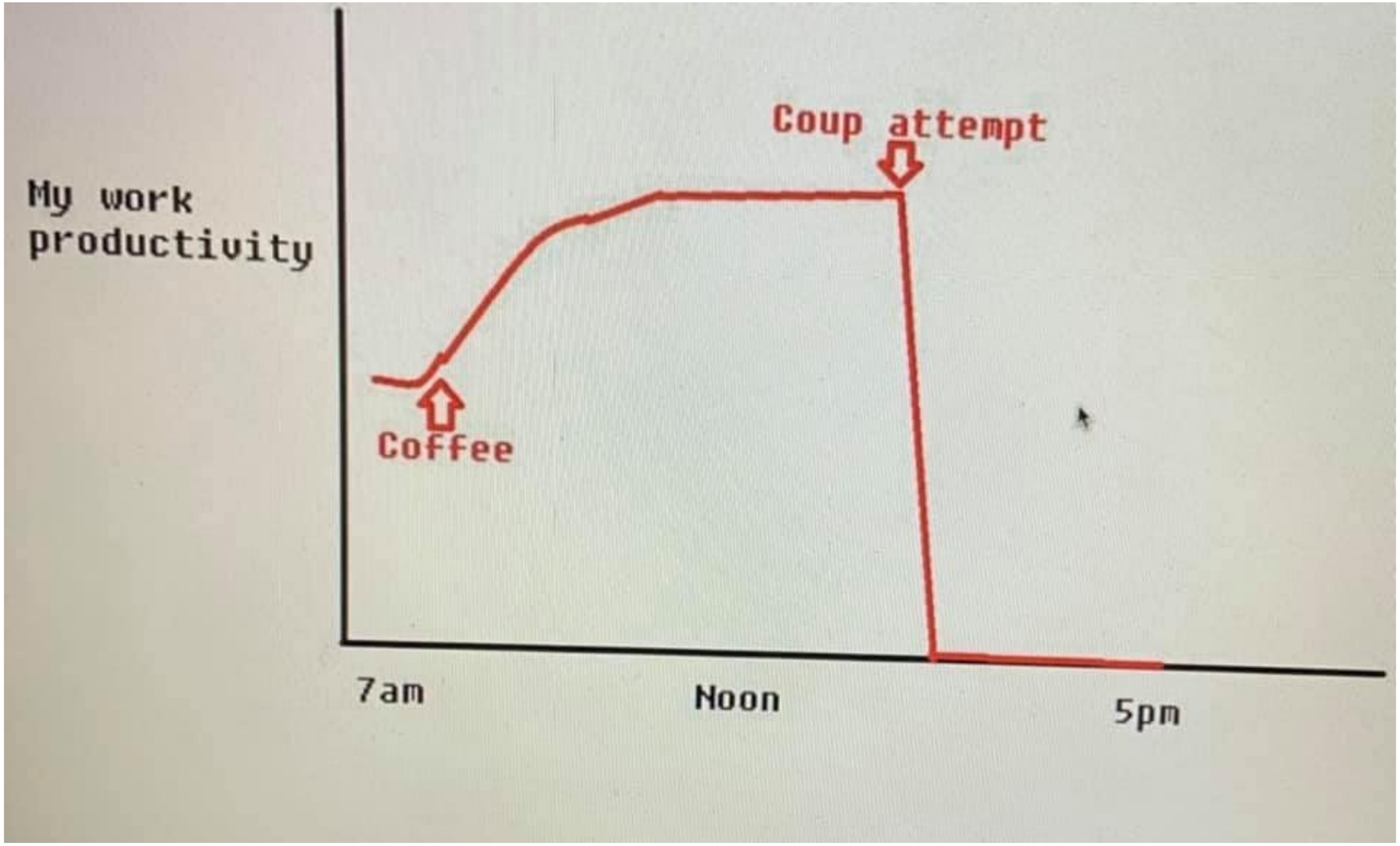

My work productivity
Coup attempt
Coffee
7am
Noon
5pm

Slides at:
dakotamurray.me/talk/2021-northwestern/

Scientific evaluation in context

Dakota Murray
@dakotasmurray
dakota.s.murray@gmail.com

SCHOOL OF
INFORMATICS, COMPUTING, AND ENGINEERING

Who was the best scientist?

Who was the best scientist?

Marie Curie

Far-reaching theoretical and experimental contributions

Who was the best scientist?

Marie Curie

Far-reaching theoretical and experimental contributions

Leonardo Da Vinci

Broad contributions to science, but also art and engineering

Who was the best scientist?

Marie Curie

Far-reaching theoretical and experimental contributions

Leonardo Da Vinci

Broad contributions to science, but also art and engineering

Zacharias Janssen

A spectacle-maker credited with inventing the microscope

How do we compare such diverse contributions?

Marie Curie

Far-reaching theoretical and experimental contributions

Leonardo Da Vinci

Broad contributions to science, but also art and engineering

Zacharias Janssen

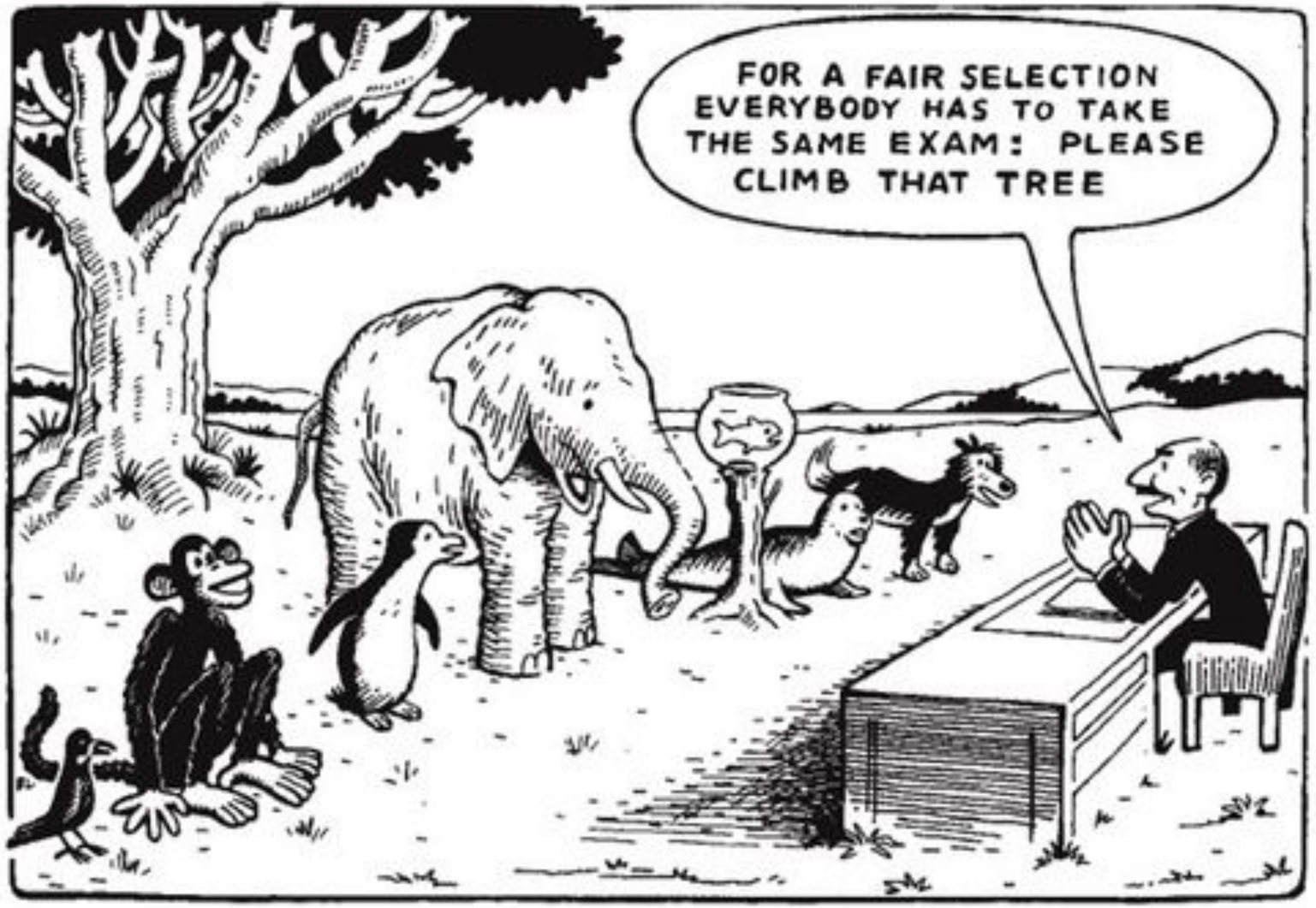
A spectacle-maker credited with inventing the microscope

Identifying excellence is essential to science

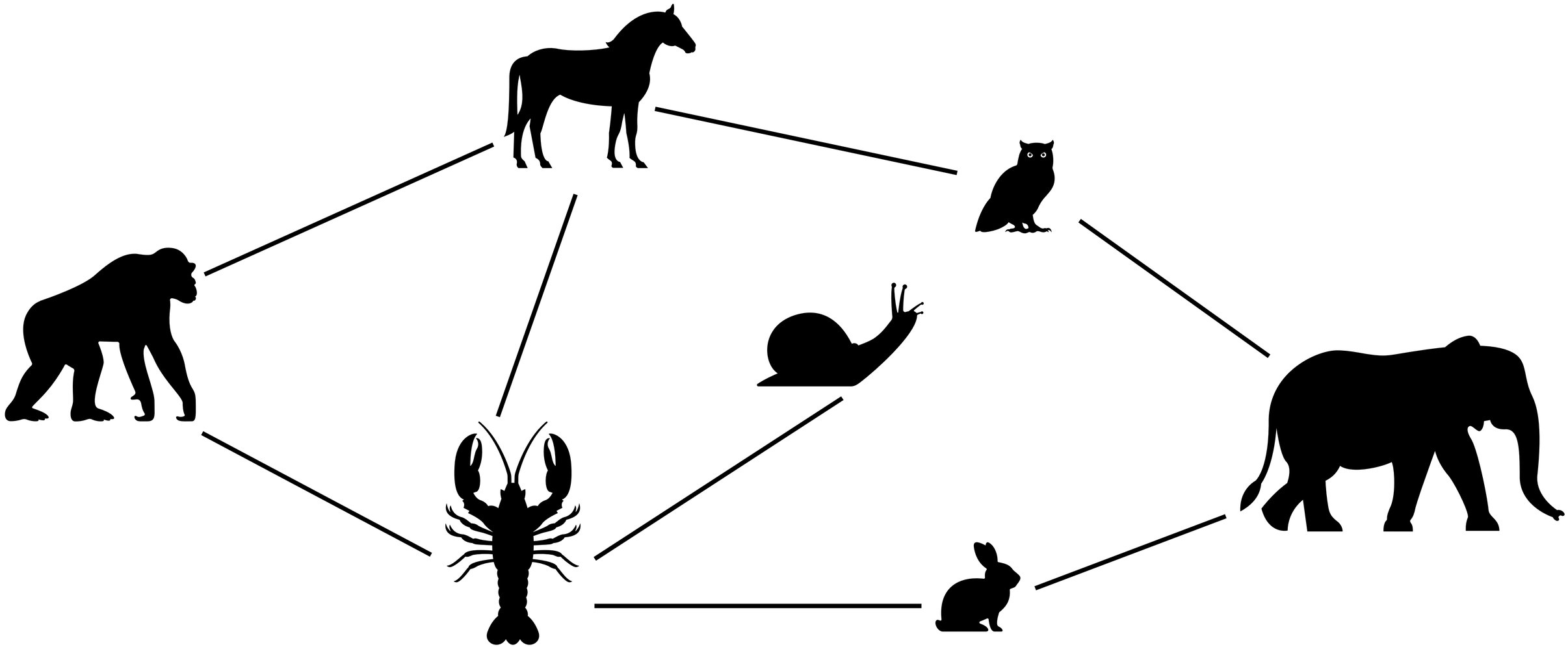
Everything from grant funding, to publication, to the Nobel prize

Evaluating performance in science is hard!

We end up judging people with limited criteria

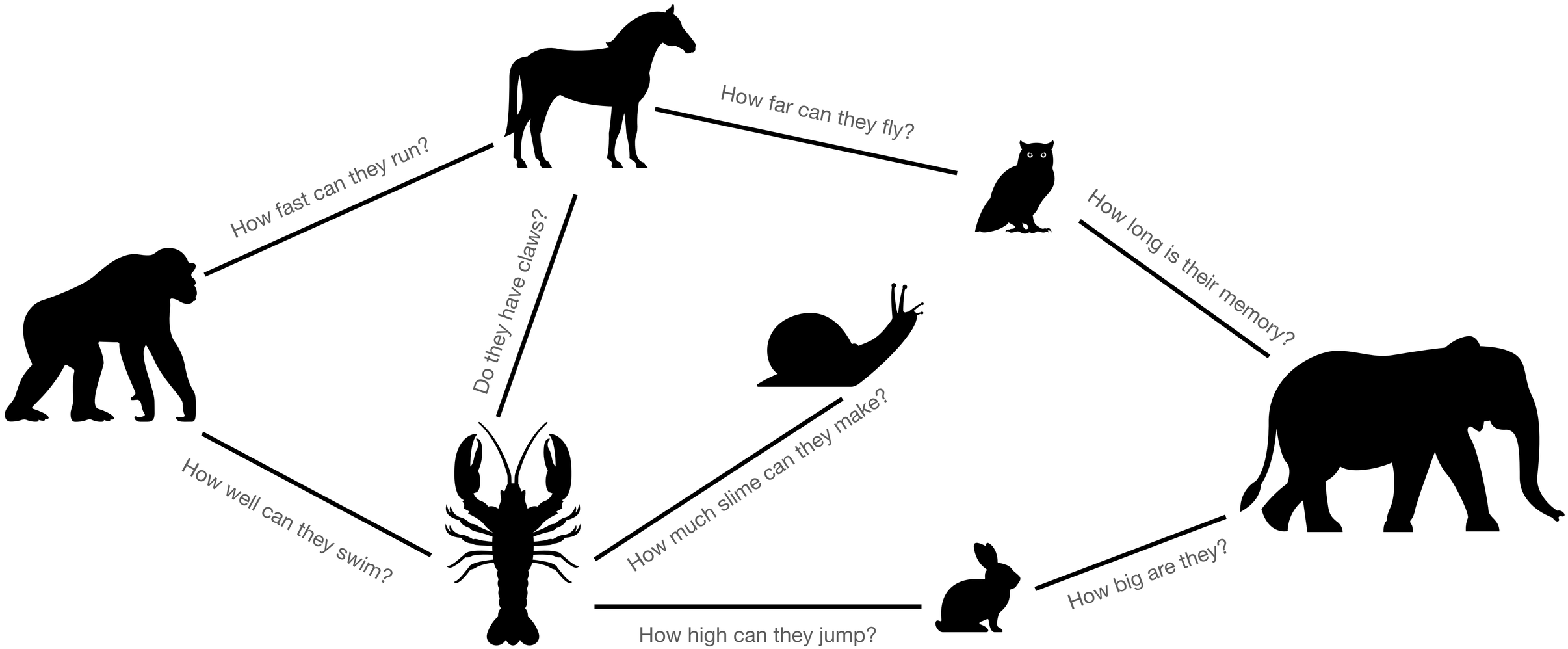


Scientific evaluation is inherently contextual



Scientific evaluation is inherently contextual

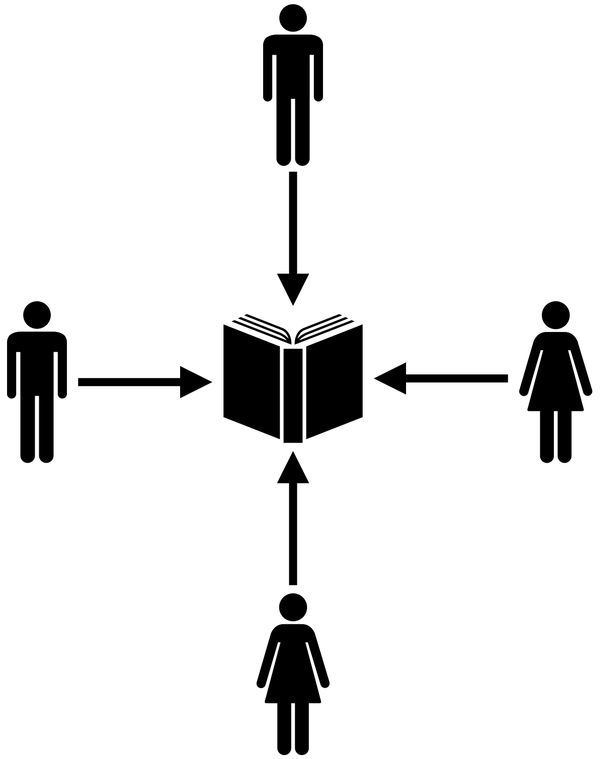
We all bring our own expectations, cultures, and biases to judge others



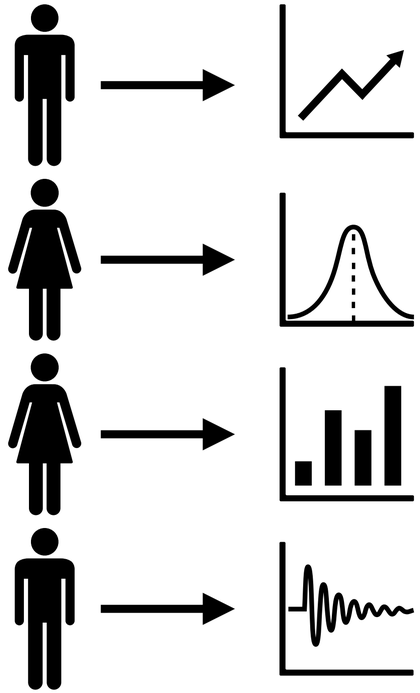
Evaluation in science

Hiring, promotion, publication, funding, and reputation

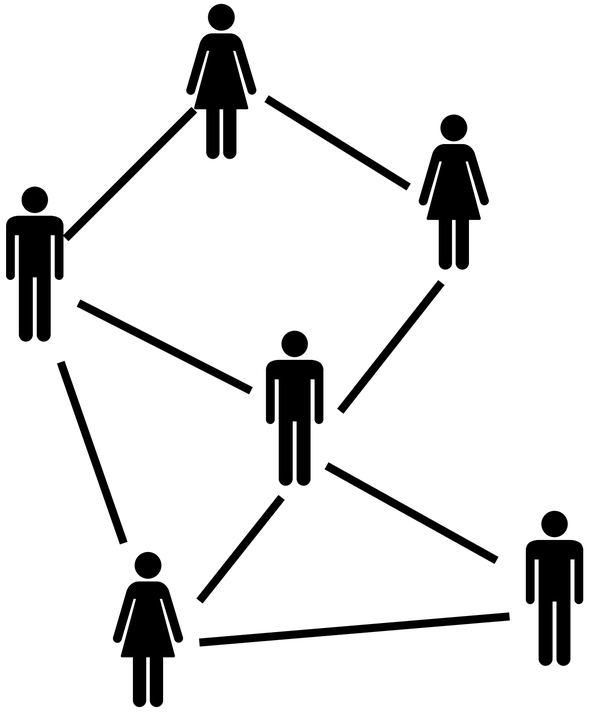
Peer review



Performance metrics



Reputation
(networks)



Ideally, evaluation should capture true merit

Ideally, evaluation should capture true merit

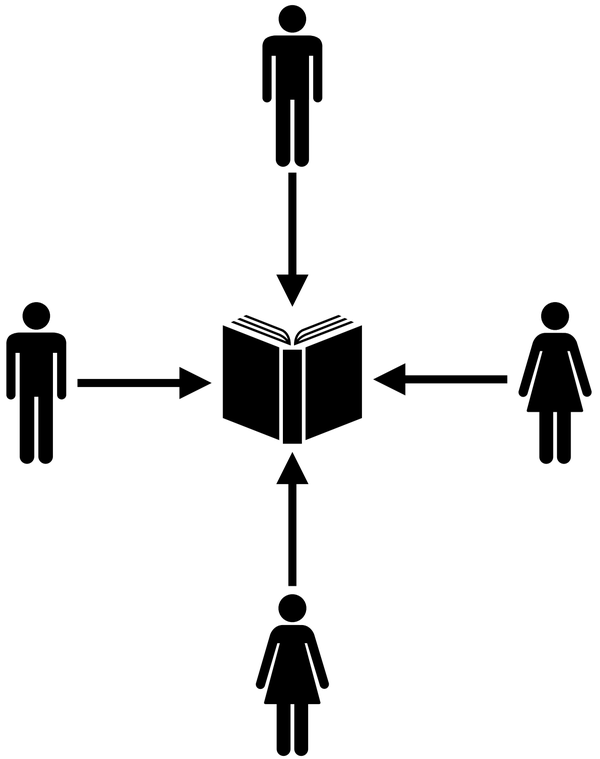
But its embedded in a social, political, and cultural context

My research direction:

Investigate the contextual factors that drive scientific evaluation

5 studies examining different areas of evaluation

Peer review

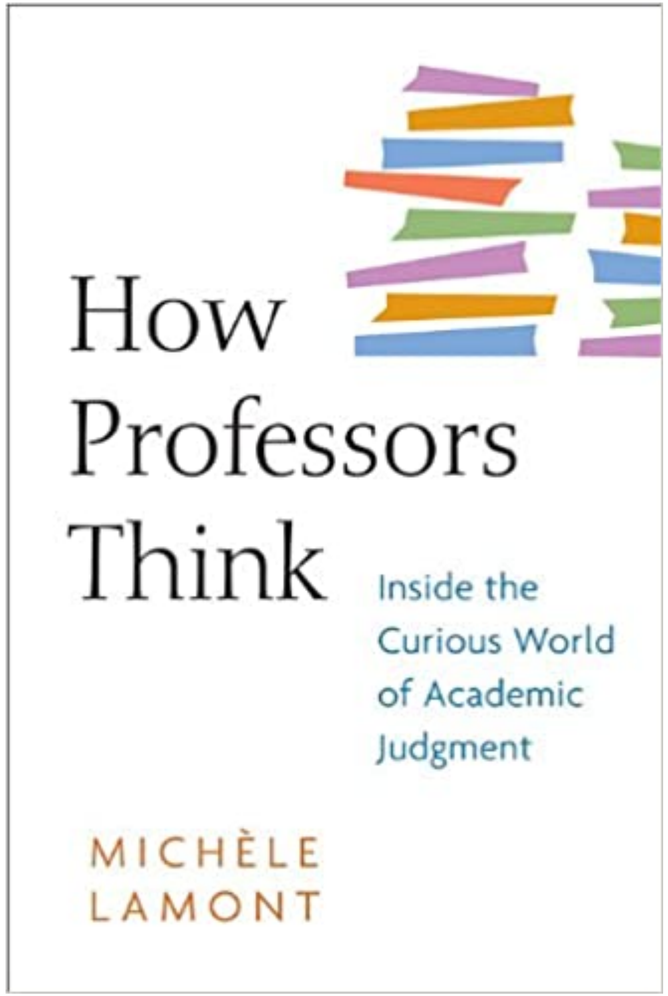


Peer evaluation

Sensitive to contextual factors

- Personal relationships
- Taste preferences
- Disciplinary cultures
- Biases and prejudices

