



# **Ciencia abierta y colaborativa: Un enfoque práctico**

Lucía Castillo Iglesias

Facultad de Educación, UCSC

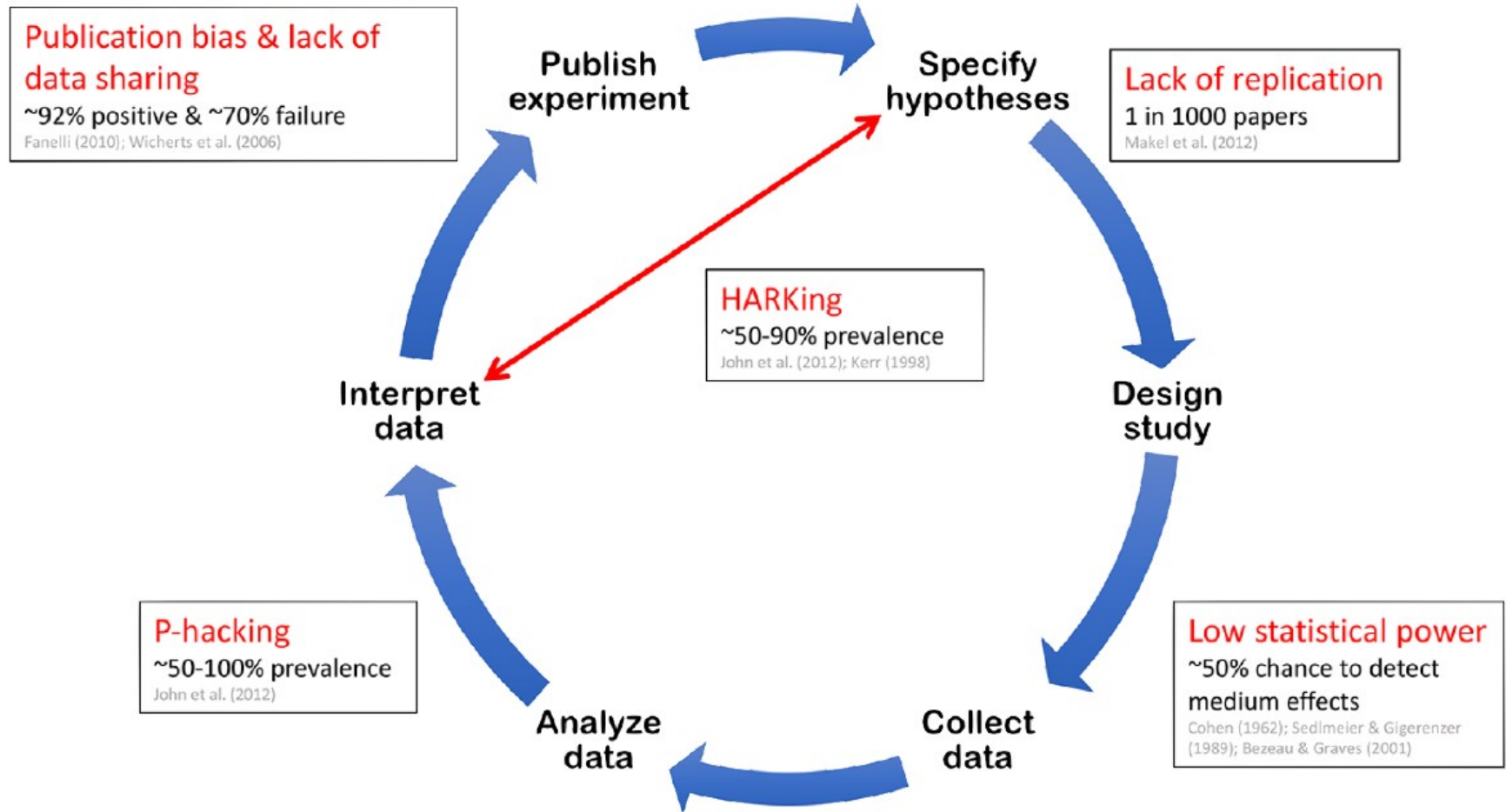
# Algunas consideraciones previas

- El estado de avance de estas discusiones varía enormemente entre disciplinas. Esta presentación se plantea *desde las ciencias sociales, en particular desde la psicología*.
- La adopción de estas prácticas depende mucho de un cambio en la cultura académica, y eso depende de la actividad concreta de todos nosotros: de nuestra práctica investigativa individual, de nuestras asociaciones, de nuestros proyectos

