2515245917751886

How to achieve cultural change

5: Policy
Require openness

4: Incentives
Reward openness

3: Norms
Communities define and communicate what is „good" scientific practice

2: Interfaces
Workflows that make it *easy* to do the behaviors

1: Implementation
Infrastructure that makes is *possible* to do the behaviors

wellcome

„We expect our researchers to maximise the availability of research data, software and materials with as few restrictions as possible**. As a minimum, the data underpinning research papers should be made available to other researchers at the time of publication. [...]**
Wellcome will also consider **whether researchers have managed and shared their research outputs** in line with our requirements**, as a critical part of the end of grant reporting process**."

NIH

„expects and supports the timely release and sharing of final research data"

FNSNF
Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung

„erwartet der SNF, dass Daten [...] auf öffentlich zugänglichen, digitalen Datenbanken archiviert werden"

ACADEMY OF FINLAND

„It is recommended to make all research data [...] available for reuse, for example under Creative Commons licence"

Input control → output control?

Funders: Require Transparency and Openness (TOP) statement in final reports

Extra cost for funders	**None** (add a few sentences to guidelines)
Extra cost for reviewers	**None** (take information into consideration)
Extra cost for researchers	**~5 min.**

Are the relevant data from the funded project accessible in an open repository?	
Yes	Provide a persistent, unique identifier and any required instructions
No	Provide justification (short free text)
Not applicable	Provide explanation (short free text)

Have you cited any previously generated data used in this project?
Yes
No
Not applicable

1. Disclose →
2. Require →
3. Verify

Aalbersberg, I. J., Appleyard, T., Brookhart, S., Carpenter, T., Clarke, M.,... Vazire, S. (2018, February 15). Making Science Transparent By Default; Introducing the TOP Statement. Retrieved from osf.io/sm78t; See also https://osf.io/n9mrh/

Action List: „Bridging the last mile“

- **Universities:** Educate and practice the values and principles of good scientific practice.
- **Universities:** Provide supporting infrastructure, such as data stewards.
- **Universities:** Add research transparency as desirable or essential job characteristic for post-doc and prof positions
- **Infrastructure:** Provide *user-friendly* tools
- **Journals:** Make open data the default; authors can opt-out with a *public* justification
- **Funders:** Appreciate openness in grant proposal (both on project level and applicant level)
- **Funders:** Require transparency and openness statement in final reports; use openness track record for future decisions

Fast adoption vs. High (FAIR) quality?

- Low hurdles, one small step at a time
- Reward small steps
  Sharing something - even badly documented data - is better than sharing nothing.
- Learning by doing
  With increasing practice, hopefully the quality gets better, too.
- But: (Initially) Low quality
  Barely reusable data sets; trying to reproduce a result is a pain in the ass or impossible; data reuse very limited.
- Risk of „open-washing“
  Pretending openness without actual value.

- High hurdles
  Mainly enthusiasts/computer scientists will able and motivated use it
- Reward big steps
  Curated repositories with input quality control.
- Instant high quality
  The data sets which are open are instantly FAIR.

https://commons.wikimedia.org/wiki/File:Soap_Bubble_-_foliage_background_-_iridescent_colours_-_Traquair_040801.jpg

PD Dr. Felix Schönbrodt
Ludwig-Maximilians-Universität
München

RESEARCH TRANSPARENCY

OSC
LMU Open Science Center

www.nicebread.de
www.researchtransparency.org
@nicebread303