- *"…ultimately, reasoned judgements are buffered by unpredictable human proclivities, agency, and improvisation". (pg 201)*

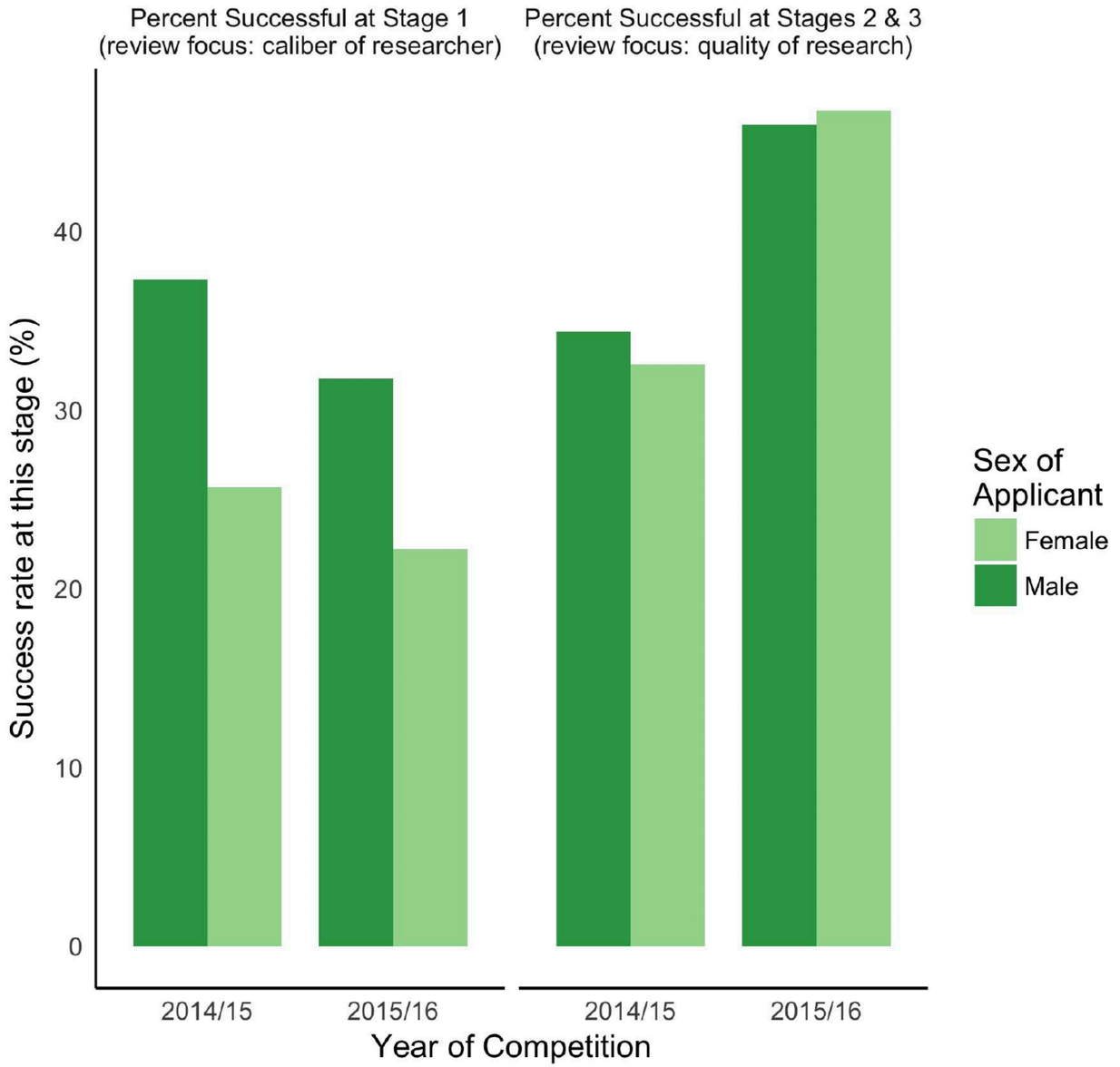


Lamont, M. (2009). *How Professors Think: Inside the Curious World of Academic Judgment*. Harvard University Press.

Gender bias

Grant peer review at CIHR

- 2-stage review process

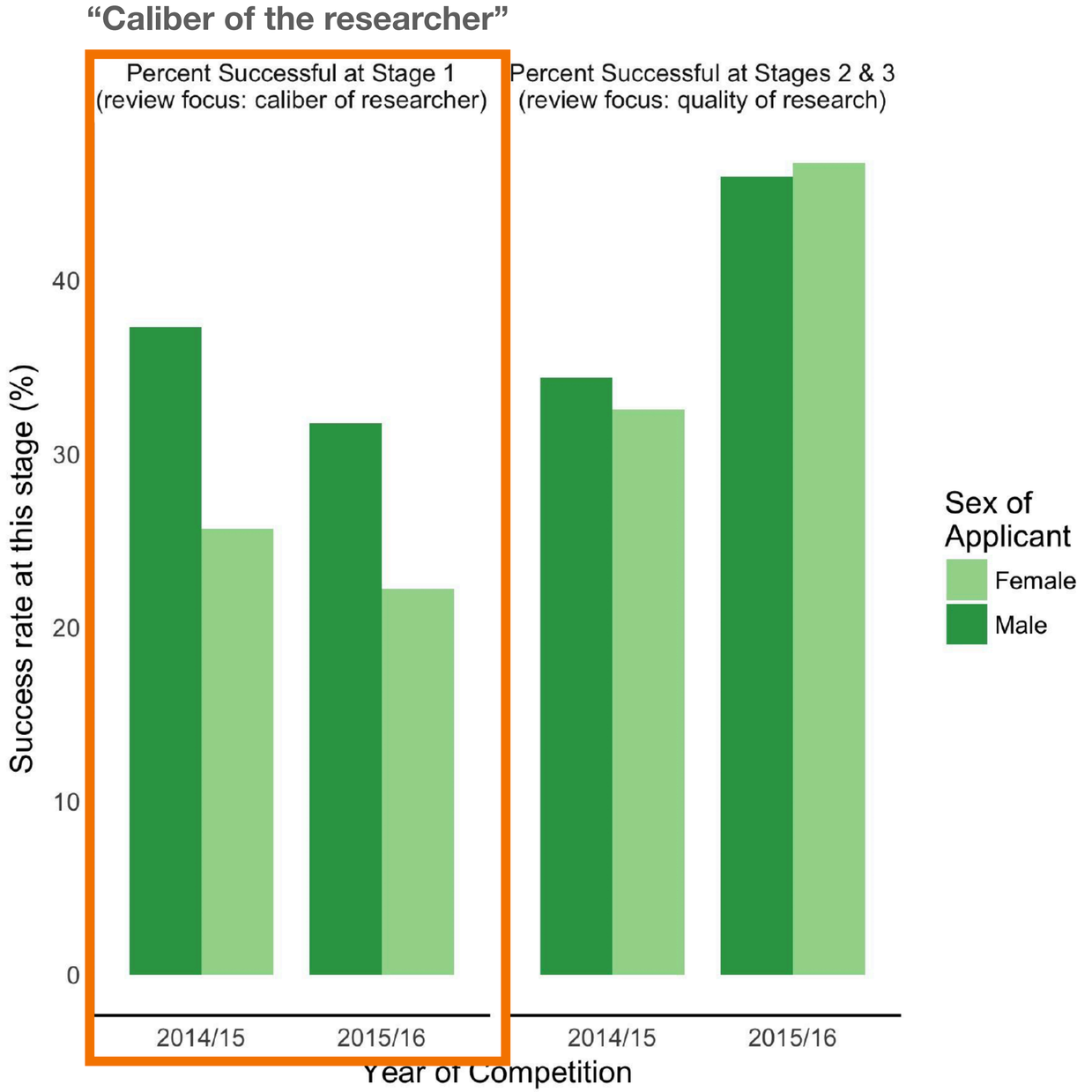


Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2017). Female grant applicants are equally successful when peer reviewers assess the science, but not when they assess the scientist. *BioRxiv*, 232868.

Gender bias

Grant peer review at CIHR

- 2-stage review process
- Women rated lower in “Caliber of researcher”

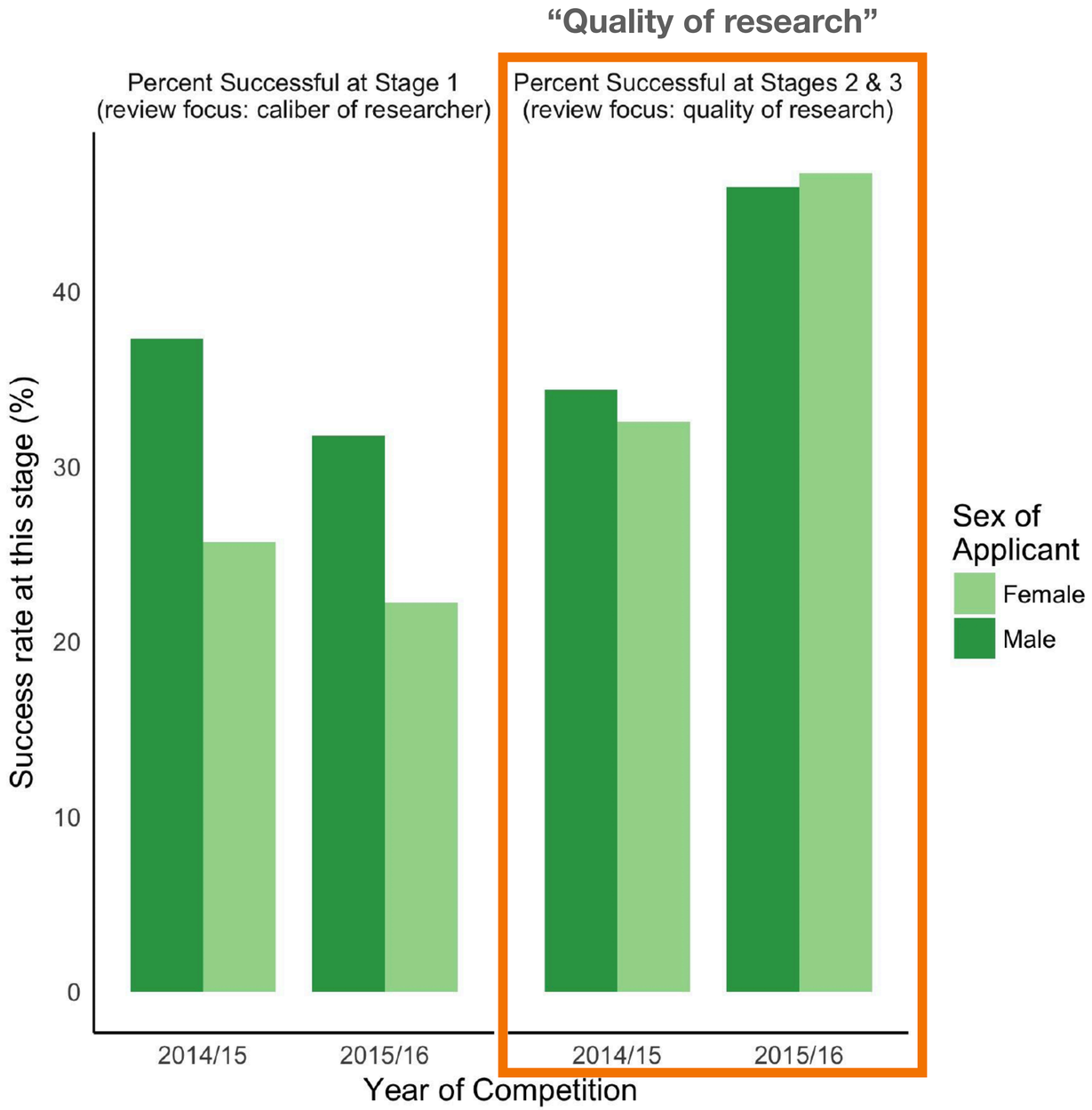


Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2017). Female grant applicants are equally successful when peer reviewers assess the science, but not when they assess the scientist. *BioRxiv*, 232868.

Gender bias

Grant peer review at CIHR

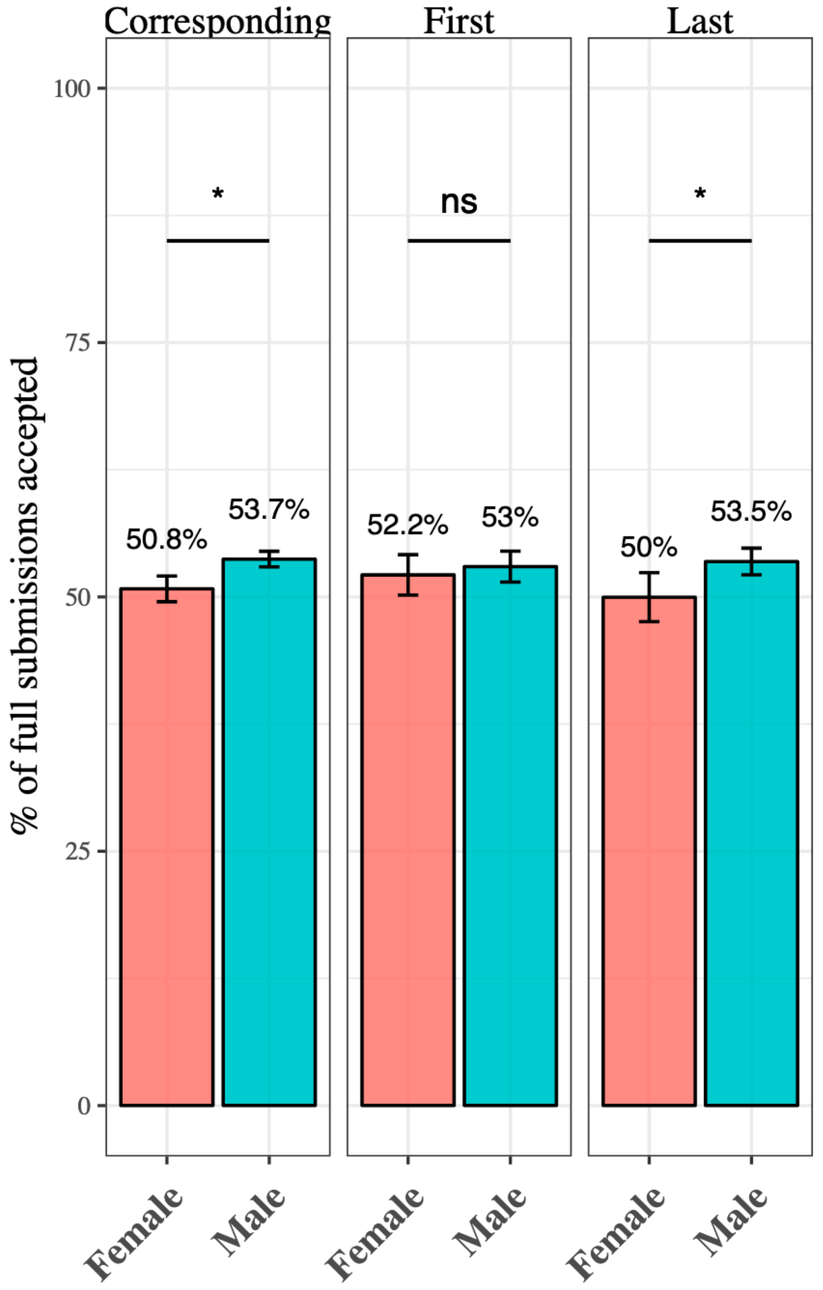
- 2-stage review process
- Women rated lower in “Caliber of researcher”
- Gender equity when rating on “quality of research”



Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2017). Female grant applicants are equally successful when peer reviewers assess the science, but not when they assess the scientist. *BioRxiv*, 232868.

Journal peer review at eLife

Women's papers accepted less

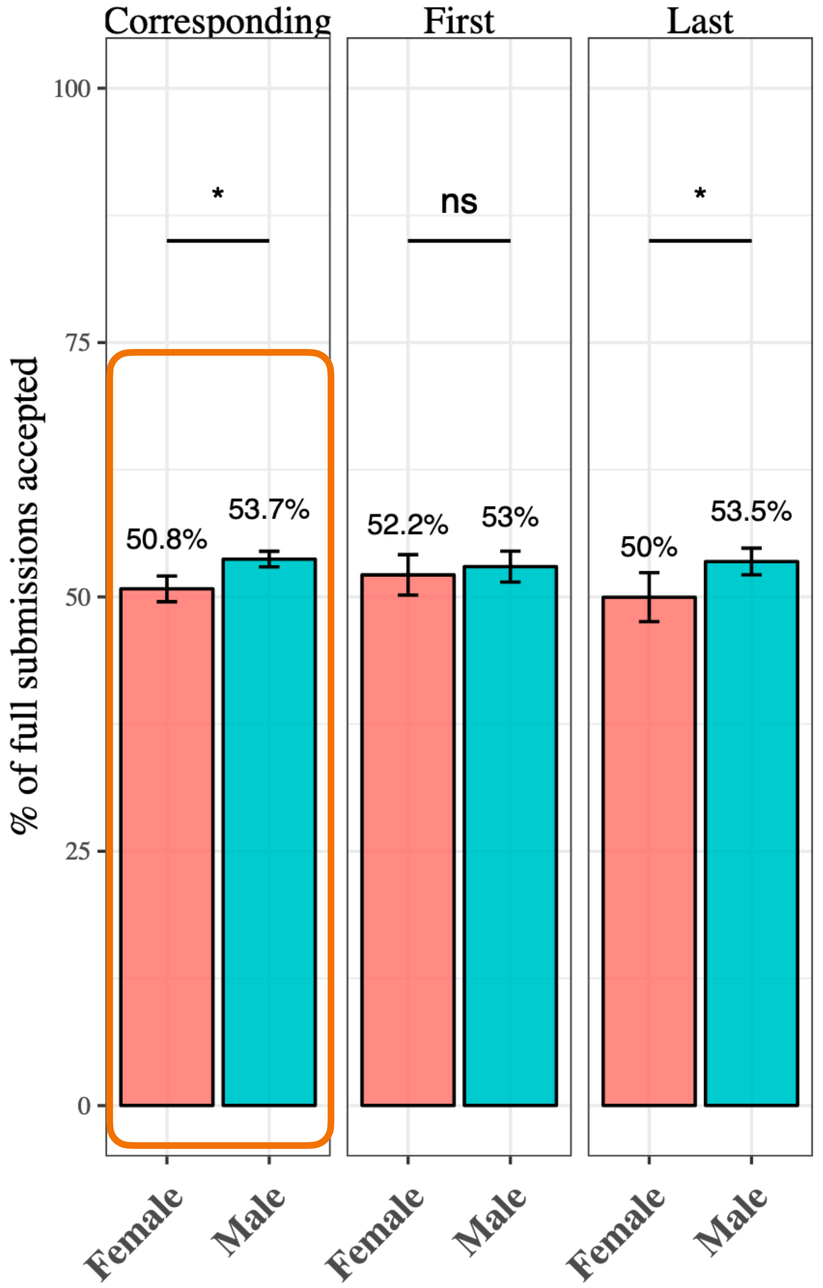


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

Women's papers accepted less

- 2.9 % point gender disparity for corresponding authorship

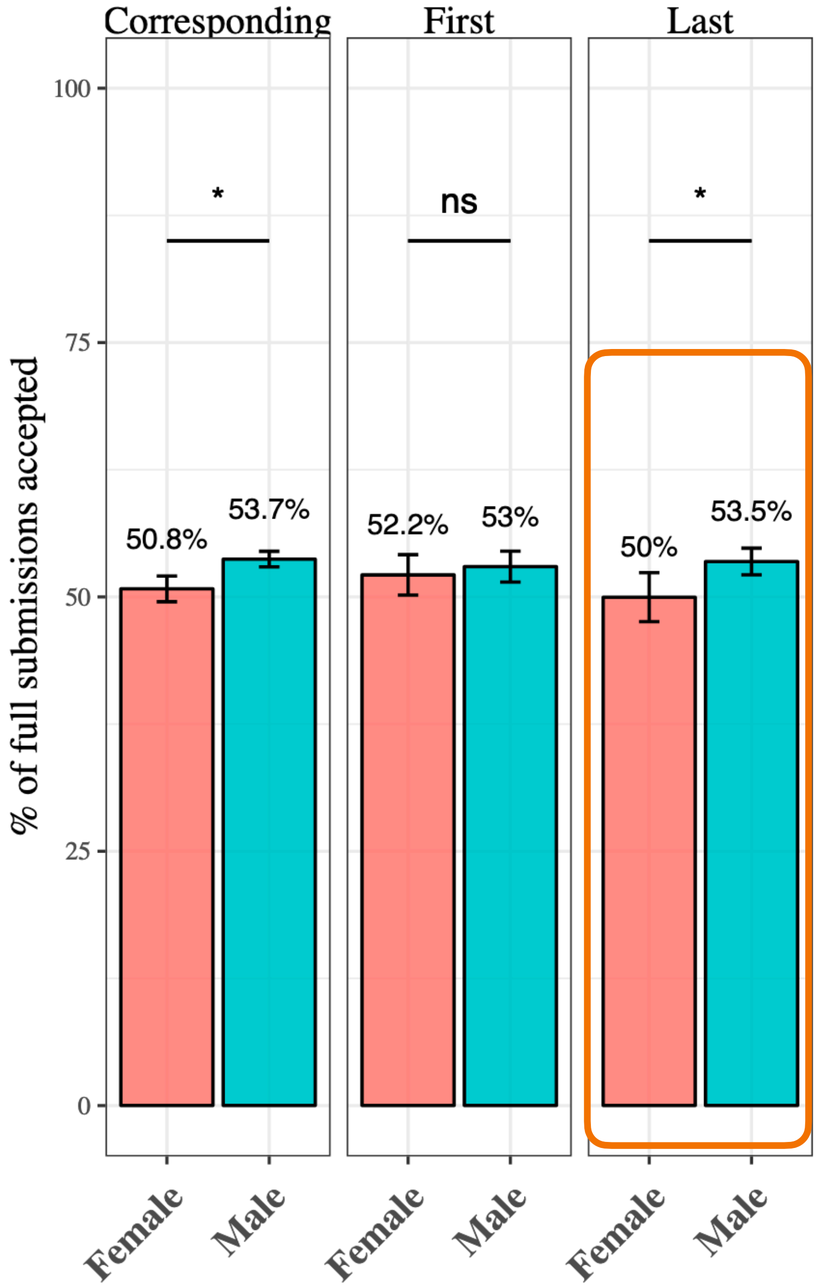


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

Women's papers accepted less

- 2.9 % point gender disparity for corresponding authorship
- 3.5% among last authors