# Distinciones conceptuales clave

		DATA	
		SAME	DIFFERENT
ANALYSIS	SAME	REPRODUCIBLE	REPLICABLE
	DIFFERENT	ROBUST	GENERALISABLE

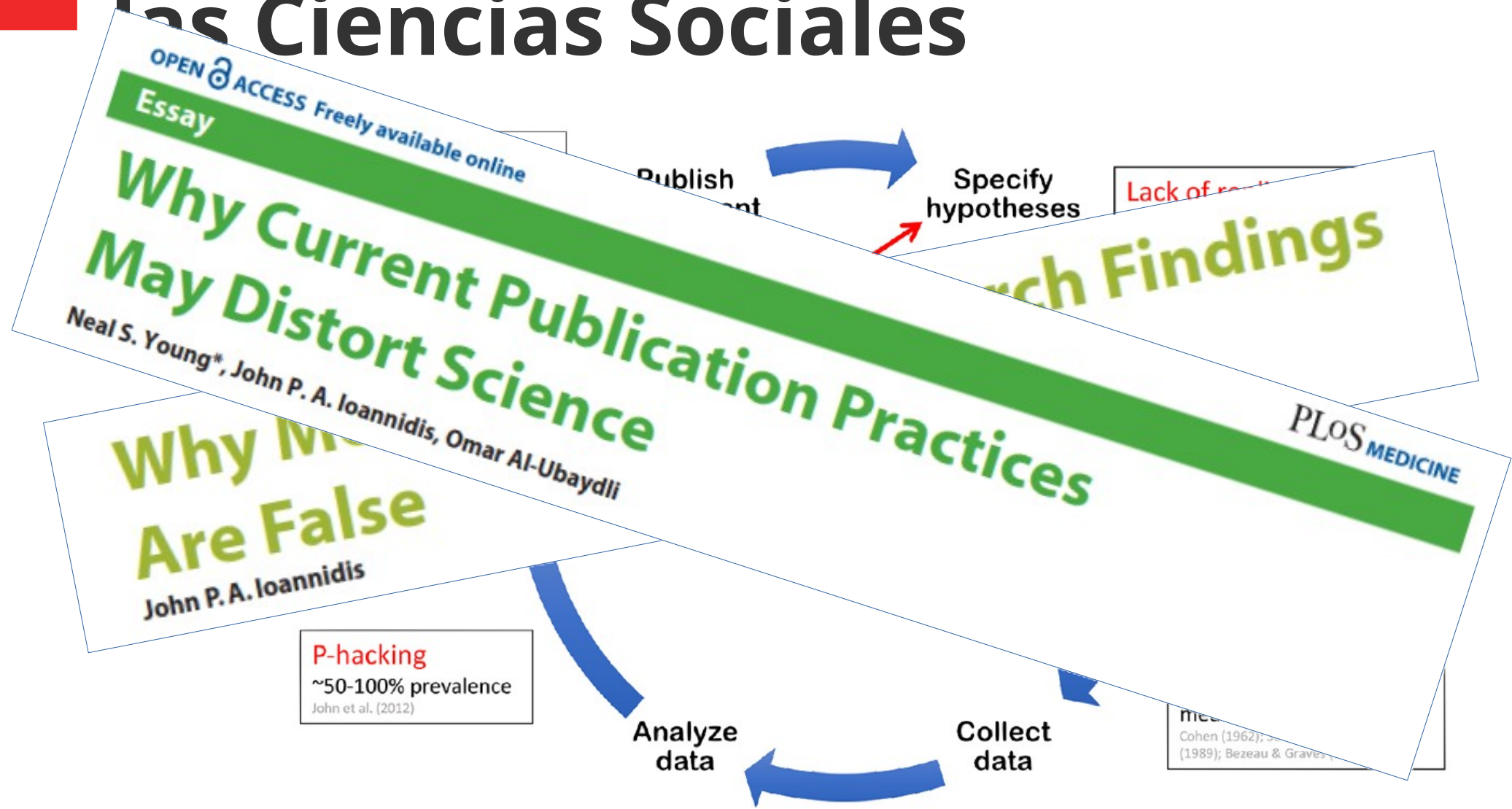
# La crisis de reproducibilidad en las Ciencias Sociales



# La crisis de reproducibilidad en las Ciencias Sociales



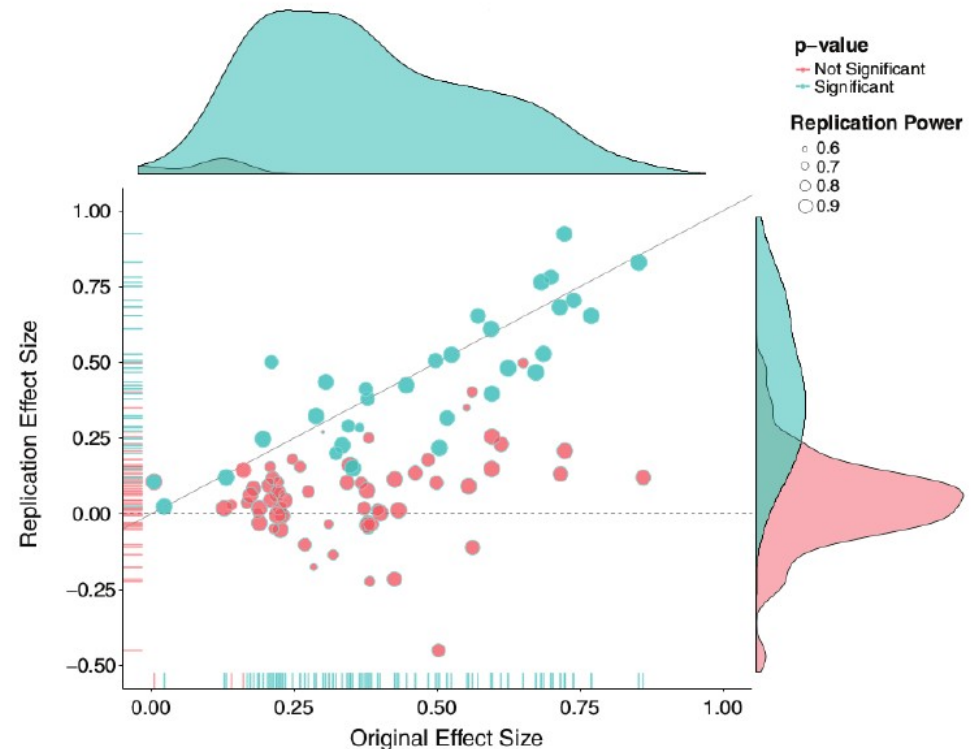
# La crisis de reproducibilidad en las Ciencias Sociales



# Baja replicabilidad

Open Science Collaboration (2015), "Estimating the reproducibility of psychological science"

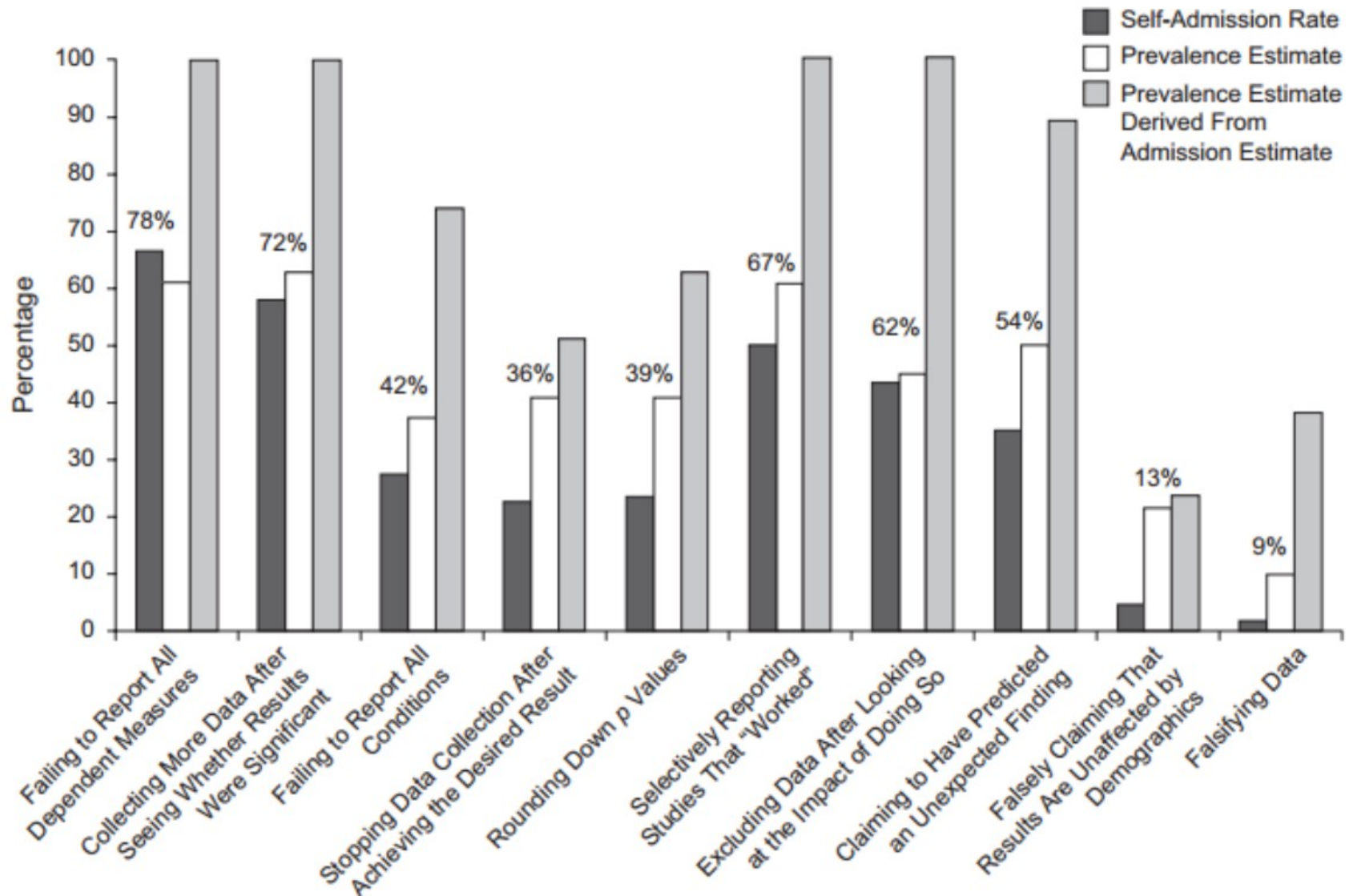
- Attempted replications of 100 studies published in 2008 in three journals (JEP:LMC, PS, JPSP)
- 97% of originals had  $p < .05$
- 36% of replications had  $p < .05$