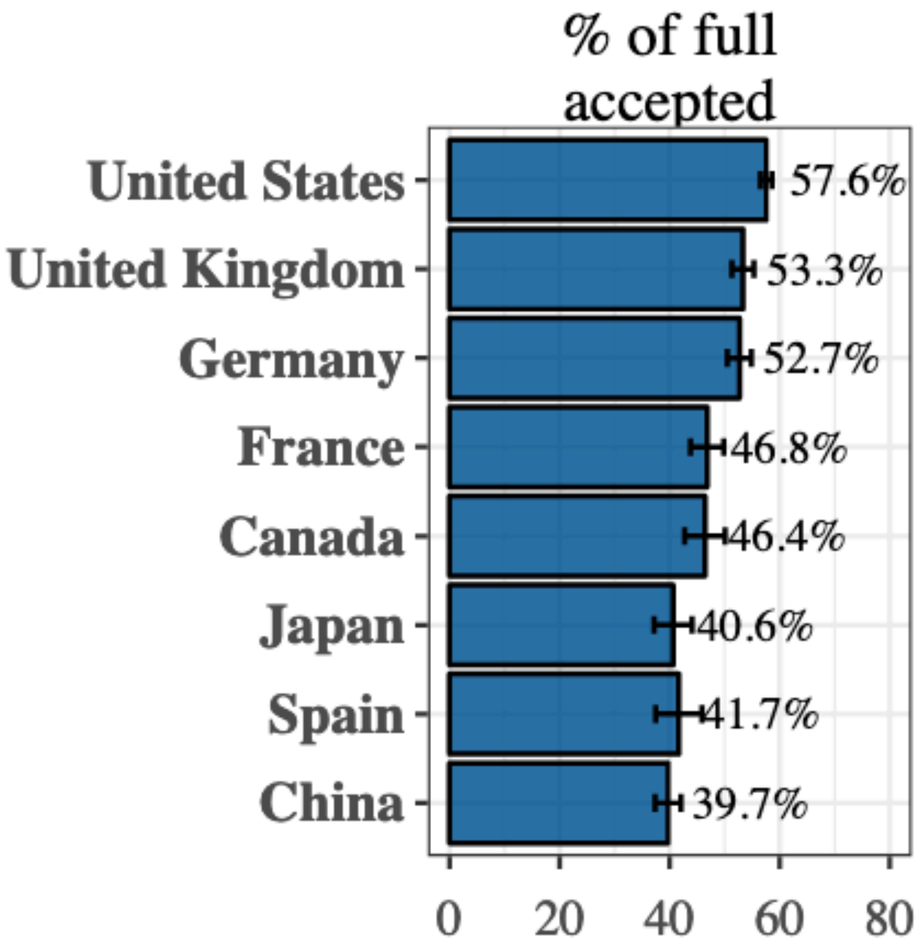


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

U.S. papers accepted more

- Country of corresponding author

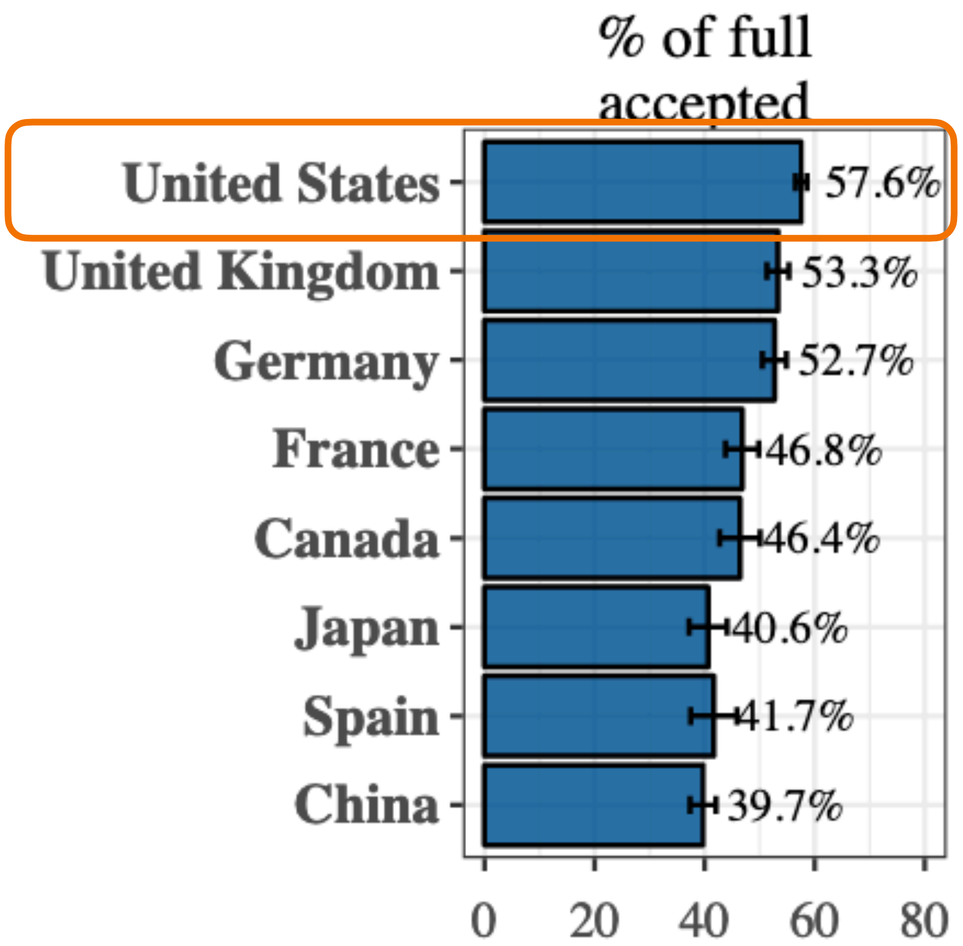


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

U.S. papers accepted more

- Country of corresponding author
- U.S. papers had the highest acceptance rates
- Beating our even other major scientific countries



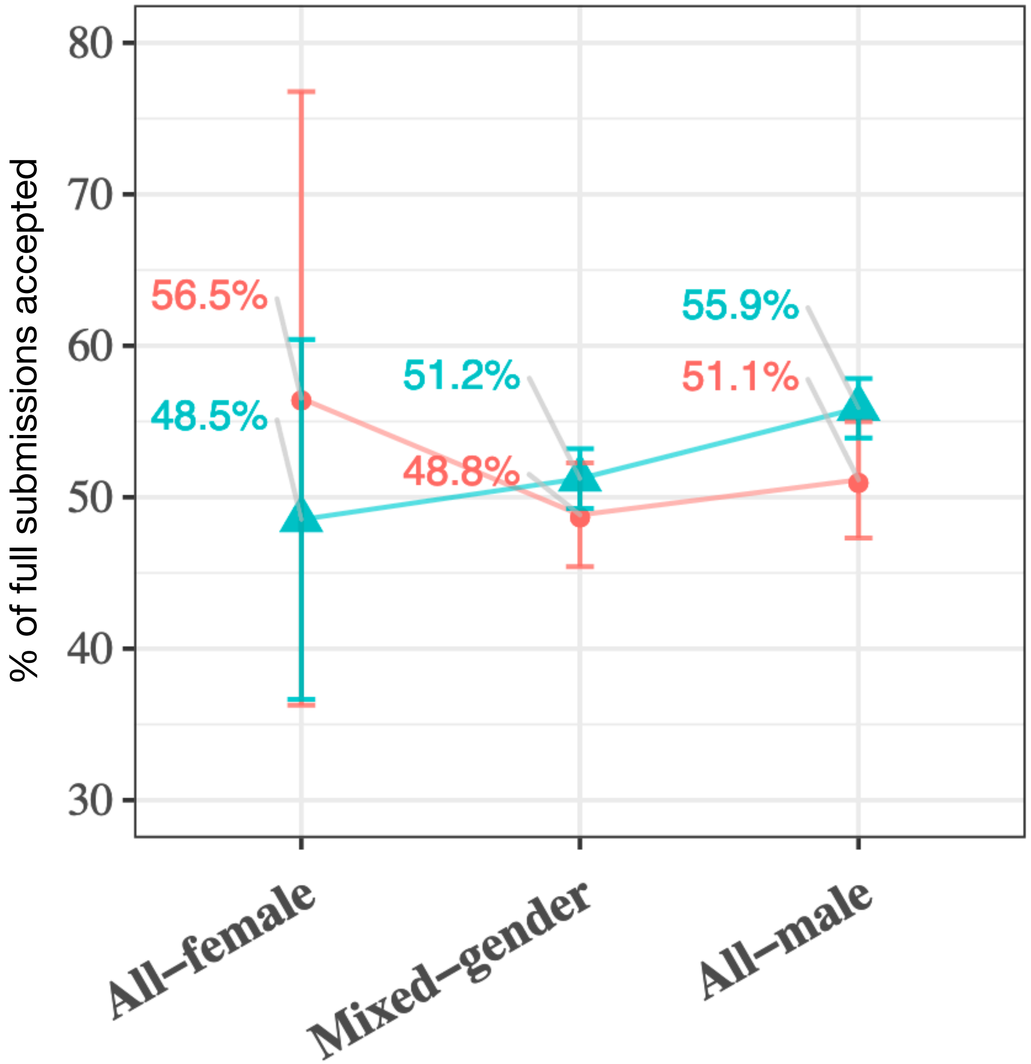
Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Bias in peer review?

What can be done to mitigate bias in peer review?

Reviewers matter!

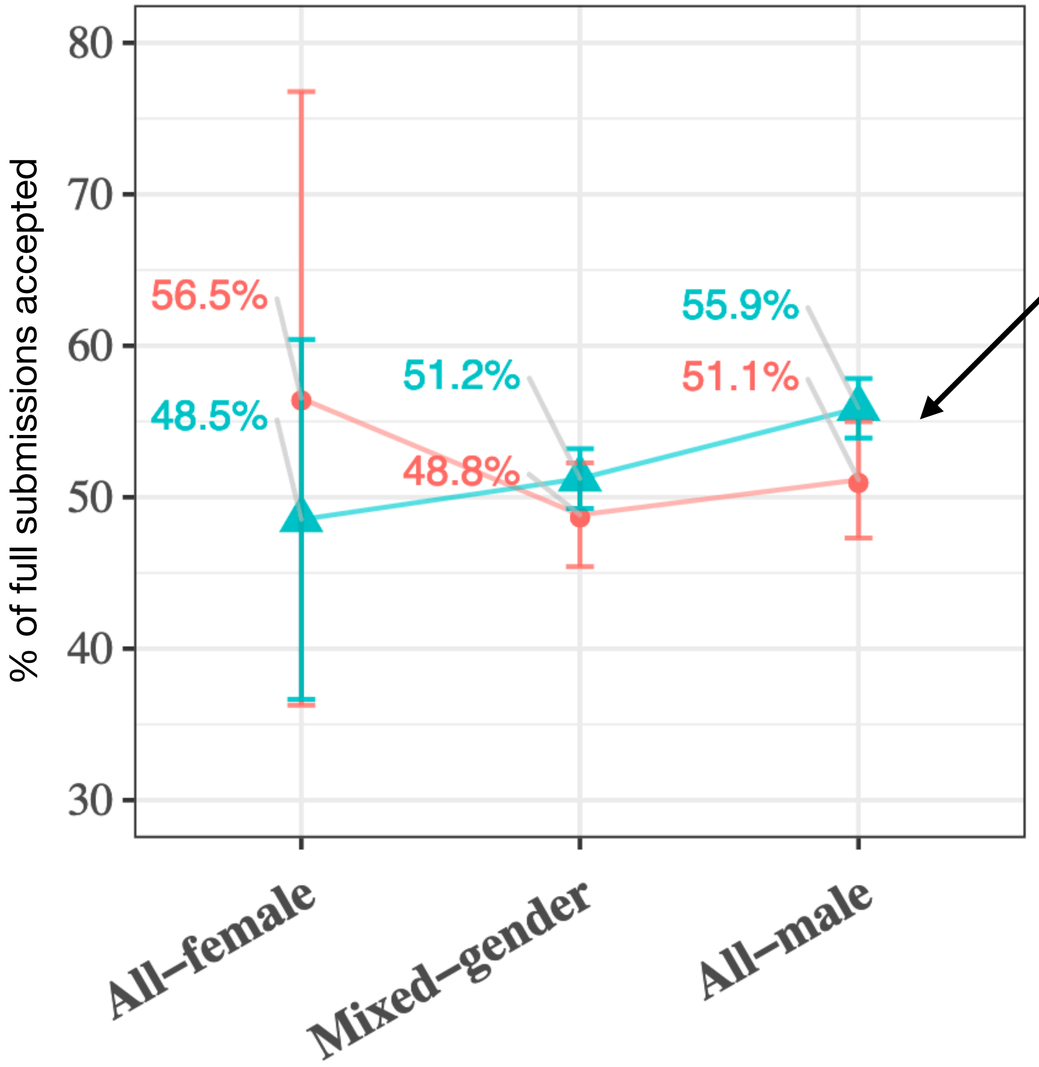
Disparities differ by the composition of the reviewer group



Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Reviewers matter!

Disparities differ by the composition of the reviewer group

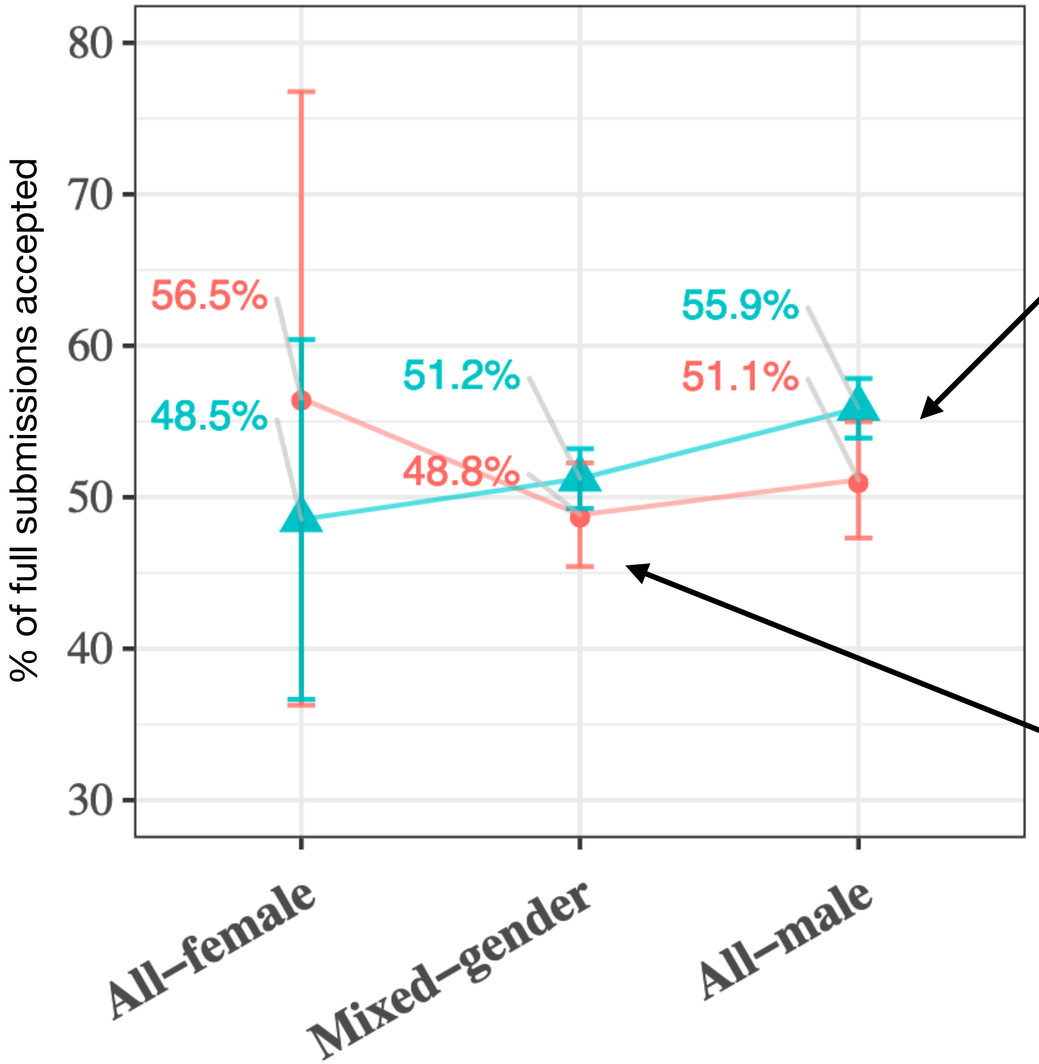


4.8% disparity for all-male teams

Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Reviewers matter!

Disparities differ by the composition of the reviewer group



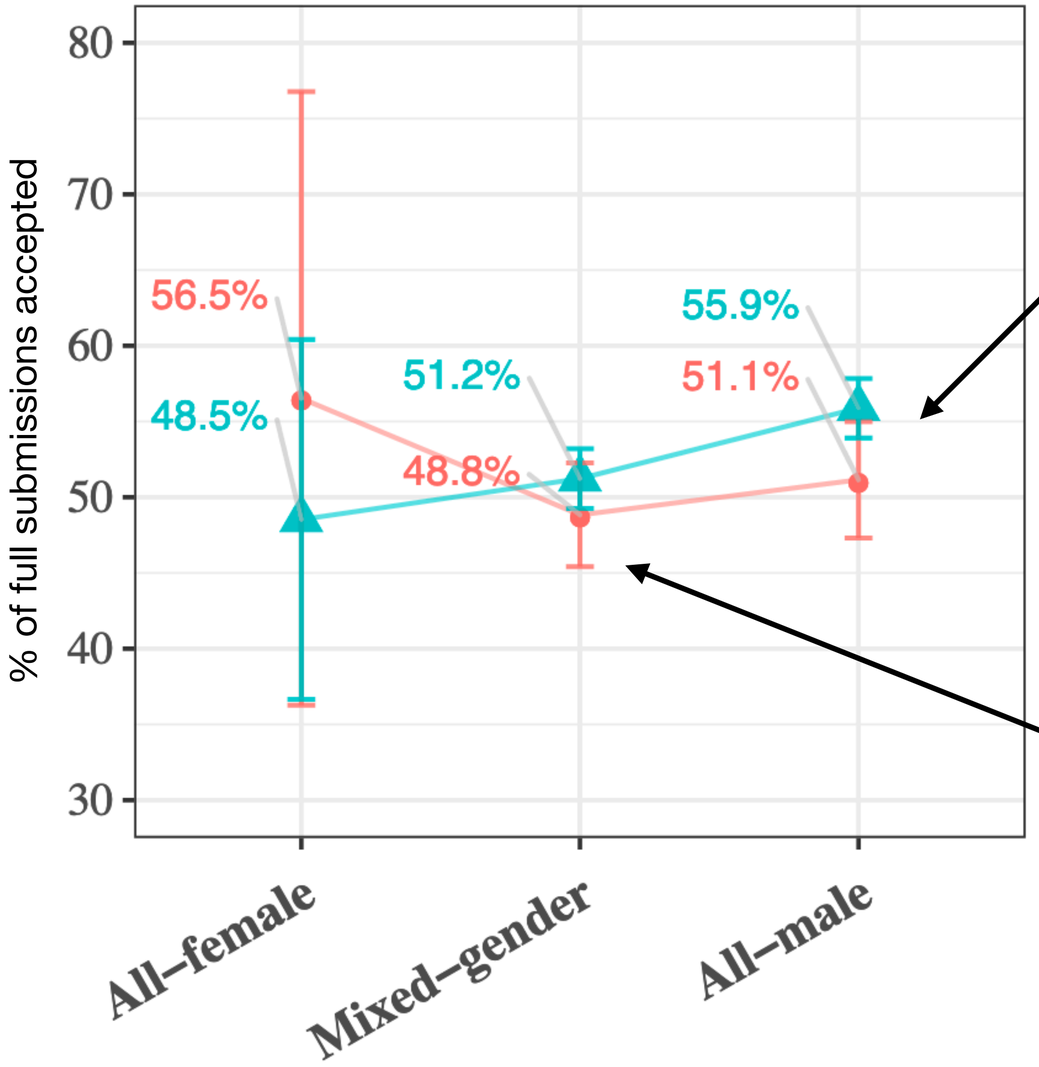
4.8% disparity for all-male teams

2.4% disparity for mixed-gender teams

Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Reviewers matter!

Disparities differ by the composition of the reviewer group



4.8% disparity for all-male teams

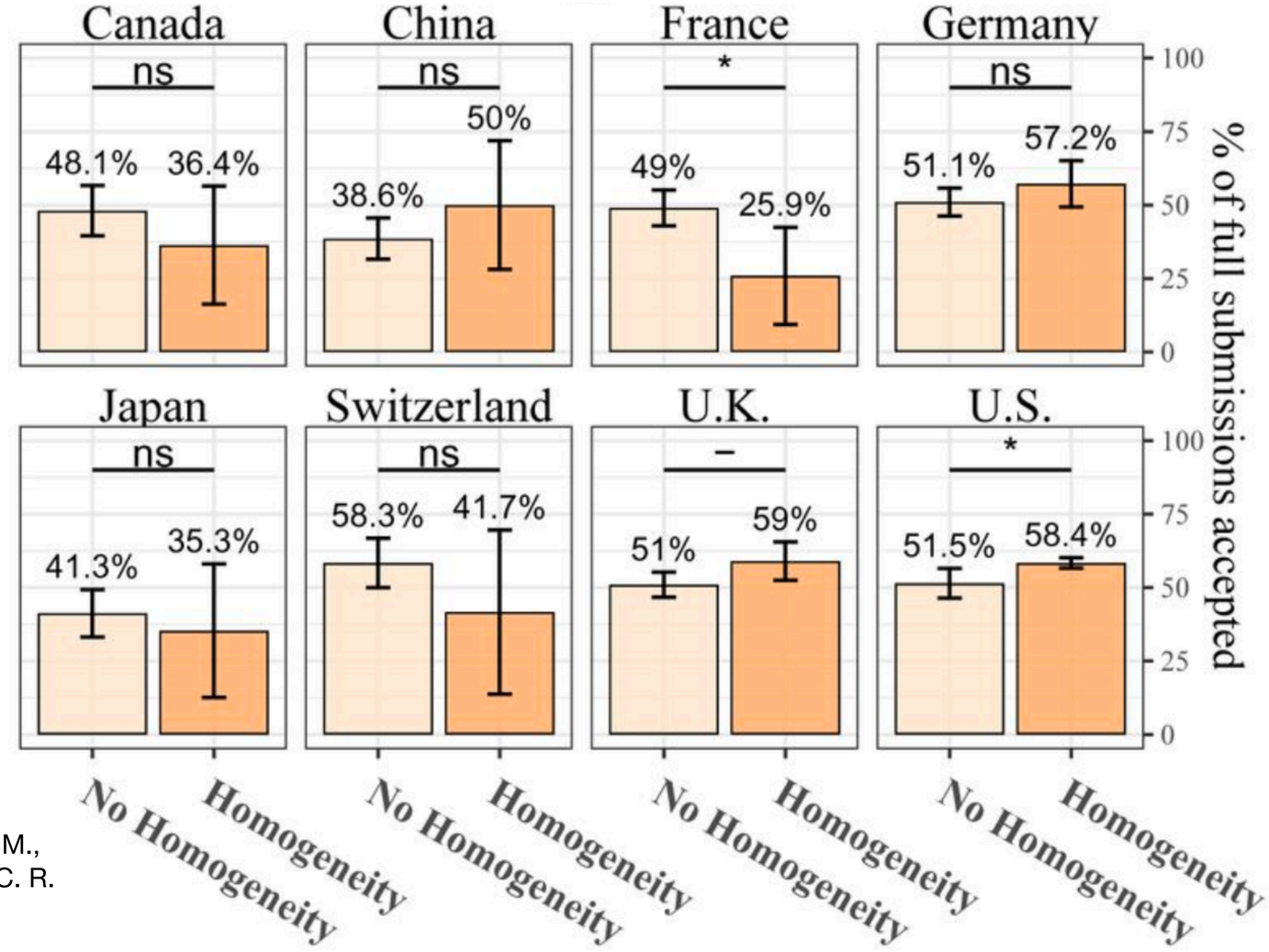
Evidence (all be it marginal) that mixed-gender reviewer teams may produce more equitable outcomes

2.4% disparity for mixed-gender teams

Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Reviewer nationality also matters

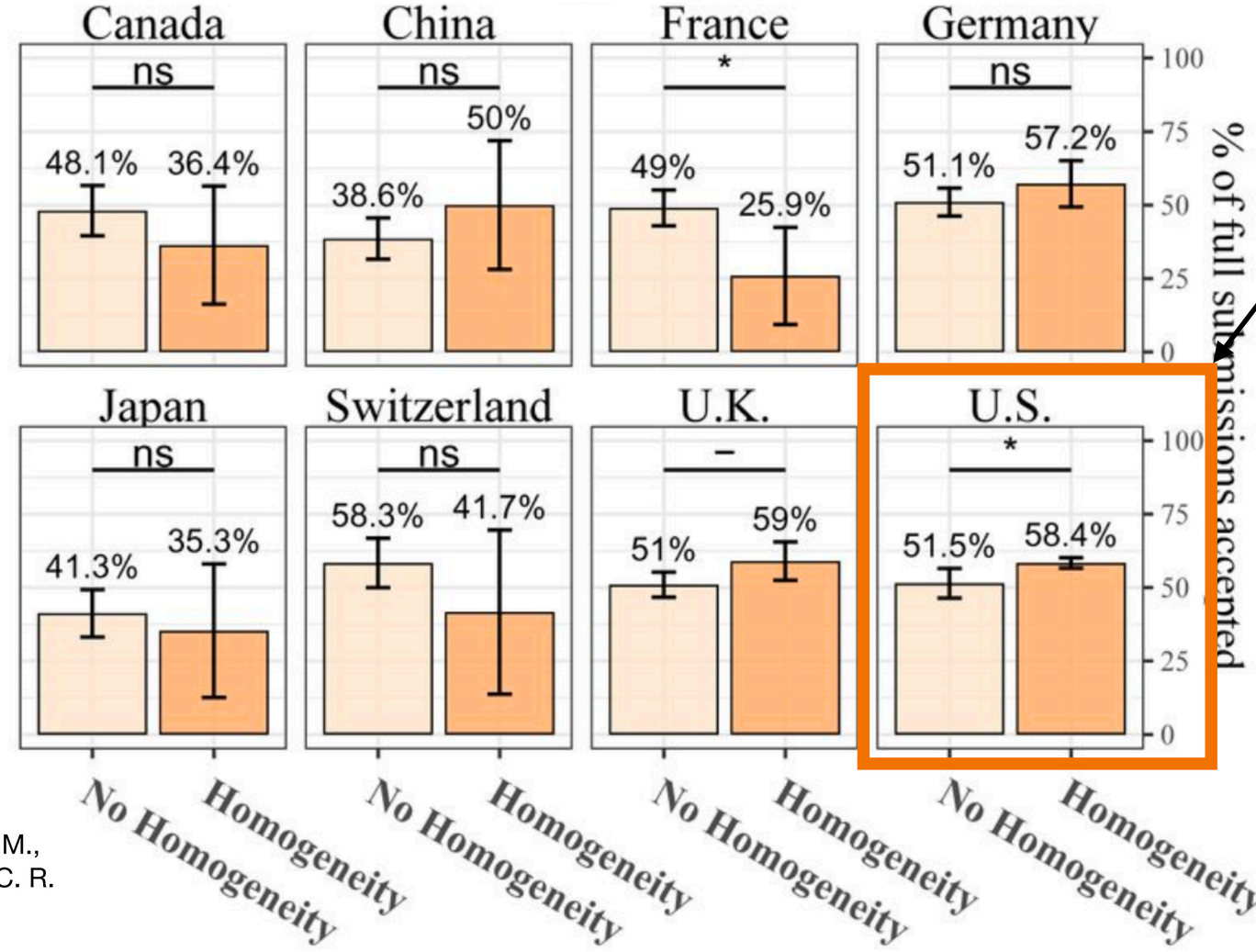
Review outcomes based on homogeny between author and reviewers



Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Reviewer nationality also matters

Review outcomes based on homogeny between author and reviewers

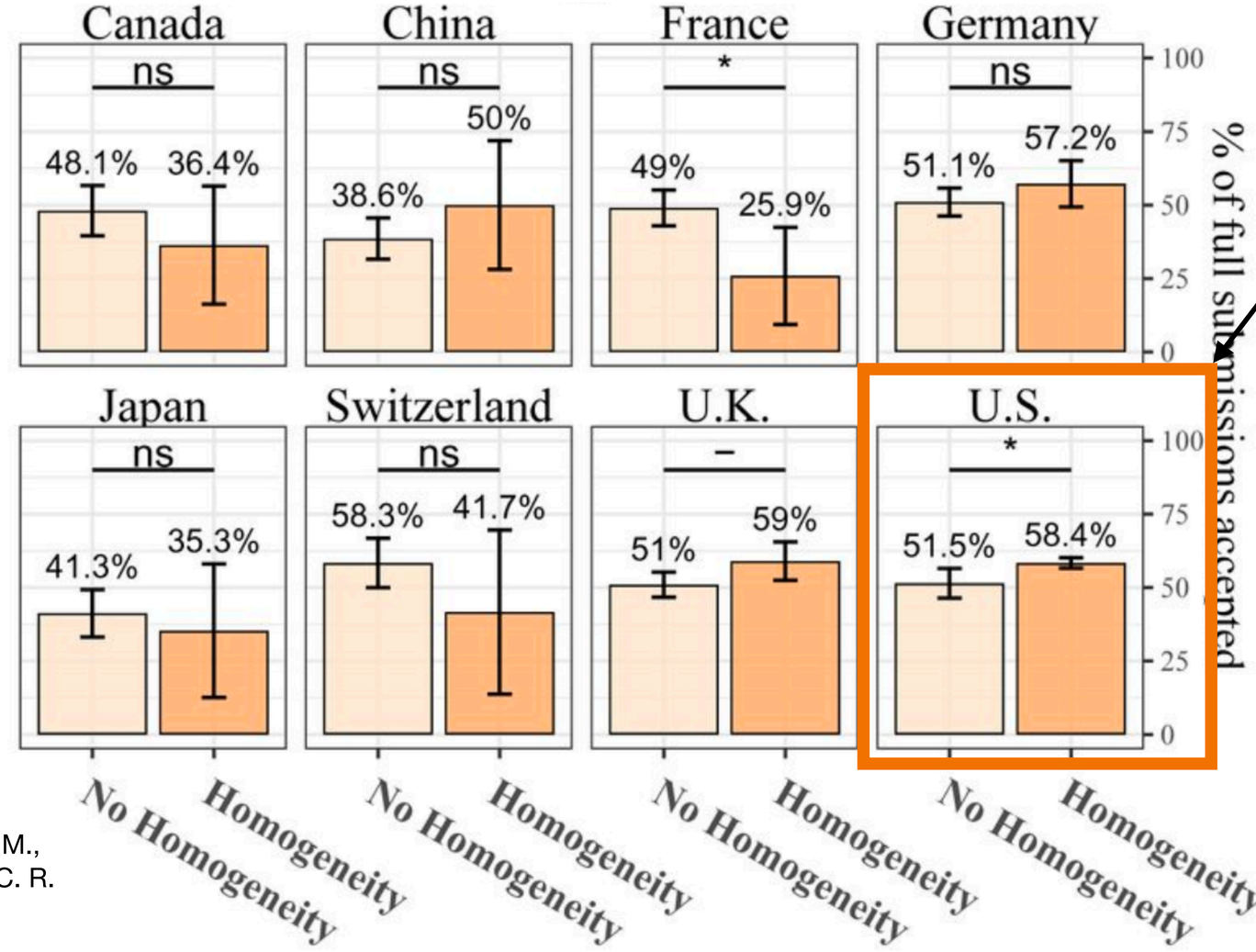


U.S. authors reviewed by U.S. reviewer accepted more often

Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Reviewer nationality also matters

Review outcomes based on homogeny between author and reviewers



U.S. authors reviewed by U.S. reviewer accepted more often

A benefit not always experienced by other nationalities

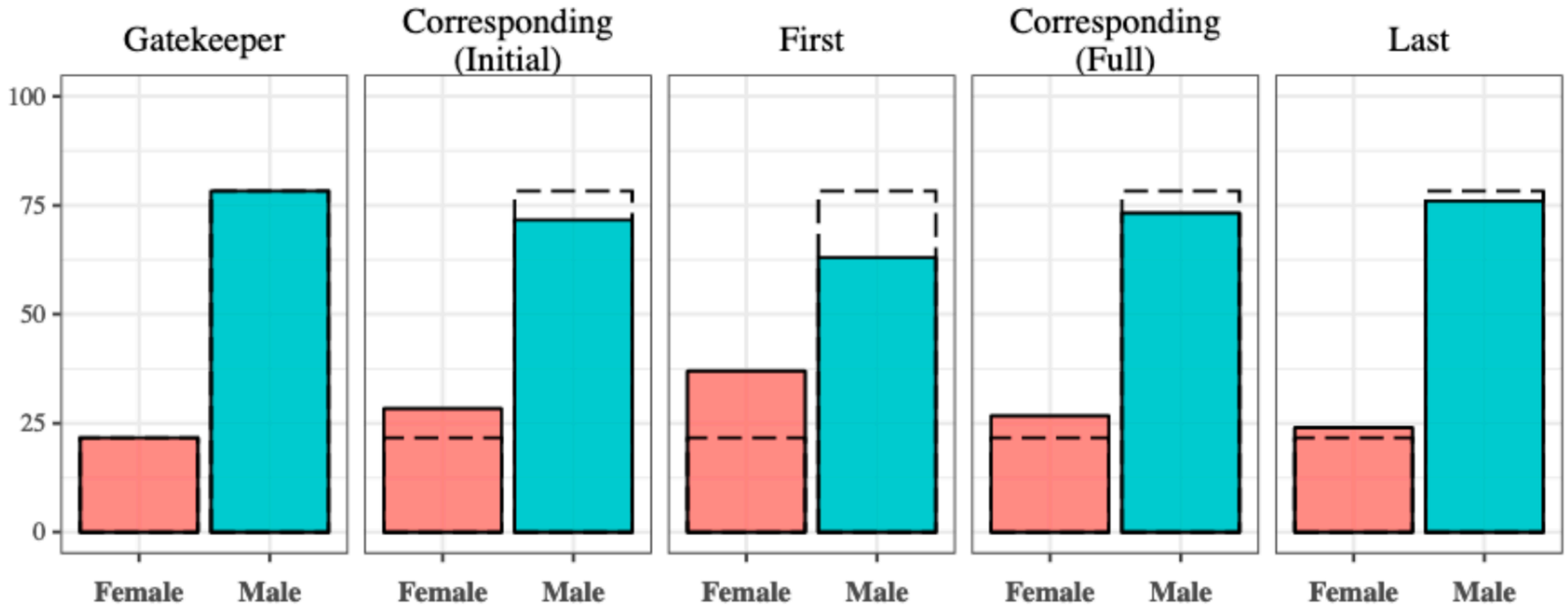
Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

A simple, low risk solution:

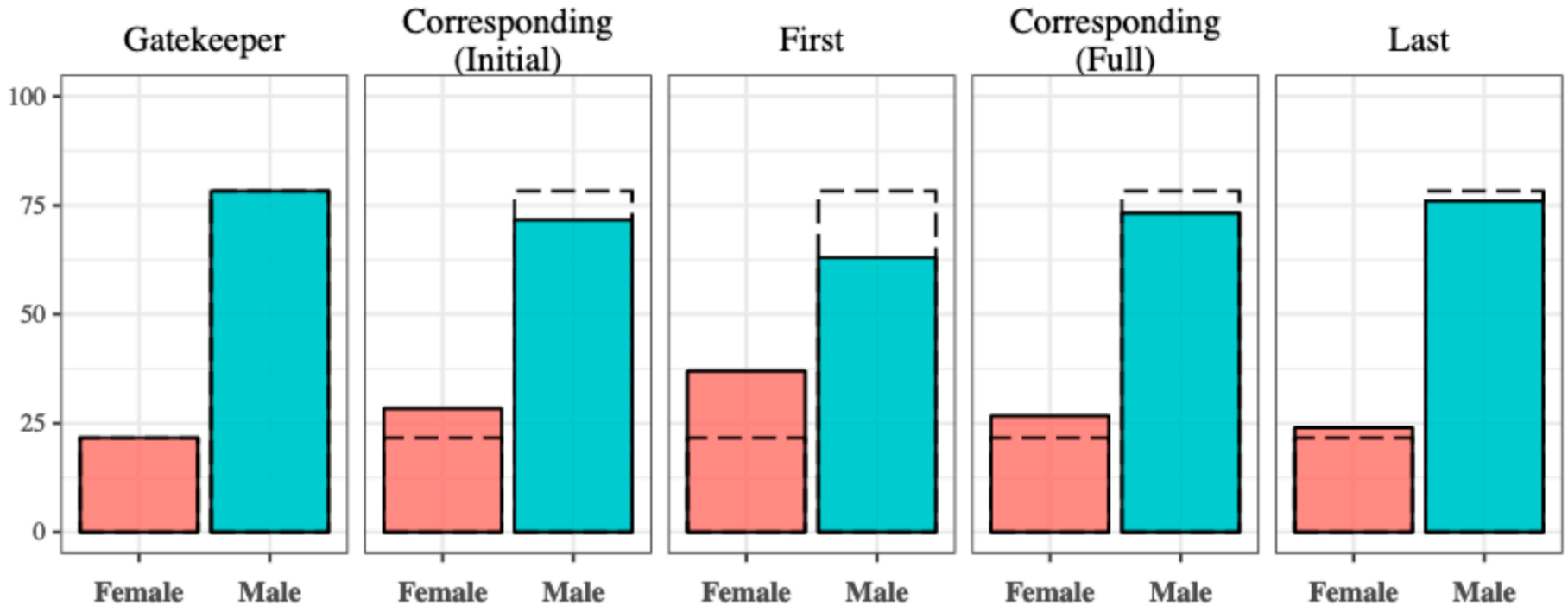
A simple, low risk solution:

Make gatekeepers more diverse!

Make reviewers represent the authorship

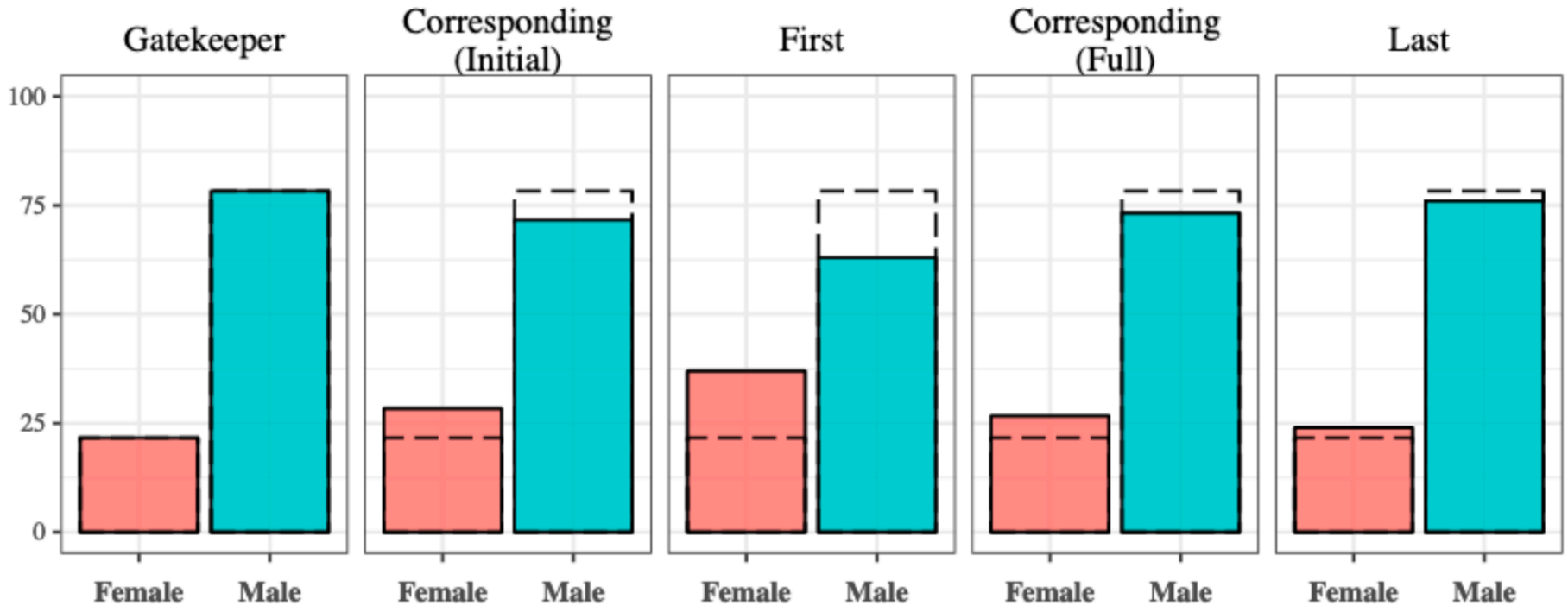


Make reviewers represent the authorship



This is something that eLife does well!

Make reviewers represent the authorship



This is something that eLife does well!

Only about half of papers were reviewed by mixed-gender teams

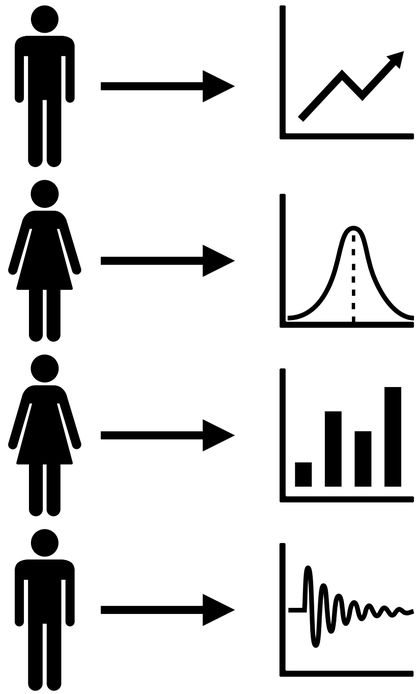
eLife is currently experimenting with policies to improve diversity and equity in peer review

Peer review is subject to contextual factors

Peer review is subject to contextual factors

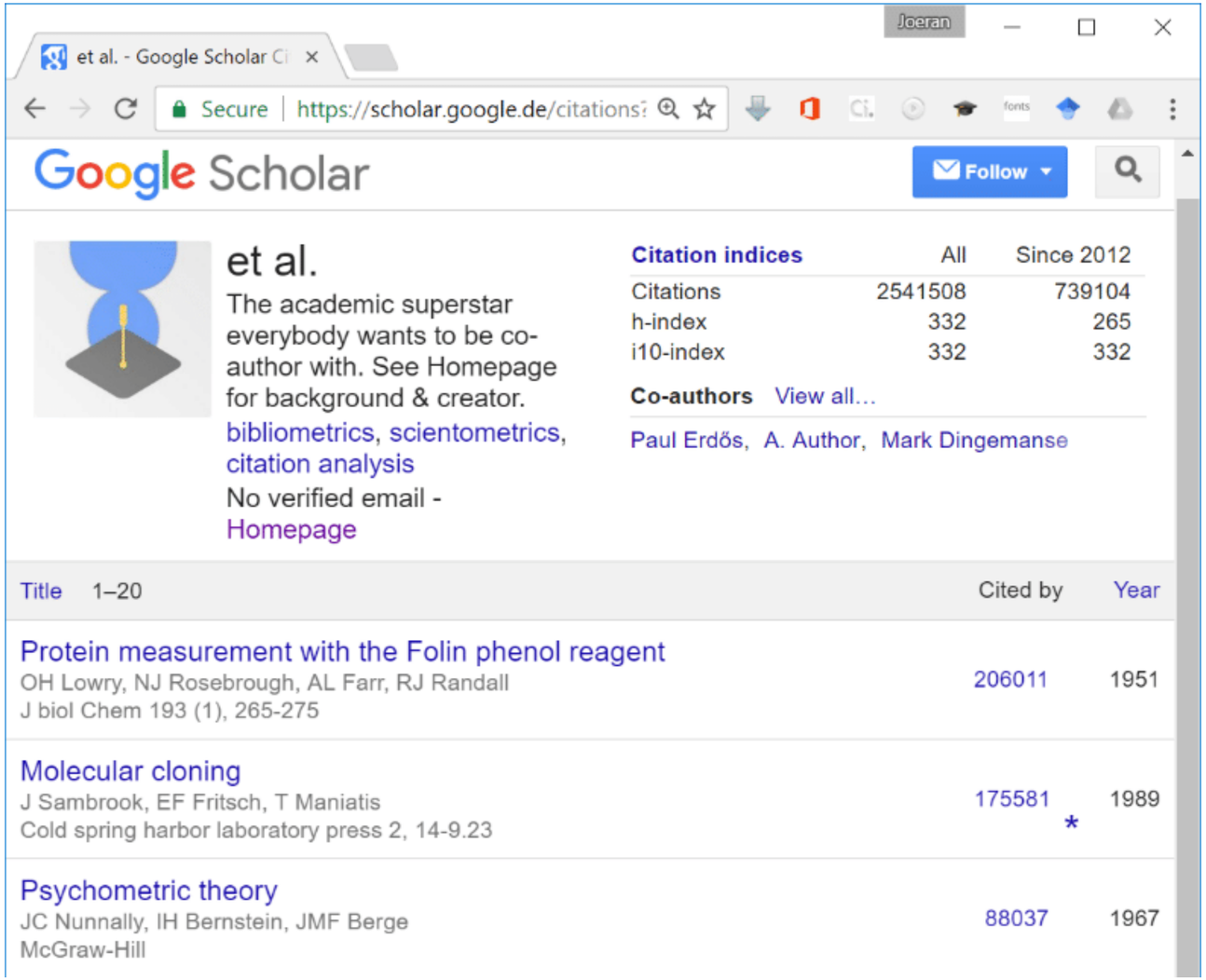
What about more “objective”, quantitative, metrics?