**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.

# Manipulación de datos

John, Loewenstein, & Prelec (2012)







# El problema del acceso a la información

- Baja credibilidad de las publicaciones
- Pérdida de esfuerzos
- Escasa colaboración
- *Publication bias* y predominio del *paper* en la evaluación académica: “Selección natural de mala ciencia” (Smaldino & MacElreath, 2016)
- Ralentiza el avance de la ciencia



# Pasos concretos hacia una cultura de ciencia abierta

- **La ciencia abierta es colaborativa y transparente:** Acabar con la lógica del genio individual
  - **Materiales abiertos**

Colaboración y avance del conocimiento científico
  - **Código abierto**

El código abierto permite no sólo reproducir análisis sino que además disminuye la complejidad de la curva de aprendizaje y democratiza el acceso
  - **Datos abiertos**

“Todo conocimiento generado con fondos públicos debe reportar beneficios para la sociedad y, en consecuencia, estar disponible para su uso” (Aisén Etcheverry, ex-directora ANID)

# Materiales abiertos

[www.nature.com/scientificdata](http://www.nature.com/scientificdata)

## scientific **data**



OPEN

### The Multilingual Picture Database

DATA DESCRIPTOR

Jon Andoni Duñabeitia *et al.*<sup>#</sup>

The growing interdisciplinary research field of psycholinguistics is in constant need of new and up-to-date tools which will allow researchers to answer complex questions, but also expand on languages other than English, which dominates the field. One type of such tools are picture datasets which provide naming norms for everyday objects. However, existing databases tend to be small in terms of the number of items they include, and have also been normed in a limited number of languages, despite the recent boom in multilingualism research. In this paper we present the Multilingual Picture (Multipic) database, containing naming norms and familiarity scores for 500 coloured pictures, in thirty-two languages or language varieties from around the world. The data was validated with standard methods that have been used for existing picture datasets. This is the first dataset to provide naming norms, and translation equivalents, for such a variety of languages; as such, it will be of particular value to psycholinguists and other interested researchers. The dataset has been made freely available.

# Materiales abiertos



Browse

Search on figshare...



Log in

Sign up

<  
1/1



DATASET

MultiPic Dataset Open M... xlsx (1015.1 kB)



DATASET

MultiPic Dataset Open Mat... csv (944.6 kB)



ARCHIVE

MultiPic Dataset Pictures.zip (14.82 MB)



ARCHIVE

Color Variations of MultiPic... zip (29.67 MB)

Switch View | 4 files

## MultiPic: Multilingual Picture Database

Cite

Download all (46.4 MB)

Share

Embed

+ Collect

Version 5 ▾ Dataset posted on 2022-06-29, 05:31 authored by [Jon Andoni Duñabeitia](#)

Database of name agreement and familiarity for a set of 500 colored pictures of concrete objects in 32 different languages (i.e., American English, Australian English, Basque, Belgium Dutch, British English, Catalan, Cypriot Greek, Czech, Finnish, French, German, Greek, Hebrew, Hungarian, Italian, Korean, Lebanese Arabic, Malay, Malay English, Mandarin Chinese, Netherlands Dutch, Norwegian, Polish, Portuguese, Quebec French, Rioplatense Spanish, Russian, Serbian, Slovak, Spanish, Turkish, Welsh).