Performance metrics



Metrics—far from ideal

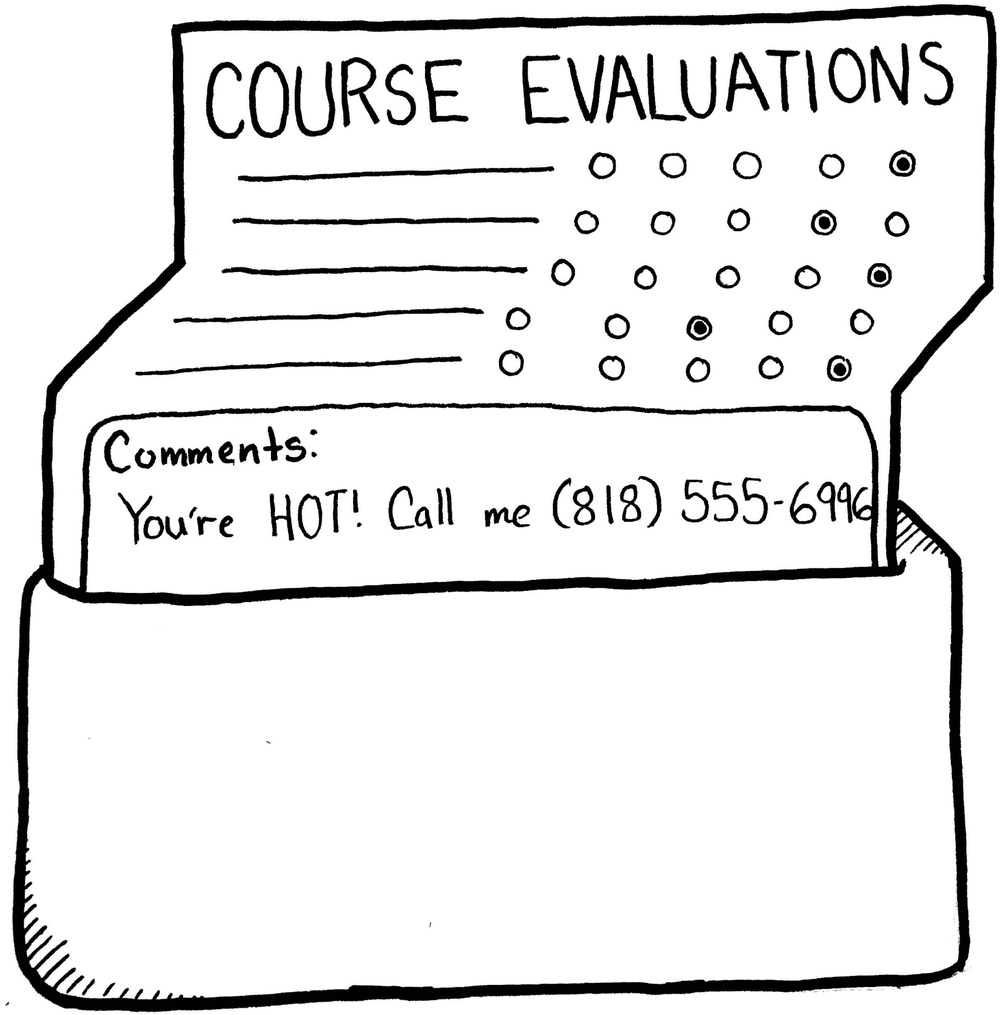
Criticisms of misincentives, validity issues, and more



An often overlooked faculty performance metric…

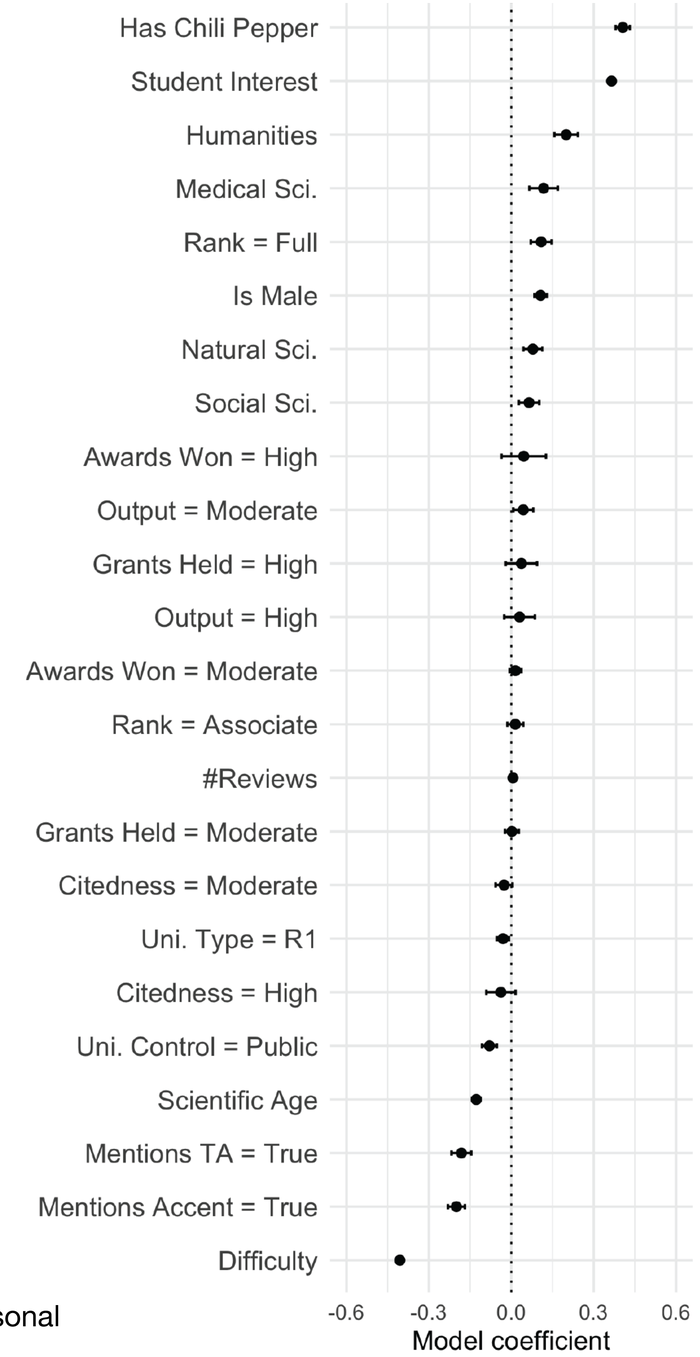
?

An often overlooked faculty performance metric…



Student ratings of teachers

Factors relating to ratings on 19,000 TT faculty on RateMyProfessors.con

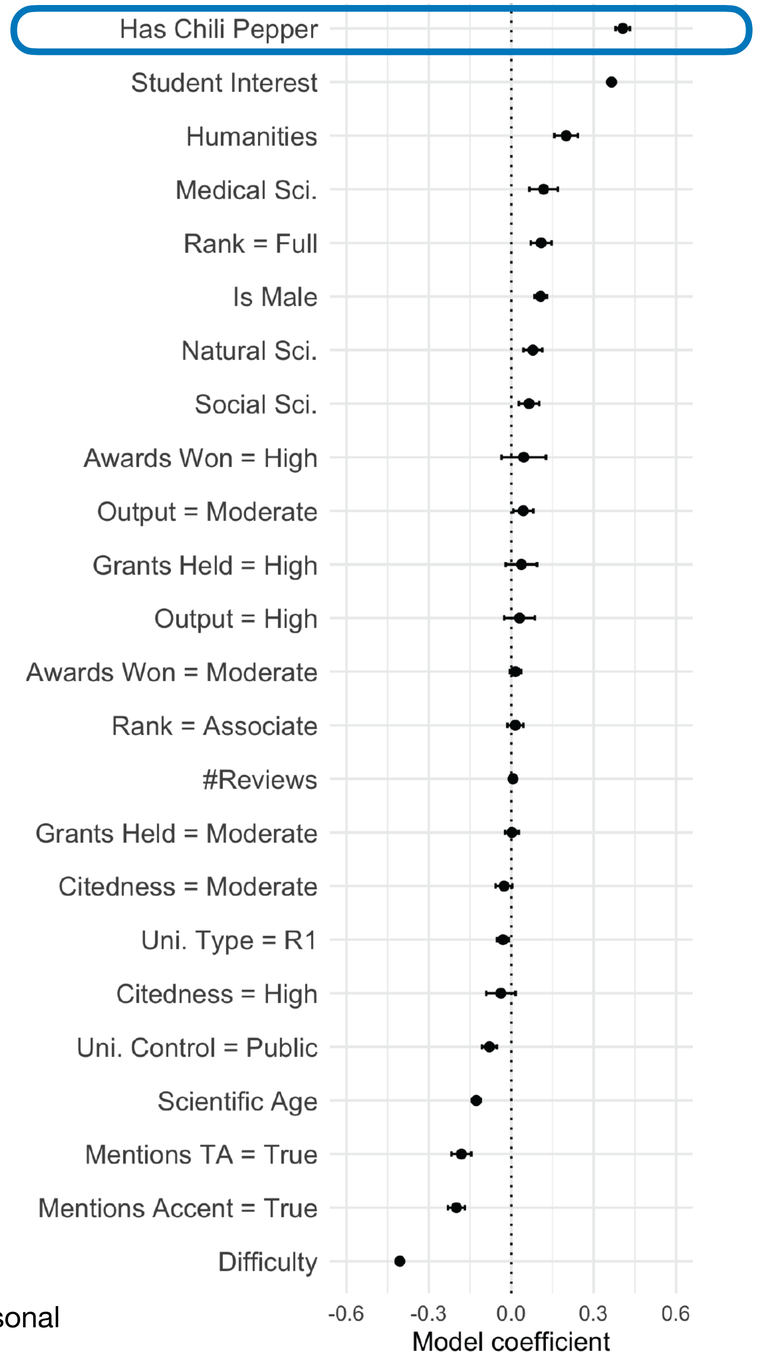


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

Student ratings of teachers

Factors relating to ratings on 19,000 TT faculty on RateMyProfessors.con

- Attractive profs rated higher

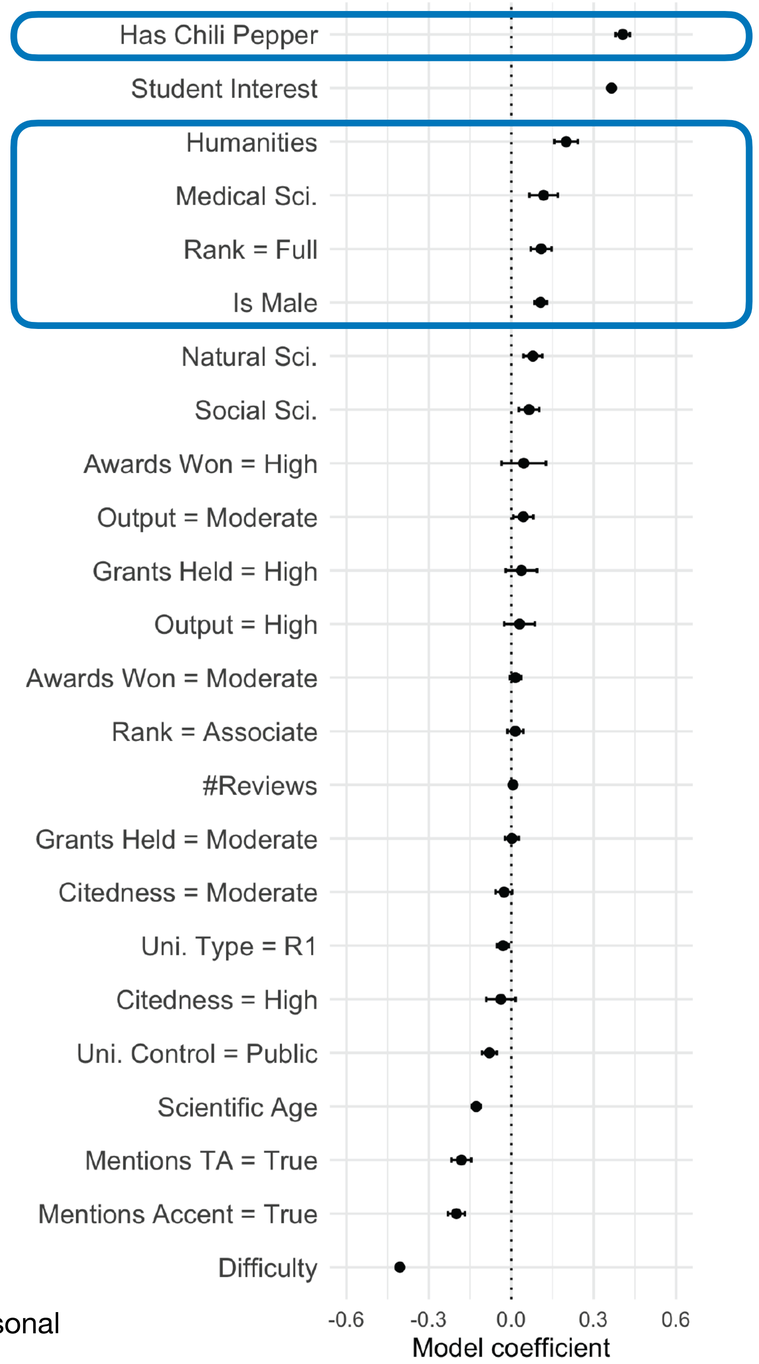


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

Student ratings of teachers

Factors relating to ratings on 19,000 TT faculty on RateMyProfessors.con

- Attractive profs rated higher
- Profs in humanities, men, and full professors have more positive reviews

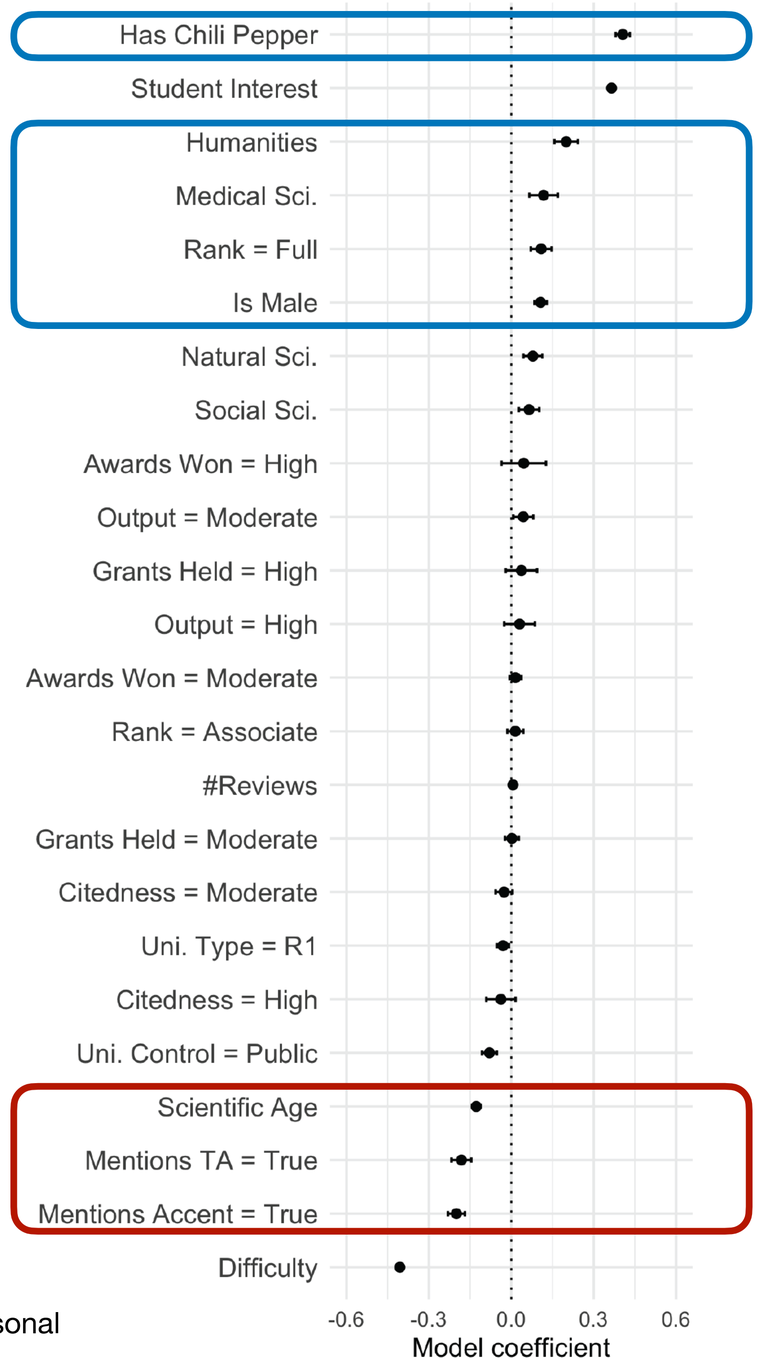


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

Student ratings of teachers

Factors relating to ratings on 19,000 TT faculty on RateMyProfessors.con

- Attractive profs rated higher
- Profs in humanities, men, and full professors have more positive reviews
- Older Profs, and those for whom an accent or TA was mentioned were rated lower

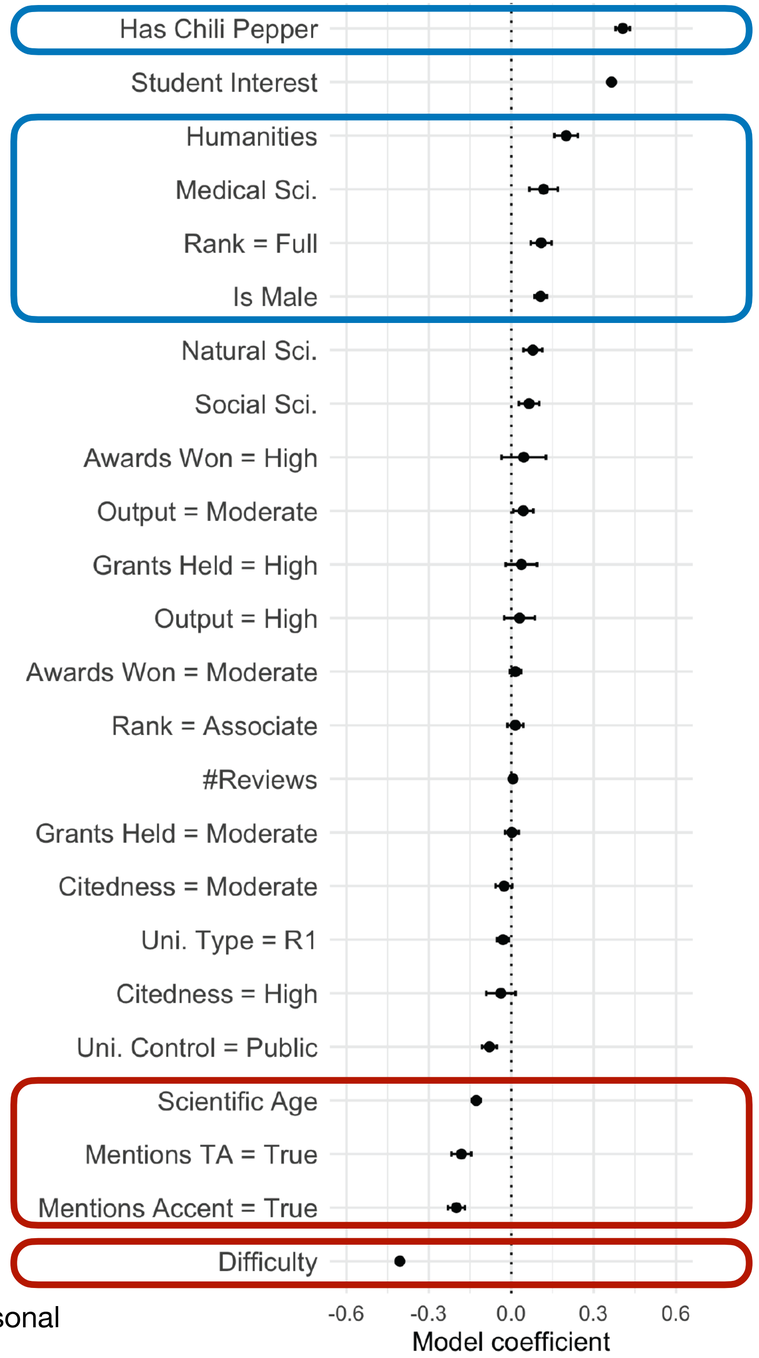


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

Student ratings of teachers

Factors relating to ratings on 19,000 TT faculty on RateMyProfessors.con

- Attractive profs rated higher
- Profs in humanities, men, and full professors have more positive reviews
- Older Profs, and those for whom an accent or TA was mentioned were rated lower
- The worst offense is teaching a difficult class