USAGE METRICS [↗](#)

1291  
views

426  
downloads

1  
citations [↗](#)



CATEGORIES


# Código abierto

nature human behaviour

View all journals

Search 

Login 

Explore content 

About the journal 

Publish with us 

Subscribe

Sign up for alerts 

RSS feed

[nature](#) > [nature human behaviour](#) > [articles](#) > article

Article | [Published: 10 August 2020](#)

## Cultural influences on word meanings revealed through large-scale semantic alignment


[Bill Thompson](#) , [Seán G. Roberts](#) & [Gary Lupyan](#)

*Nature Human Behaviour* **4**, 1029–1038 (2020) | [Cite this article](#)

3242 Accesses | 5 Citations | 224 Altmetric | [Metrics](#)

### Abstract

If the structure of language vocabularies mirrors the structure of natural divisions that are universally perceived, then the meanings of words in different languages should closely align. By contrast, if shared word meanings are a product of shared culture, history and geography, they may differ between languages in substantial but predictable ways. Here, we analysed the semantic neighbourhoods of 1,010 meanings in 41 languages. The most-aligned words were from semantic domains with high internal structure (number, quantity and kinship). Words denoting natural kinds, common actions and artefacts aligned much less well. Languages that are more geographically proximate, more historically related and/or spoken by more-similar cultures had more aligned word meanings. These results provide evidence that the meanings of common words vary in ways that reflect the culture, history and geography of their users.

 Access through your institution

Buy or subscribe

Sections

Figures

References

Abstract

[Data availability](#)

[Code availability](#)

[References](#)

[Acknowledgements](#)

[Author information](#)

[Ethics declarations](#)

[Additional information](#)

[Supplementary information](#)

[Rights and permissions](#)

[About this article](#)

# Código abierto

## Access options

### Rent or Buy article

Get time limited or full article access on ReadCube.

from **\$8.99**

Rent or Buy

All prices are NET prices.

### Subscribe to Journal

Get full journal access for 1 year

**\$119.00**

only \$9.92 per issue

Subscribe

All prices are NET prices.  
VAT will be added later in the checkout.  
Tax calculation will be finalised during checkout.

### Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

### Data availability

Data and reproducible analyses are available at <https://osf.io/tngba/>.

### Code availability

Code to implement the alignment algorithm is available at <https://osf.io/tngba/>.

### References

1. Gleitman, L. & Fisher, C. In *The Cambridge Companion to Chomsky* (ed. McGilvray, J.) 123–142 (Cambridge Univ. Press, 2005).

#### Sections

#### Figures

#### References

#### Abstract

[Data availability](#)

[Code availability](#)

[References](#)

[Acknowledgements](#)

[Author information](#)

[Ethics declarations](#)

[Additional information](#)

[Supplementary information](#)

[Rights and permissions](#)

[About this article](#)

# Datos abiertos

OSFHOME Semantic Alignment

Contributors: Bill Thompson, Gary Luppyan, Sean Roberts  
Date created: 2019-08-27 06:59 PM | Last Updated: 2020-08-10 07:13 AM  
Identifier: DOI 10.17605/OSF.IO/TNGBA  
Category: Project

Name	Modified
additional-datasets.zip	2020-06-09 12:26 AM
alignments-nel-wiki-tr1.zip	2019-08-27 07:06 PM
analyze_alignment.zip	2020-04-08 12:37 PM
clean-data-for-release.py	2019-08-27 07:07 PM
compute-alignment.zip	2020-04-07 06:24 PM
cultural-analyses.zip	2020-06-09 12:41 AM
README.md	2019-08-29 09:30 AM

Recent Activity

- Bill Thompson created external identifier(s) doi:10.17605/OSF.IO/TNGBA on Semantic Alignment 2020-08-10 07:13 AM
- Bill Thompson made Semantic Alignment public 2020-06-09 12:44 AM
- Bill Thompson updated file cultural-analyses.zip in OSF Storage in Semantic Alignment 2020-06-09 12:41 AM
- Bill Thompson updated file additional-datasets.zip in OSF Storage in Semantic Alignment 2020-06-09 12:26 AM
- Bill Thompson updated file analyze\_alignment.zip in OSF Storage in Semantic Alignment 2020-04-08 12:37 PM
- Bill Thompson updated file compute-alignment.zip in OSF Storage in Semantic Alignment 2020-04-07 06:24 PM