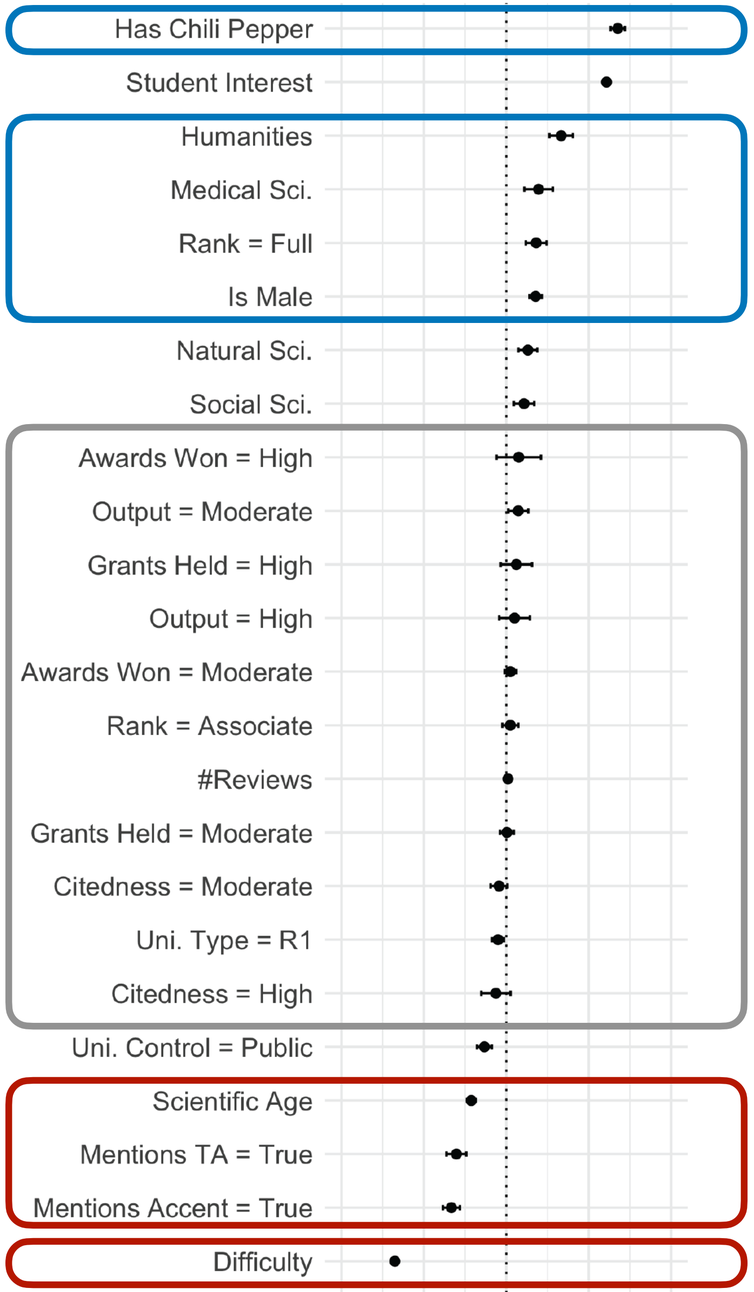


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

Student ratings of teachers

Factors relating to ratings on 19,000 TT faculty on RateMyProfessors.con

- Attractive profs rated higher
- Profs in humanities, men, and full professors have more positive reviews
- Older Profs, and those for whom an accent or TA was mentioned were rated lower
- The worst offense is teaching a difficult class
- No relation with research performance



Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

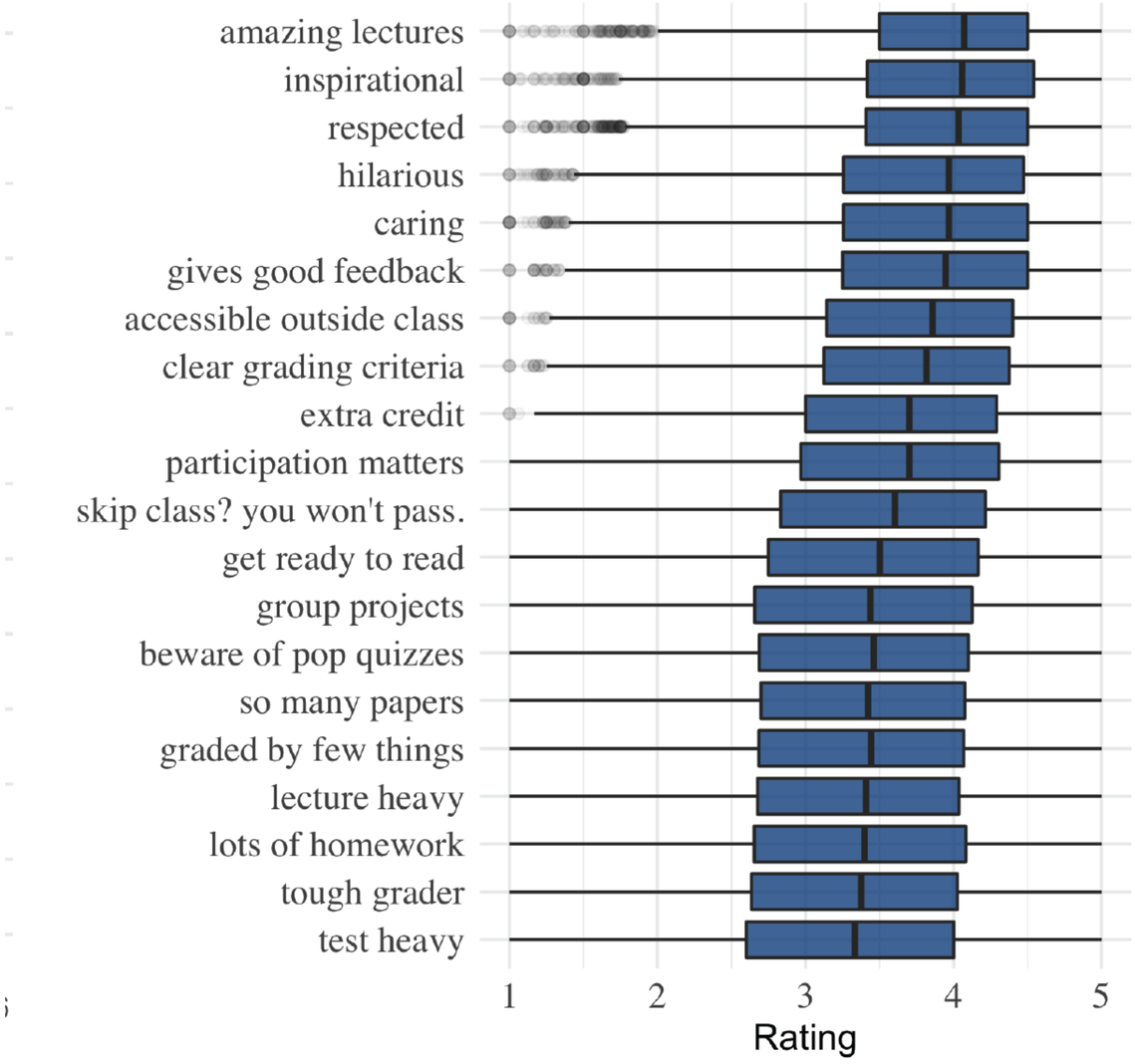
Not good indicators of teaching quality

Not good indicators of teaching quality

So what are they actually measuring?

Tags attached to reviews

Emotive tags (positive review) and workload related tags (low reviews)



Metrics often stem from human judgements,

Have all the same subjectivity

Are not always measuring what you hope they would

Citation metrics

A more objective approach?

Dakota Murray

FOLLOW

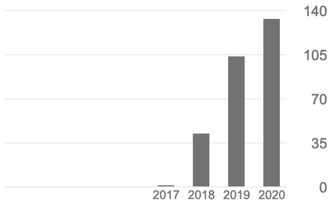
Graduate Student at Indiana University Bloomington
Verified email at iu.edu - Homepage

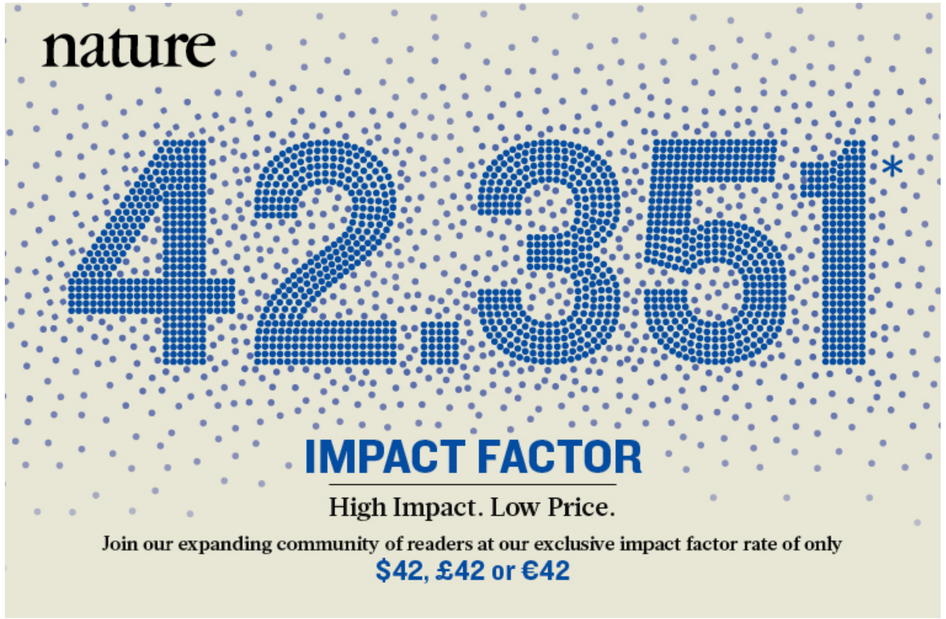
Scientometrics Scholarly Communication Data Science Science Policy

TITLE	CITED BY	YEAR
Scientists have most impact when they're free to move CR Sugimoto, N Robinson-García, DS Murray, A Yegros-Yegros, ... Nature News 550 (7674), 29	90	2017
Gender and international diversity improves equity in peer review D Murray, K Siler, V Larivière, WM Chan, AM Collings, J Raymond, ... BioRxiv, 400515	57	2019
The many faces of mobility: Using bibliometric data to measure the movement of scientists	30	2019

Cited by

	All	Since 2015
Citations	286	286
h-index	8	8
i10-index	7	7





Not all citations are endorsements!

Citations have many functions

Table 1. An Analysis of References in 30 Articles in *Physical Review*, Published on Theoretical High Energy Physics from 1968 to 1972, Inclusive.

	Total	'Big' papers	'Small' papers
Total number of references	706	333	373
Total number of papers referred to	575	292	283
Extraneous references (books, footnotes, experimental papers, private communications, etc.)	292	147	145
1. Conceptual	306 (53%)	158 (54%)	148 (52%)
Operational	245 (43%)	120 (41%)	125 (44%)
Neither	41 (7%)	21 (7%)	20 (7%)
2. Organic	345 (60%)	167 (57%)	178 (63%)
Perfunctory	238 (41%)	125 (43%)	113 (40%)
Neither	5 (1%)	3 (1%)	2 (1%)
3. Evolutionary	338 (59%)	168 (57%)	170 (60%)
Juxtapositional	229 (40%)	120 (41%)	109 (39%)
Neither	13 (2%)	11 (4%)	2 (1%)
4. Confirmative	502 (87%)	264 (90%)	238 (84%)
Negational	83 (14%)	39 (13%)	44 (16%)
Neither	26 (5%)	8 (3%)	22 (8%)
Redundant	177 (31%)	97 (33%)	80 (28%)

Moravcsik, M. J., & Murugesan, P. (1975). Some Results on the Function and Quality of Citations. *Social Studies of Science*, *5*(1), 86–92.

Norms of citation differ between disciplines

Individual, low-consensus, sometimes antagonistic philosophy

Foucault

Chomsky

Tannen, D. (2002). Agonism in academic discourse. *Journal of Pragmatics*, *34*(10), 1651–1669.

Collaborative, high-consensus, high-author-count High-Energy Physics

Knorr-Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press.

Disagreement citations

Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Disagreement citations

- Develop a cue-word based approach to identify 450,000 instances of disagreement across 3 million articles

Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

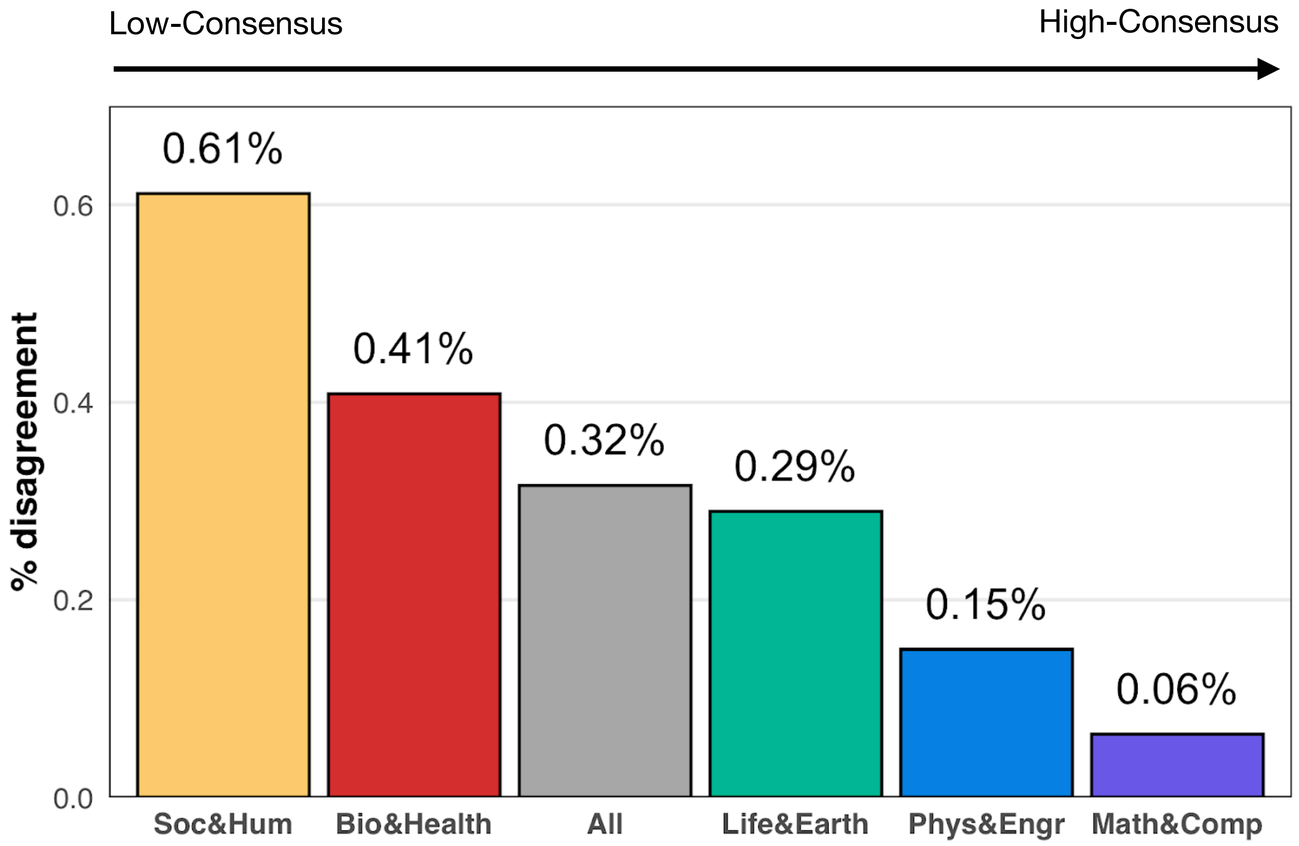
Disagreement citations

- Develop a cue-word based approach to identify 450,000 instances of disagreement across 3 million articles
- Do fields disagree…differently?

Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Disagreement citations

- Develop a cue-word based approach to identify 450,000 instances of disagreement across 3 million articles
- Do fields disagree…differently?

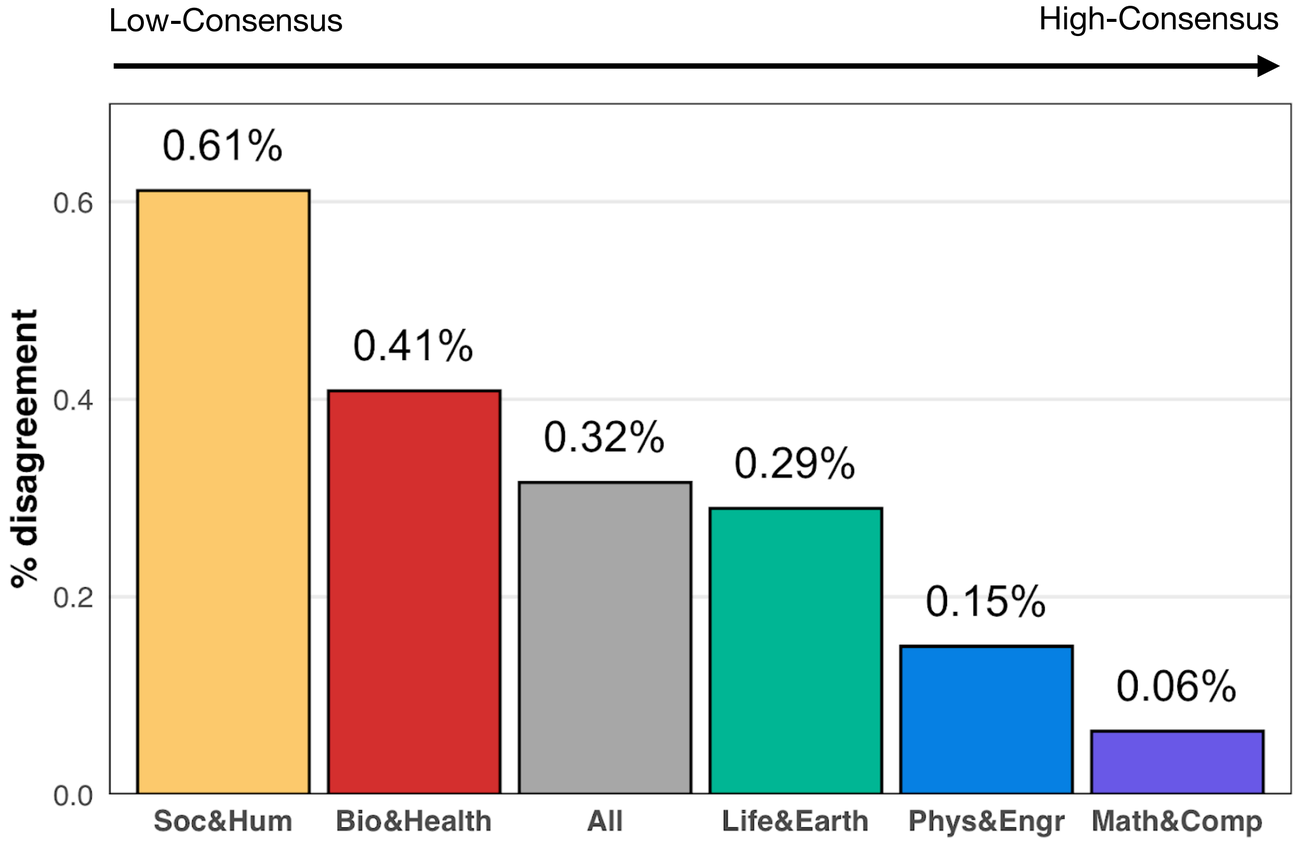


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Disagreement citations

- Develop a cue-word based approach to identify 450,000 instances of disagreement across 3 million articles
- Do fields disagree…differently?

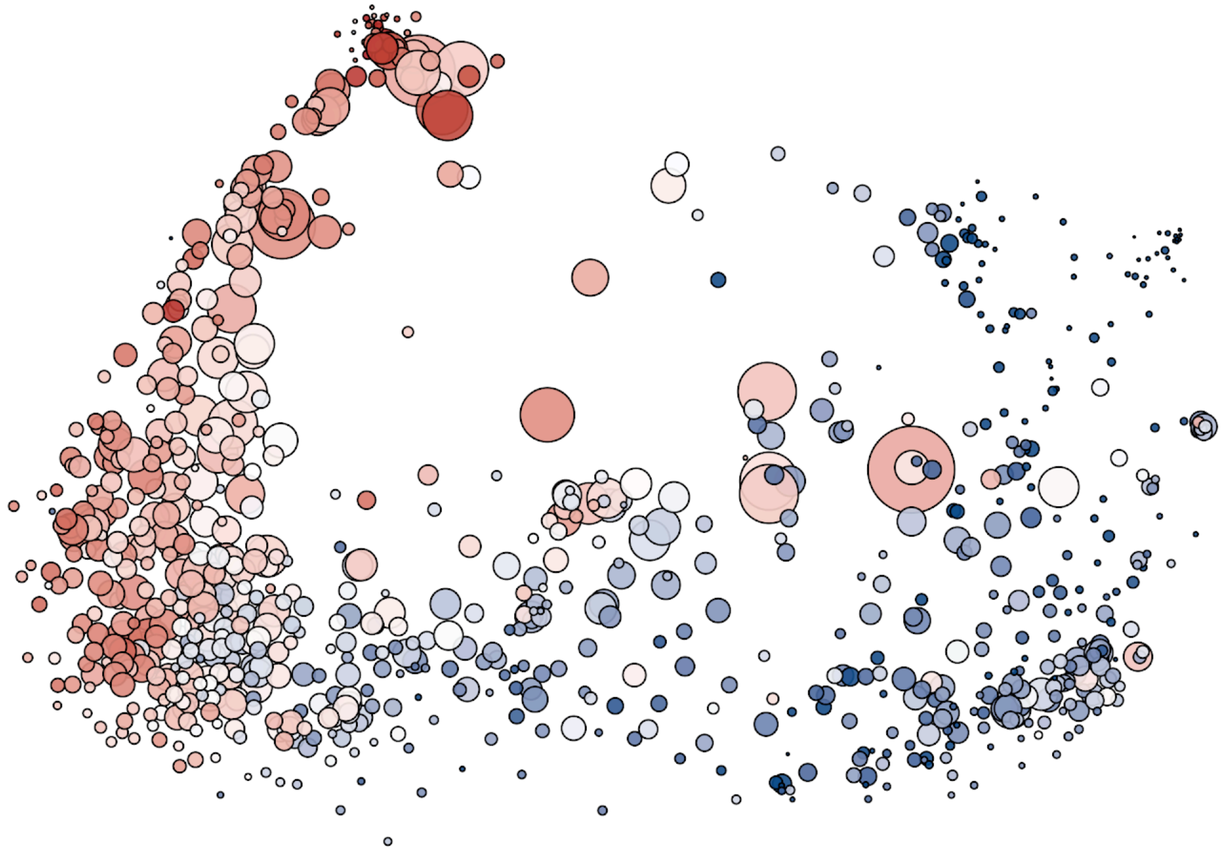
Too coarse?



Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Digging deeper

886 meso-level fields

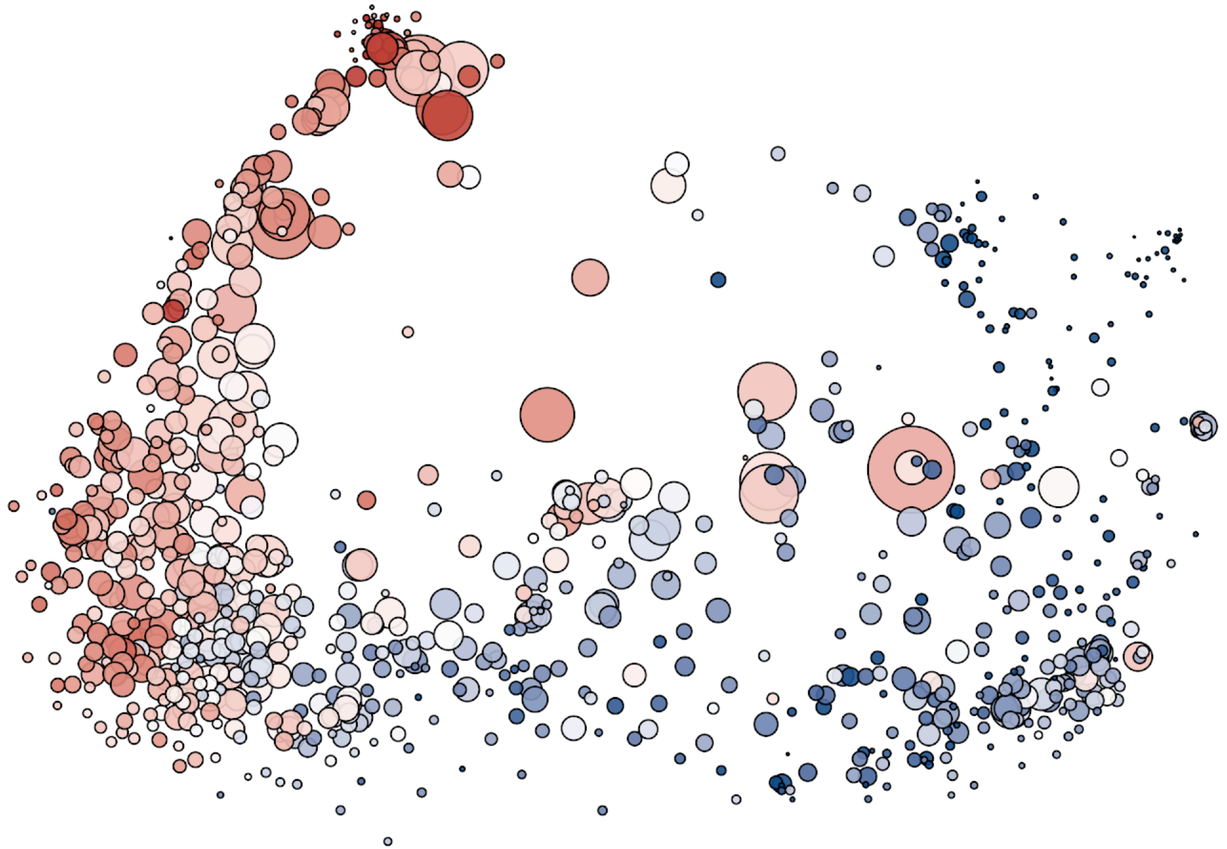


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Digging deeper

886 meso-level fields

- Each dot is a cluster of papers
- Area maps to size of field
- Distance reflects relatedness

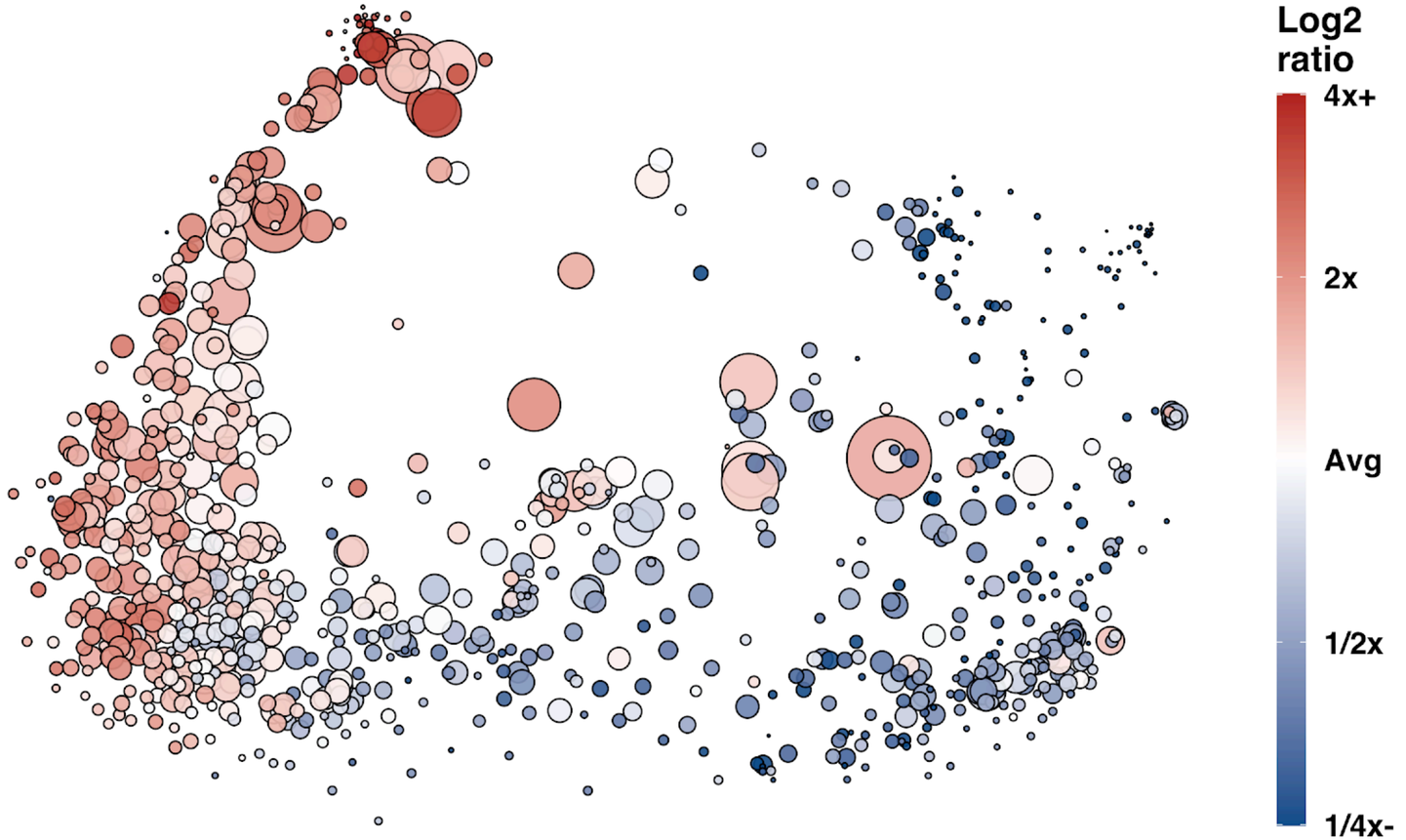


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Digging deeper

886 meso-level fields

- Each dot is a cluster of papers
- Area maps to size of field
- Distance reflects relatedness
- Color reflects ratio of disagreement to overall average

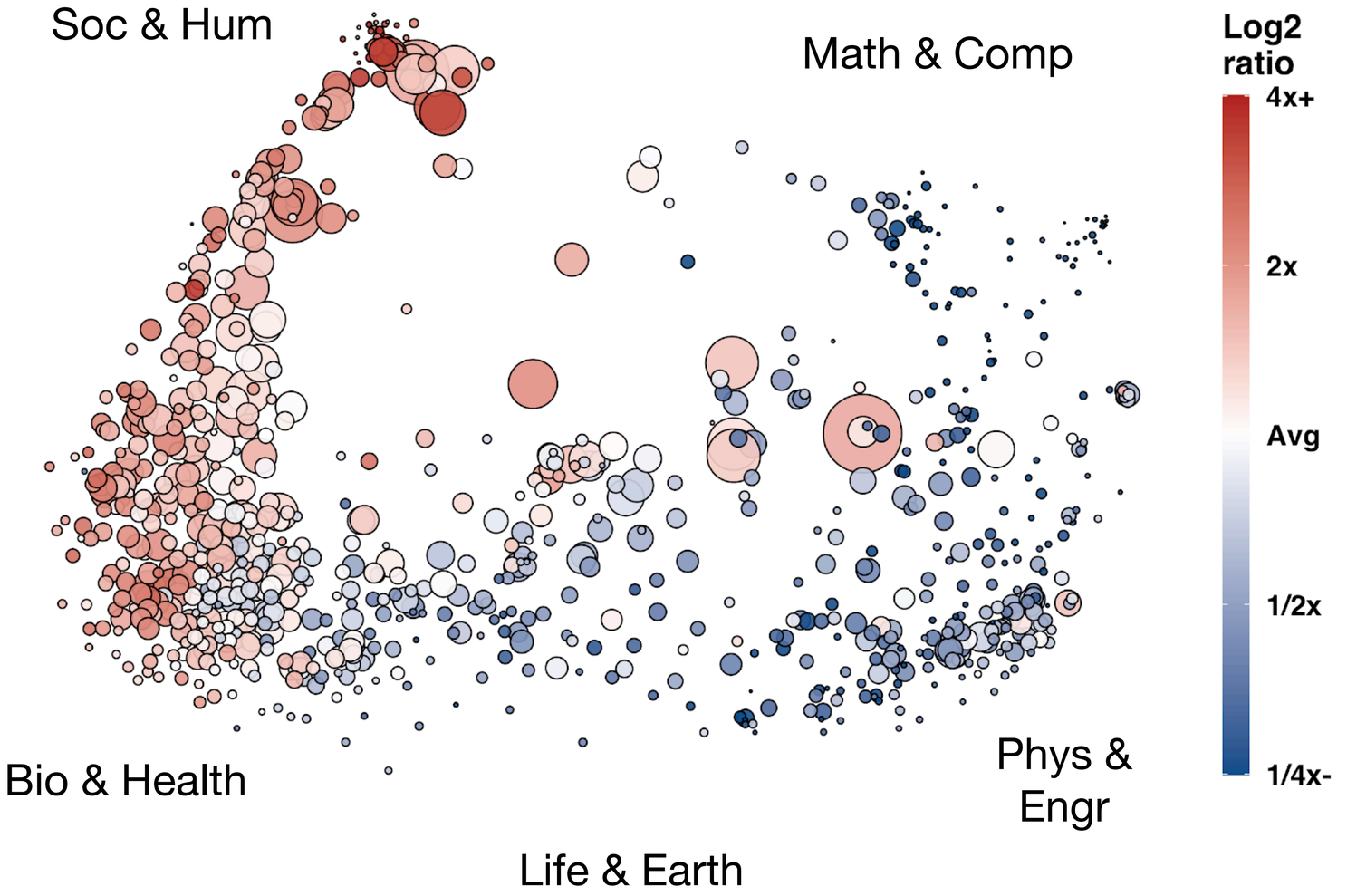


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

Digging deeper

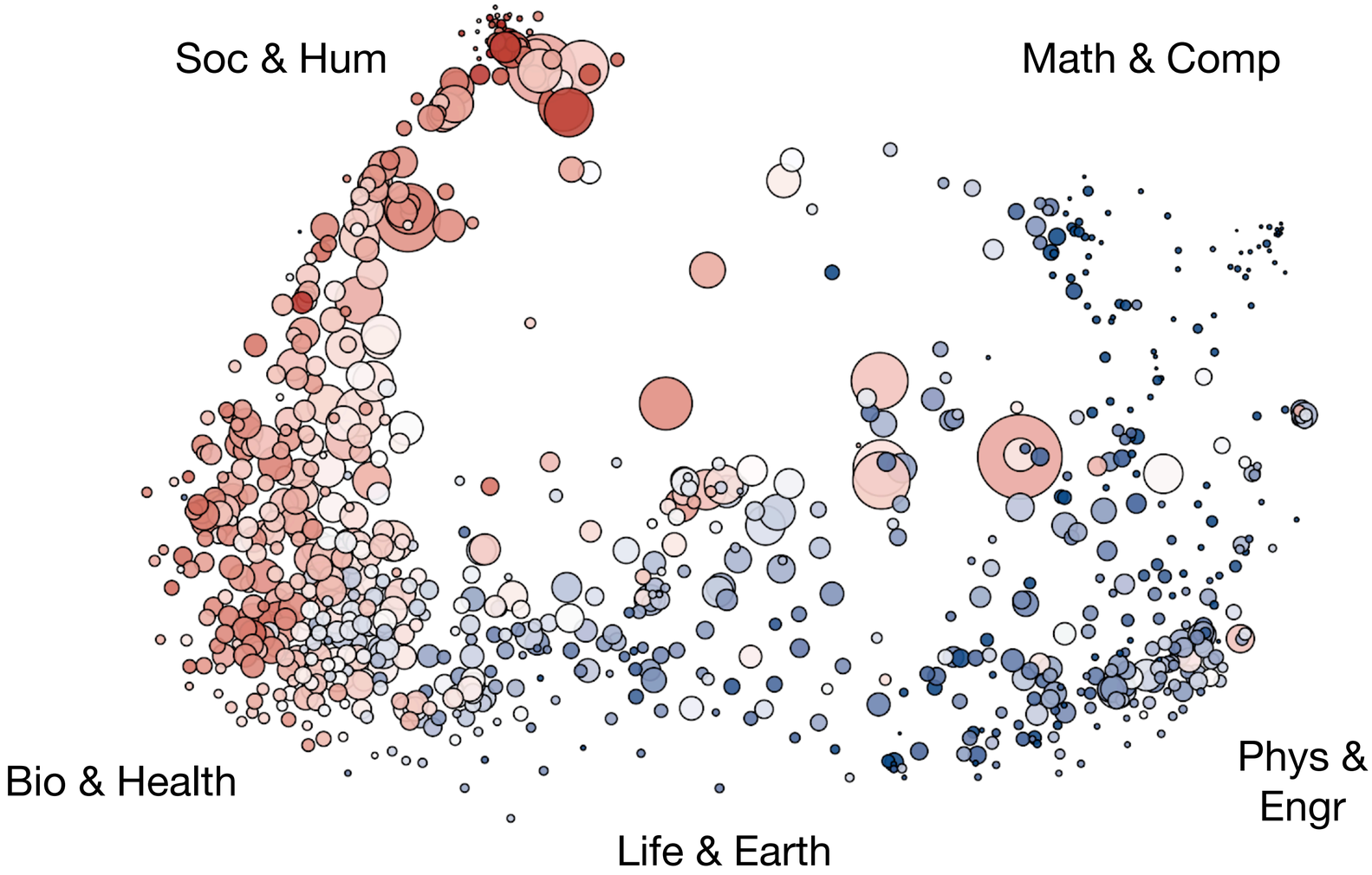
886 meso-level fields

- Each dot is a cluster of papers
- Area maps to size of field
- Distance reflects relatedness
- Color reflects ratio of disagreement to overall average

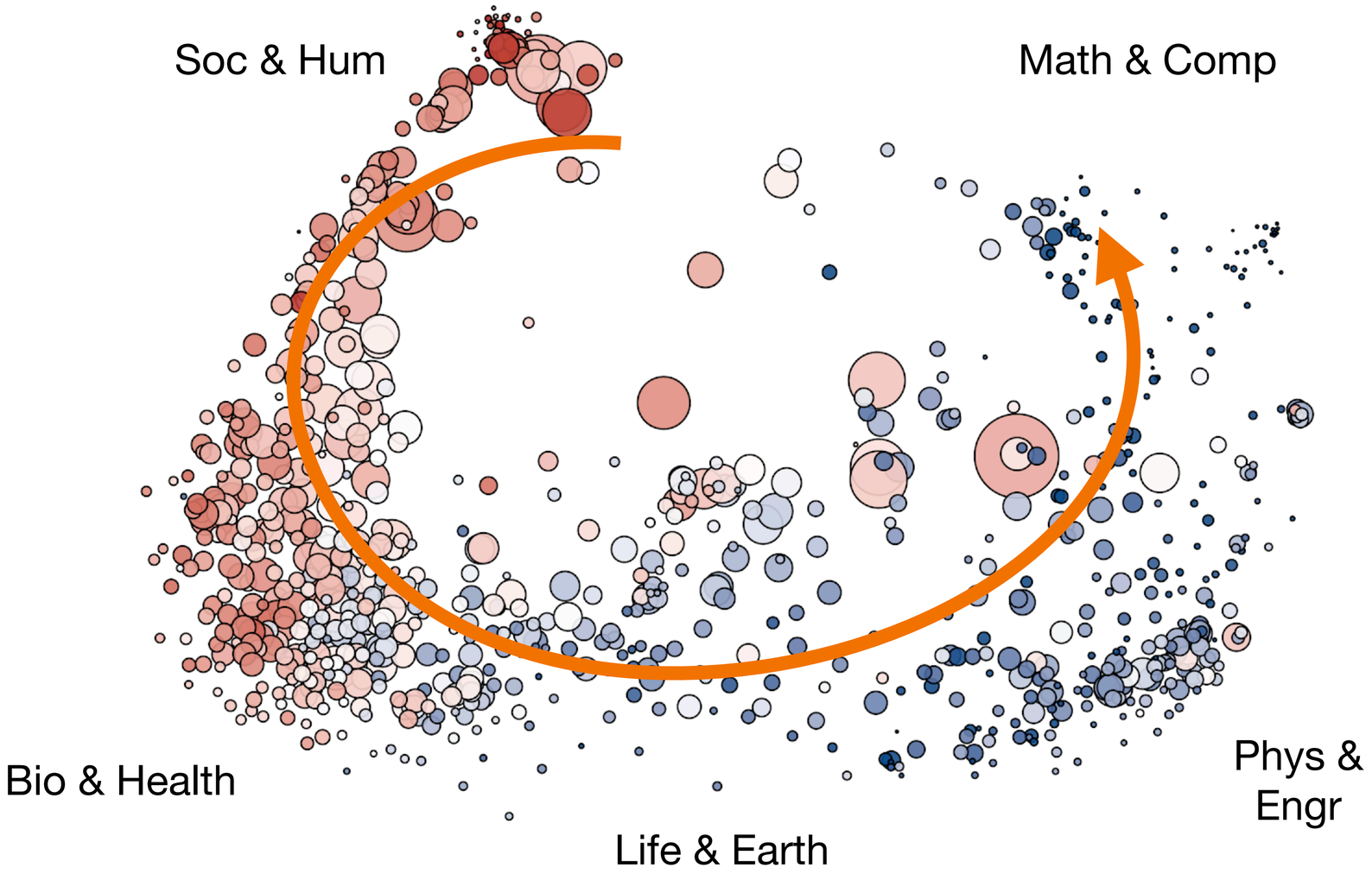


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation)*.