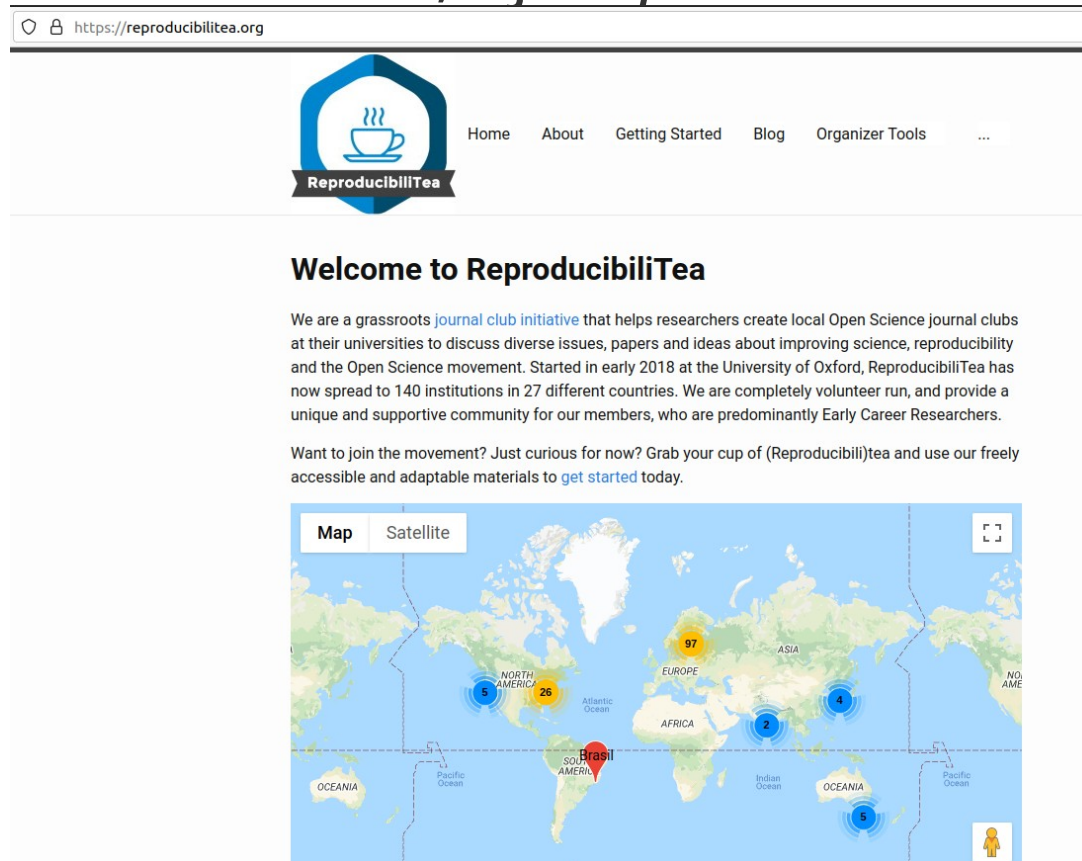
Name
analyses
data
</> 0_munge_data.html
0_munge_data.Rmd
</> 1_validating_alignment.html
1_validating_alignment.Rmd
2_multipic_analyses.Rmd
</> 3_psychological_predictors.html
3_psychological_predictors.Rmd
</> 4_rate_of_change.html
4_rate_of_change.Rmd
</> 5_polysemy_alignment.html
5_polysemy_alignment.Rmd

Análisis y datos disponibles en repositorios de acceso abierto: reproducibilidad, replicabilidad, avance de la ciencia



# Pasos concretos hacia una cultura de ciencia abierta

- Grupos transversales a nivel local:
  - Clubes de lectura, ej. *ReproducibiliTea*



The screenshot shows the homepage of the ReproducibiliTea website. At the top, there is a navigation menu with links for Home, About, Getting Started, Blog, Organizer Tools, and a dropdown menu. The main heading is "Welcome to ReproducibiliTea". Below this, there is a paragraph describing the organization as a grassroots journal club initiative that helps researchers create local Open Science journal clubs. It mentions that the organization was started in early 2018 at the University of Oxford and has spread to 140 institutions in 27 different countries. A second paragraph encourages joining the movement and provides a link to get started. At the bottom, there is a world map showing the number of clubs in various regions: North America (5), Europe (97), Asia (4), Africa (2), Oceania (5), and South America (26). A red pin is placed on Brazil in South America.


https://reproducibilitea.org

Home About Getting Started Blog Organizer Tools ...

## Welcome to ReproducibiliTea

We are a grassroots [journal club initiative](#) that helps researchers create local Open Science journal clubs at their universities to discuss diverse issues, papers and ideas about improving science, reproducibility and the Open Science movement. Started in early 2018 at the University of Oxford, ReproducibiliTea has now spread to 140 institutions in 27 different countries. We are completely volunteer run, and provide a unique and supportive community for our members, who are predominantly Early Career Researchers.

Want to join the movement? Just curious for now? Grab your cup of (Reproducibili)tea and use our freely accessible and adaptable materials to [get started](#) today.



Region	Number of Clubs
North America	5
Europe	97
Asia	4
Africa	2
Oceania	5
South America	26

# Pasos concretos hacia una cultura de ciencia abierta

- Nuevos criterios en estadística
  - El reinado de los *p-values* ha generado una disciplina dispuesta a obviar criterios de realidad y estimaciones robustas
  - “Pensamiento dicotómico” en torno a un valor único que pasa a determinar la validez del estudio completo

6. Build a cumulative quantitative discipline.
7. Whenever possible, adopt estimation thinking and avoid dichotomous thinking.
8. Remember that obtained results are one possibility from an infinite sequence.
9. Do not trust any *p* value.
10. Whenever possible, avoid using statistical significance or *p* values; simply omit any mention of null-hypothesis significance testing (NHST).
11. Move beyond NHST and use the most appropriate methods, whether estimation or other approaches.
12. Use knowledgeable judgment in context to interpret observed effect sizes (ESs).
13. Interpret your single confidence interval (CI), but bear in mind the dance. Your 95% CI just might be one of the 5% that miss. As Figure 1 illustrates, it might be red!
14. Prefer 95% CIs to *SE* bars. Routinely report 95% CIs, and use error bars to depict them in figures.

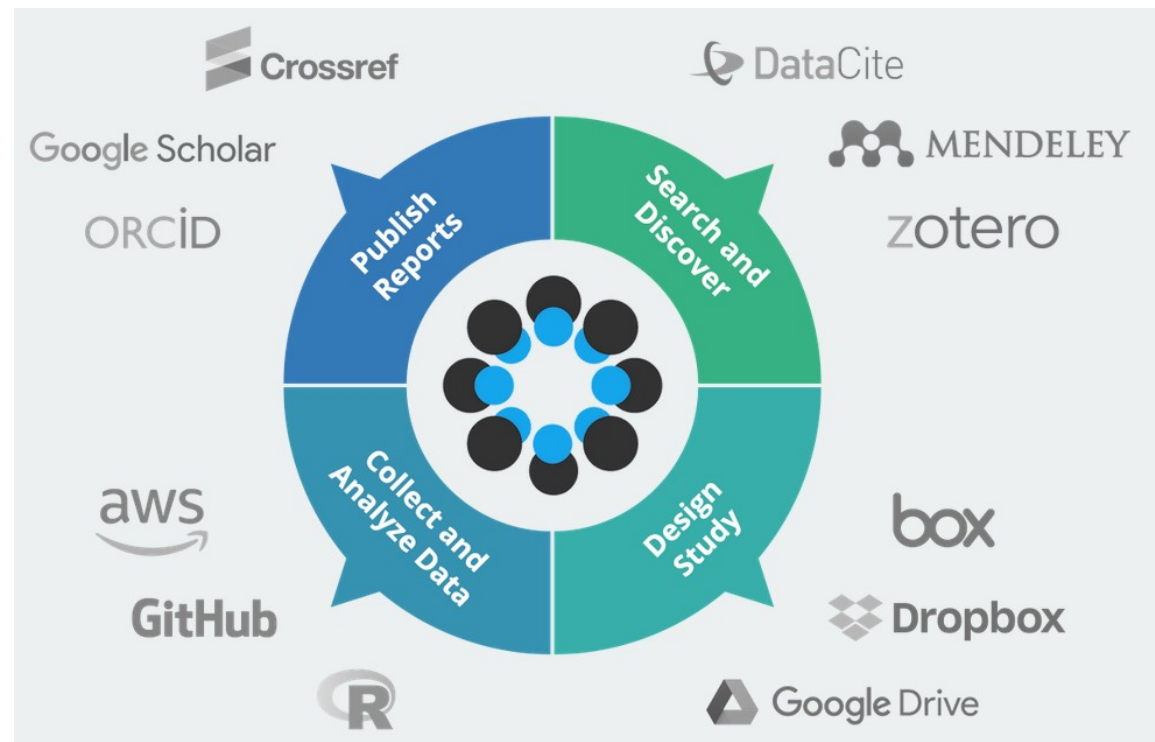
General Article

## The New Statistics: Why and How

Geoff Cumming  
La Trobe University

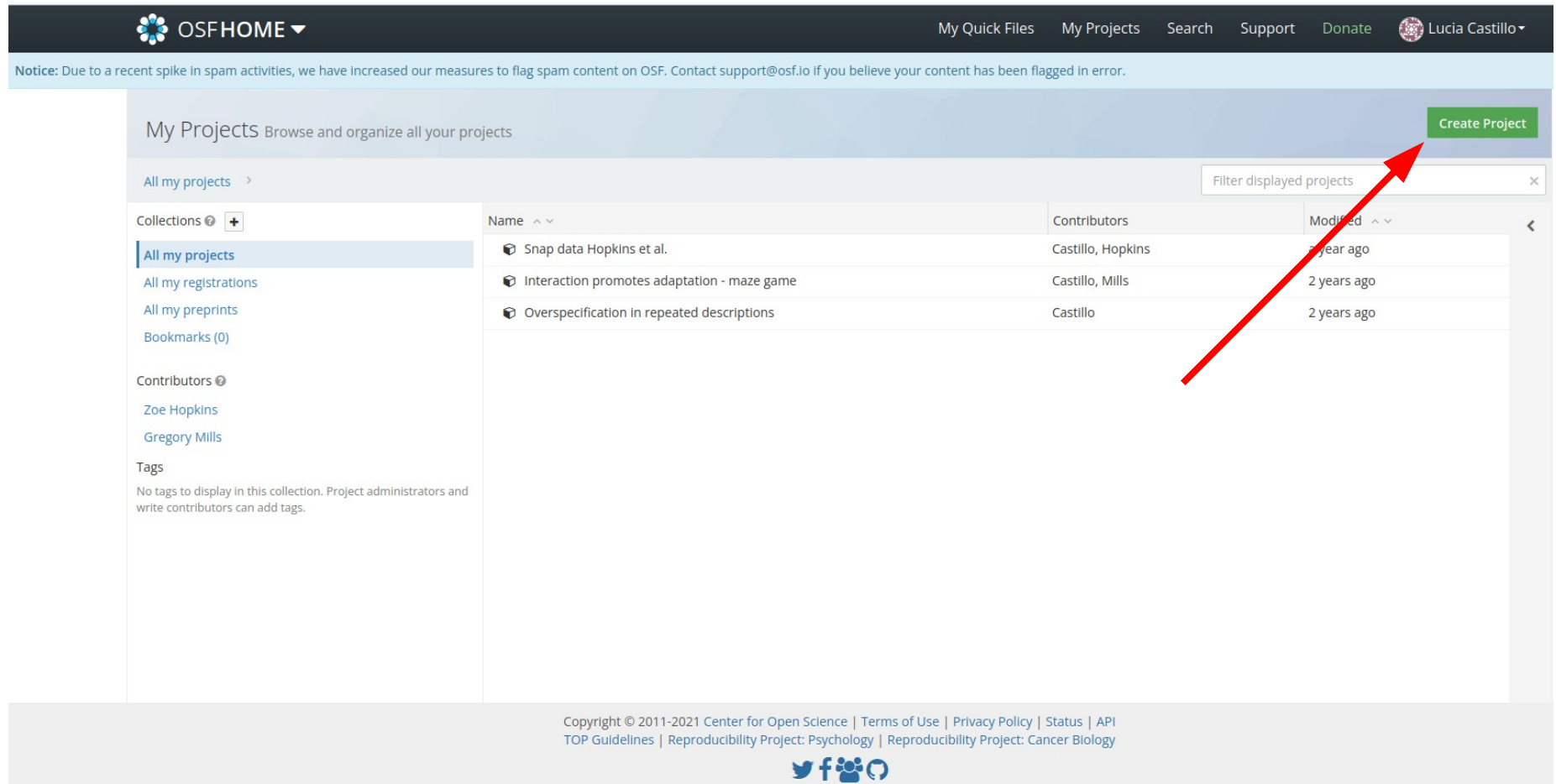
# Pasos concretos hacia una cultura de ciencia abierta

- Incorporación de software y programación de acceso abierto desde el pregrado
  - **Uso de datos reales**
  - Uso de software abierto (R, OpenSesame, etc.)
  - Uso de *notebooks* para trabajo colaborativo (R Markdown, Jupyter)
  - Cursos transversales de ética de los datos y ciencia abierta



# Pasos concretos hacia una cultura de ciencia abierta

- Dar el primer paso!



The screenshot shows the OSFHOME interface. At the top, there is a navigation bar with the OSFHOME logo, user profile 'Lucia Castillo', and links for 'My Quick Files', 'My Projects', 'Search', 'Support', and 'Donate'. A notice below the navigation bar states: 'Notice: Due to a recent spike in spam activities, we have increased our measures to flag spam content on OSF. Contact support@osf.io if you believe your content has been flagged in error.'

The main content area is titled 'My Projects Browse and organize all your projects'. On the right side of this header, there is a green 'Create Project' button, which is highlighted by a red arrow. Below the header, there is a search bar labeled 'Filter displayed projects' and a list of projects. The list has columns for 'Name', 'Contributors', and 'Modified'. The projects listed are:

Name	Contributors	Modified
📁 Snap data Hopkins et al.	Castillo, Hopkins	5 year ago
📁 Interaction promotes adaptation - maze game	Castillo, Mills	2 years ago
📁 Overspecification in repeated descriptions	Castillo	2 years ago

On the left side, there are sections for 'Collections' (with a plus icon), 'Contributors' (listing Zoe Hopkins and Gregory Mills), and 'Tags' (with a note: 'No tags to display in this collection. Project administrators and write contributors can add tags.').

At the bottom of the page, there is a footer with copyright information: 'Copyright © 2011-2021 Center for Open Science | Terms of Use | Privacy Policy | Status | API TOP Guidelines | Reproducibility Project: Psychology | Reproducibility Project: Cancer Biology' and social media icons for Twitter, Facebook, and GitHub.



# Recursos

- **Open Science Framework**

<http://osf.io>

- **The Turing Way**

<https://the-turing-way.netlify.com>

- **Data Skills for Reproducible Science**

[https://gupsych.github.io/data\\_skills/](https://gupsych.github.io/data_skills/)

- **R for Reproducible Scientific Analysis**

<http://swcarpentry.github.io/r-novice-gapminder/>

- **Datacamp: Reporting with R Markdown**

<https://www.datacamp.com/courses/reporting-with-r-markdown> 21 / 21