The pattern repeats

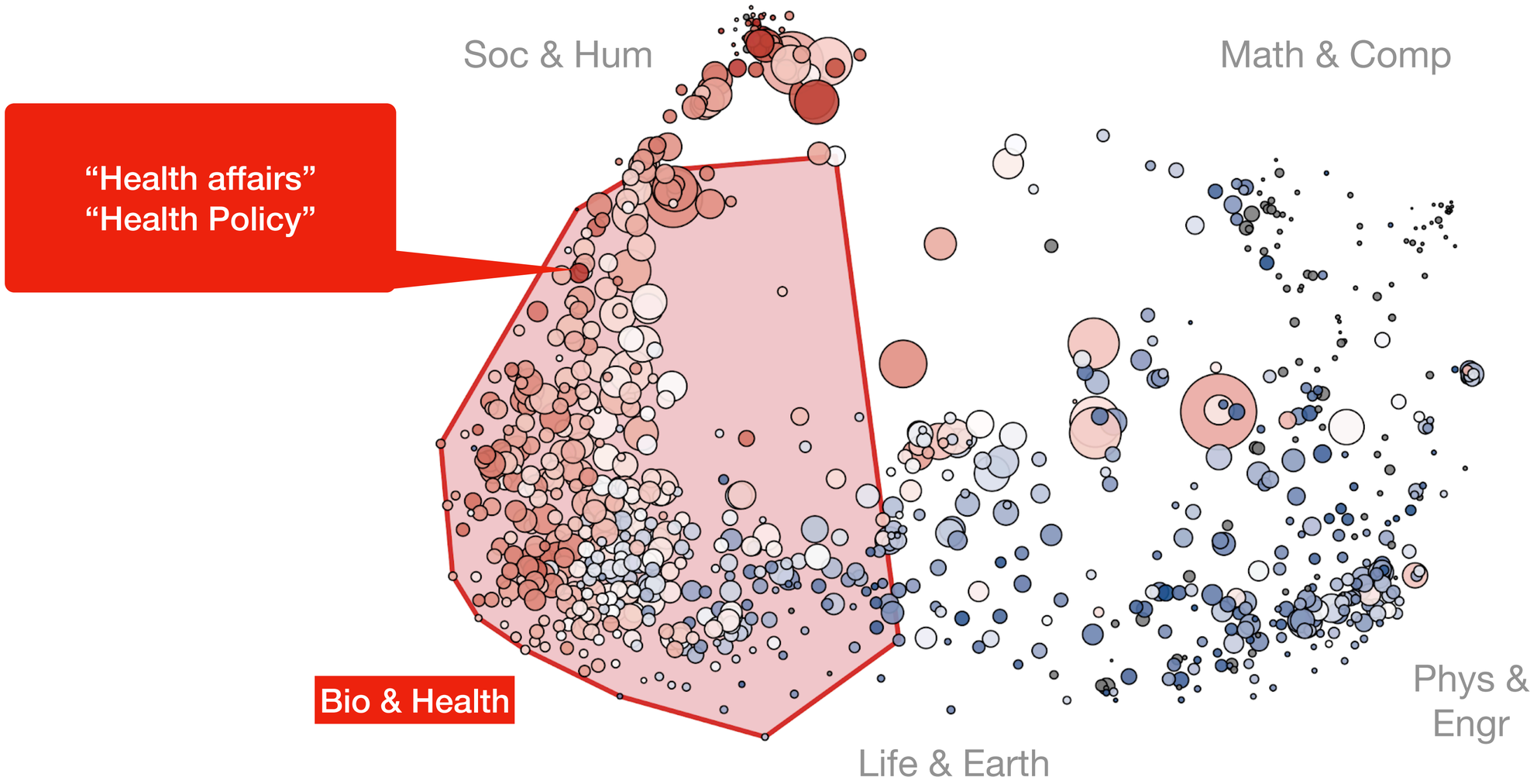


The pattern repeats



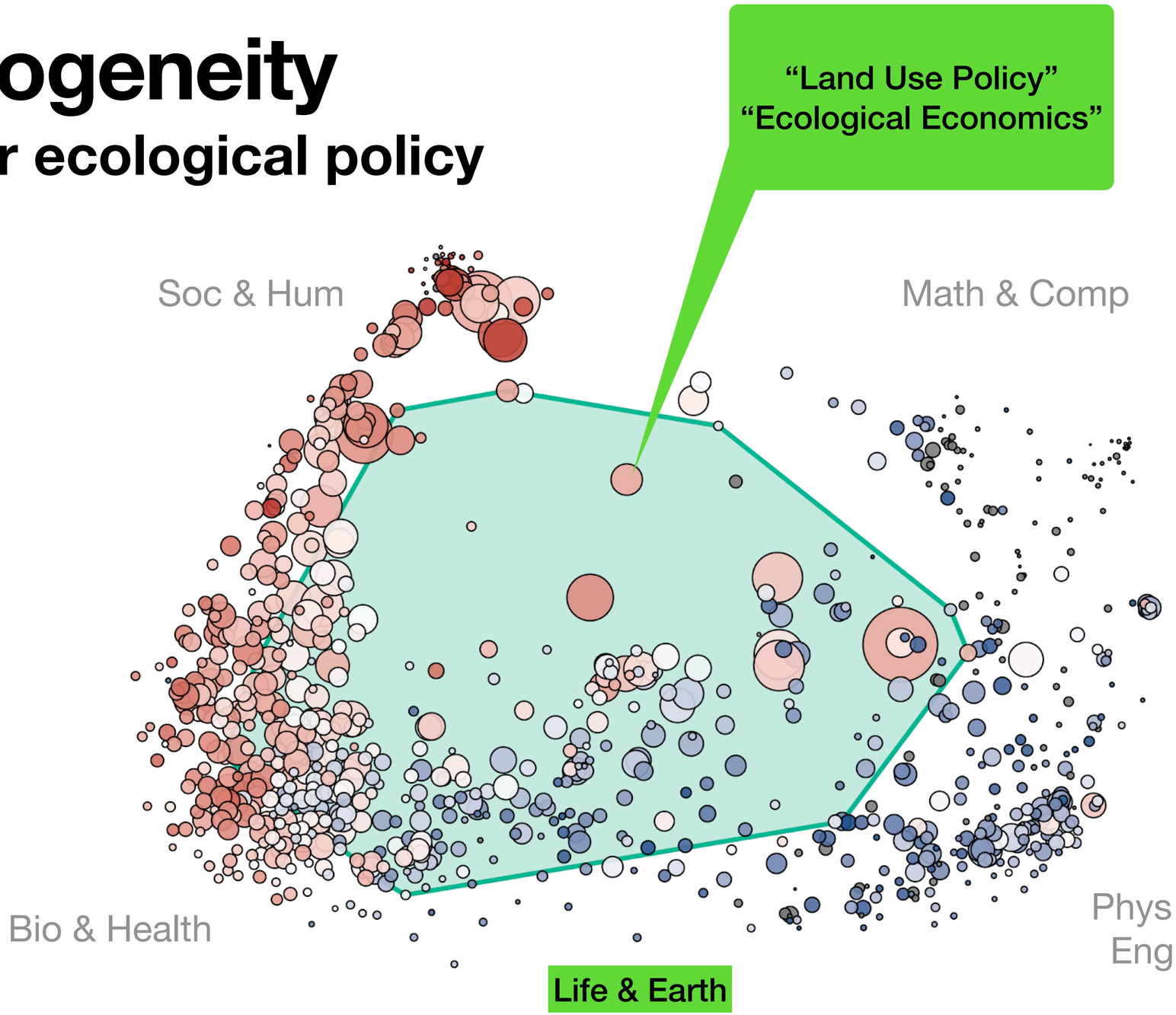
Heterogeneity

Health policy has high-disagreement relative to surrounding fields



Heterogeneity

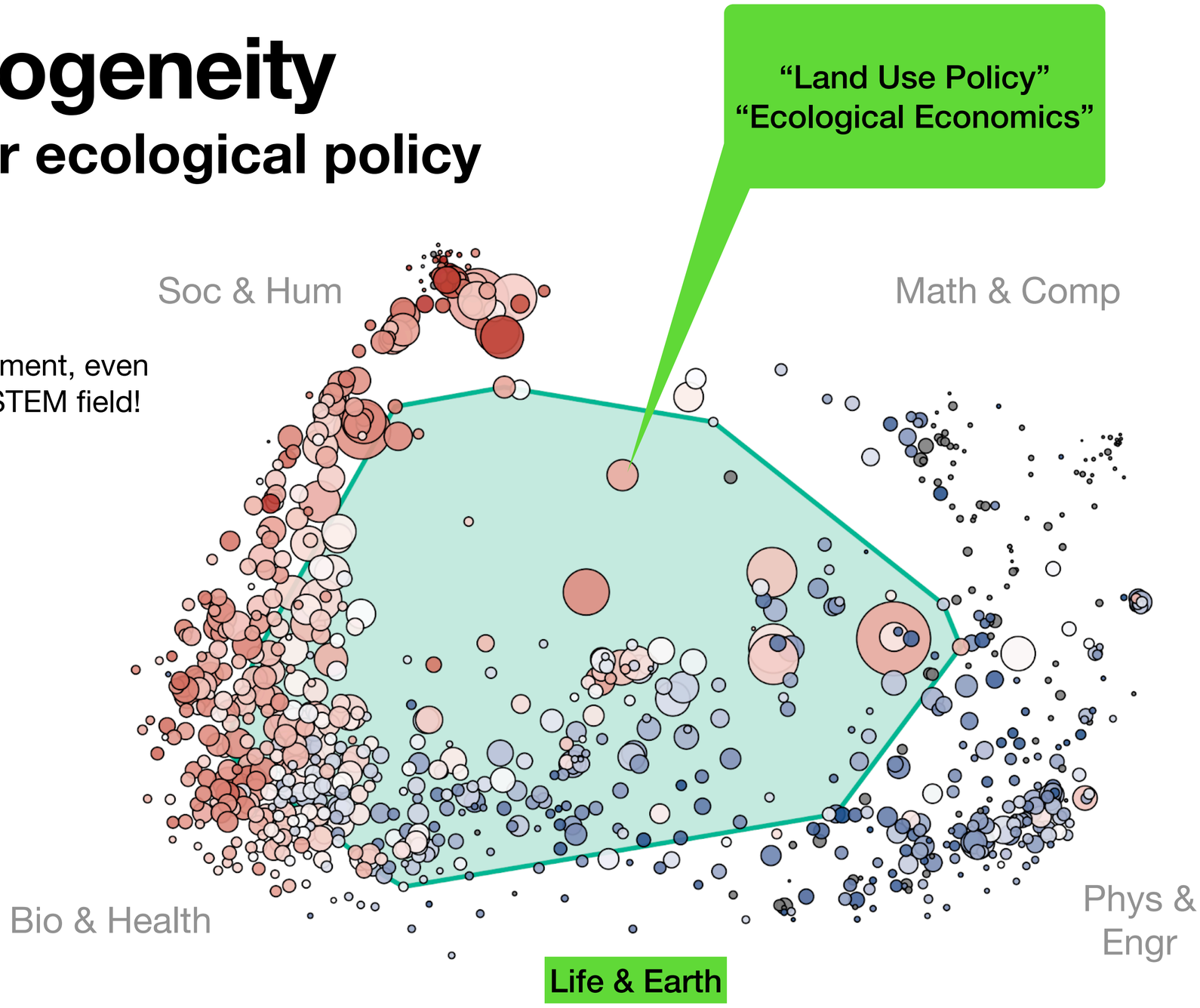
Same for ecological policy



Heterogeneity

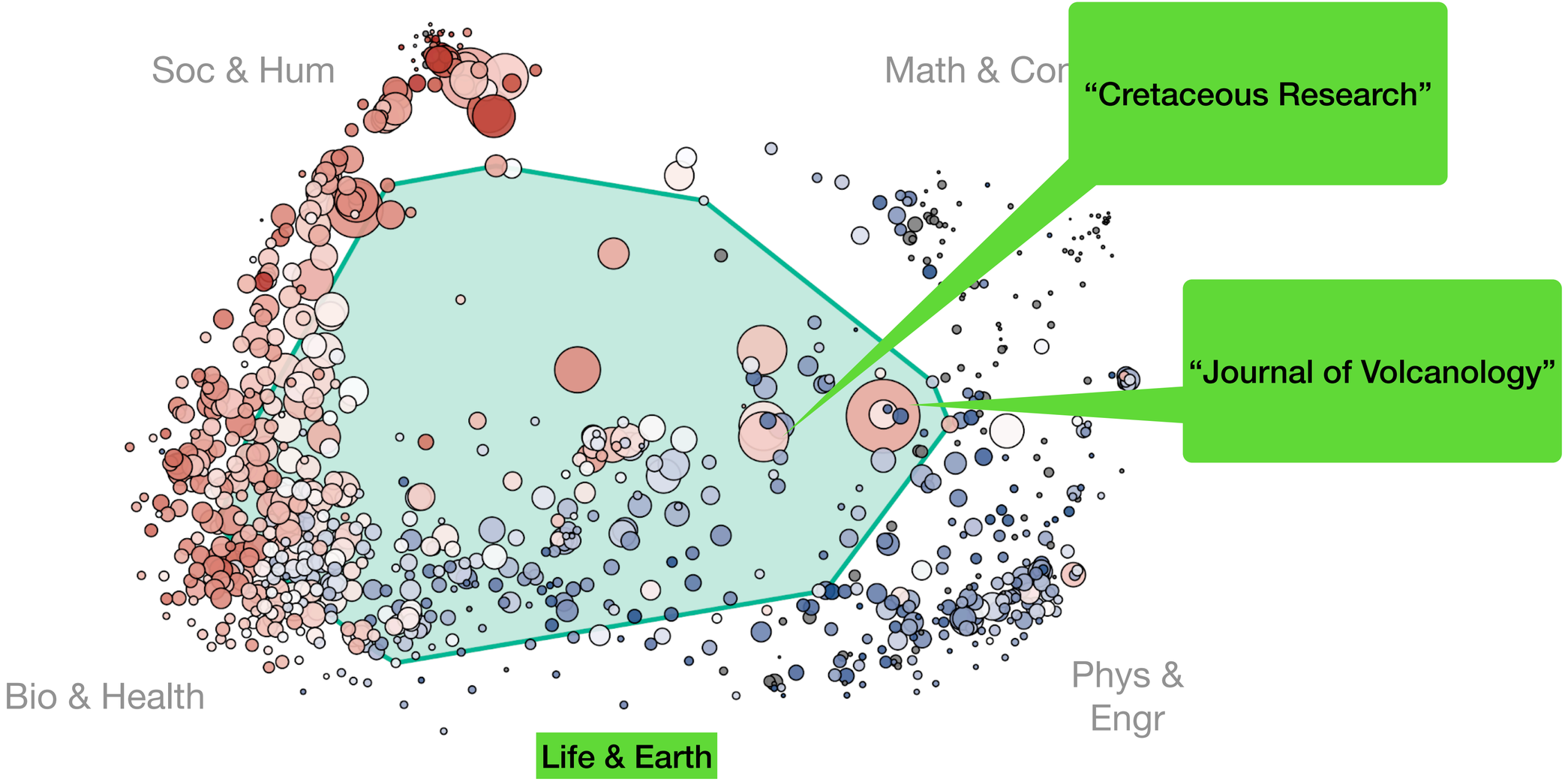
Same for ecological policy

Policy is high-disagreement, even when nested under a STEM field!



Heterogeneity

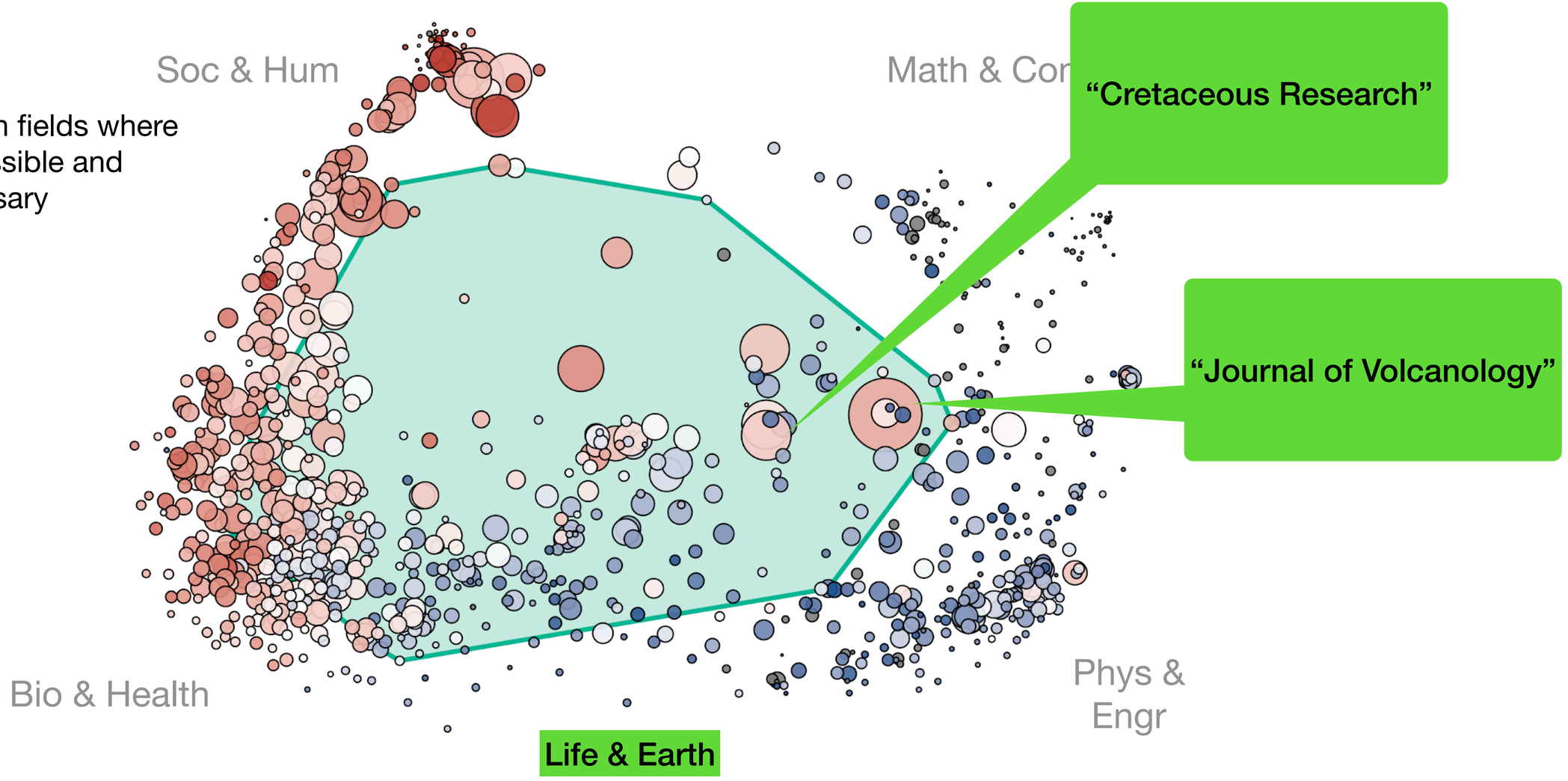
Fields depending on historical records



Heterogeneity

Fields depending on historical records

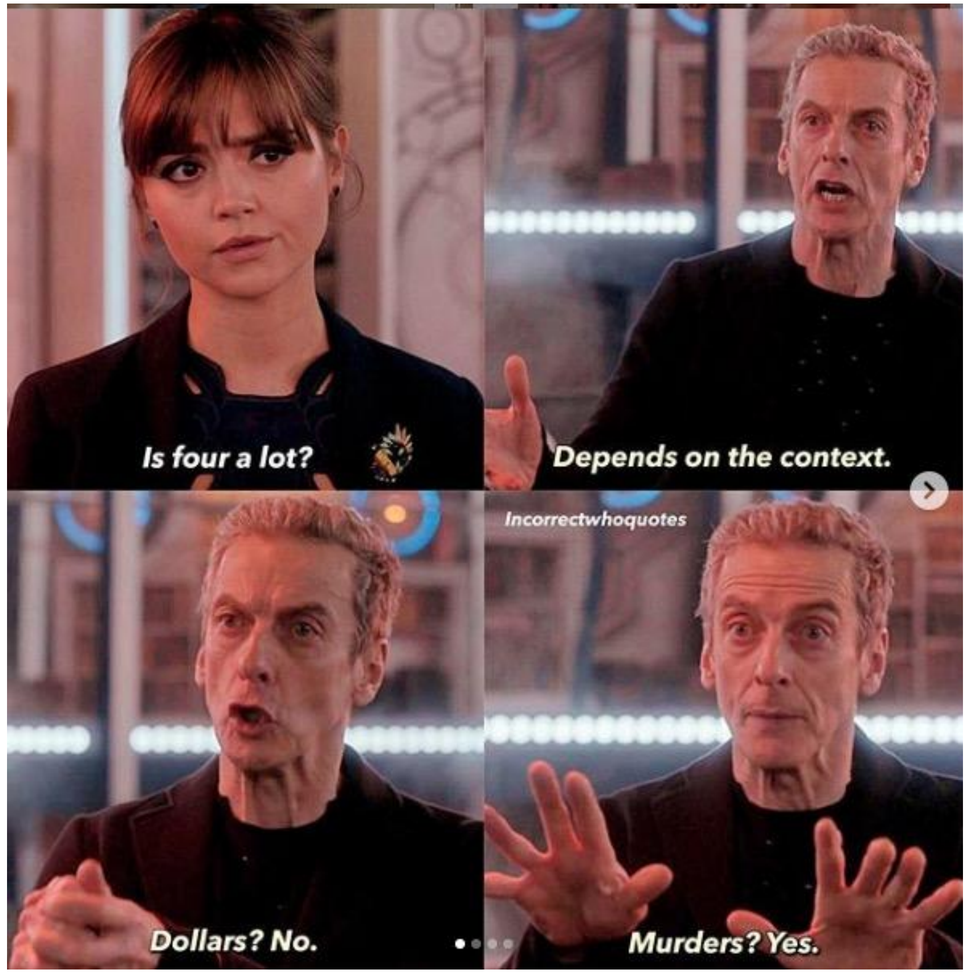
Disagreement is high in fields where experiments are impossible and interpretation is necessary



Disciplines have different norms, expectations, and cultures that shape how and why they cite

In metrics…a number is just a number

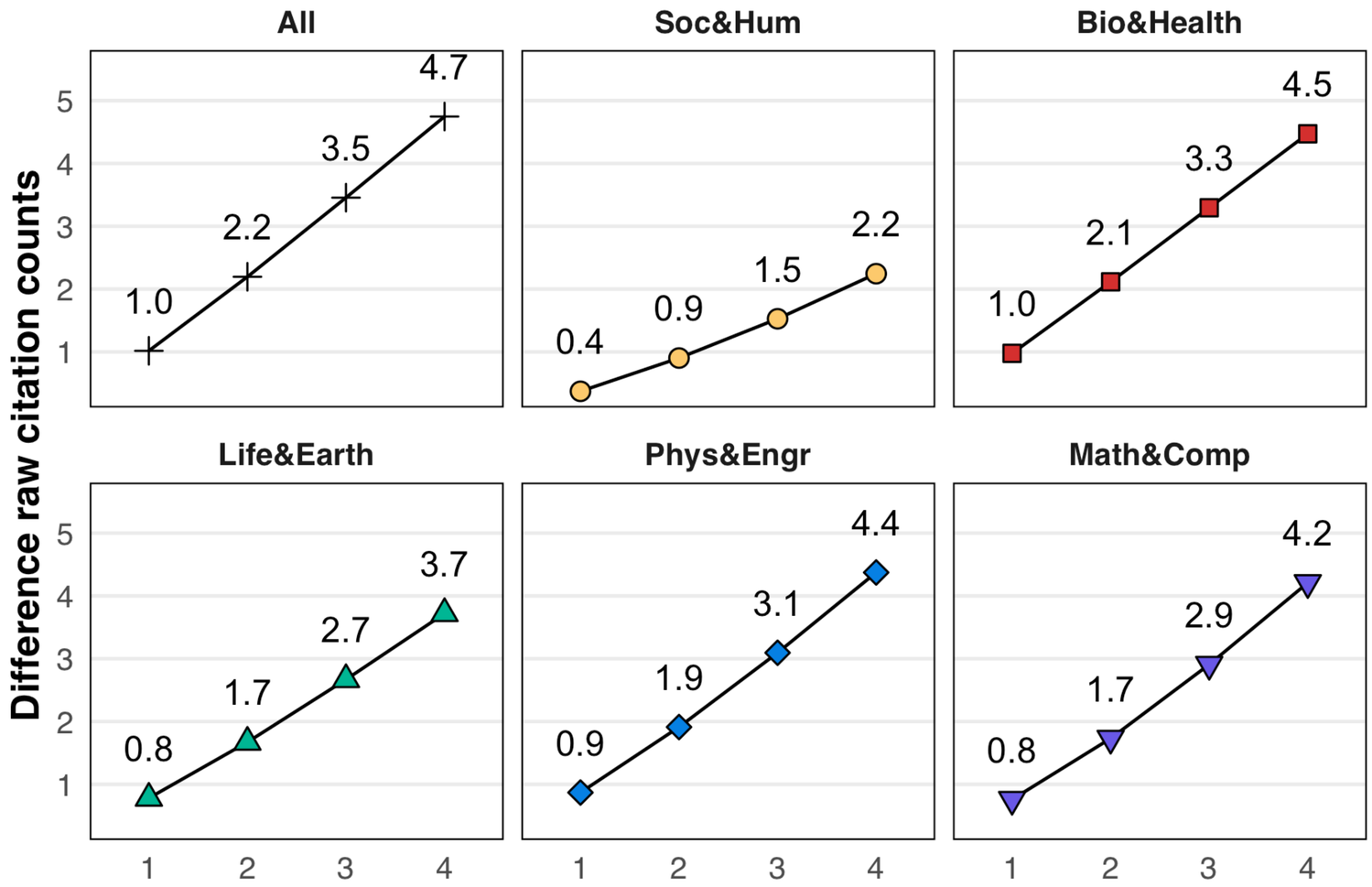
Context is needed to interpret it



An aside:

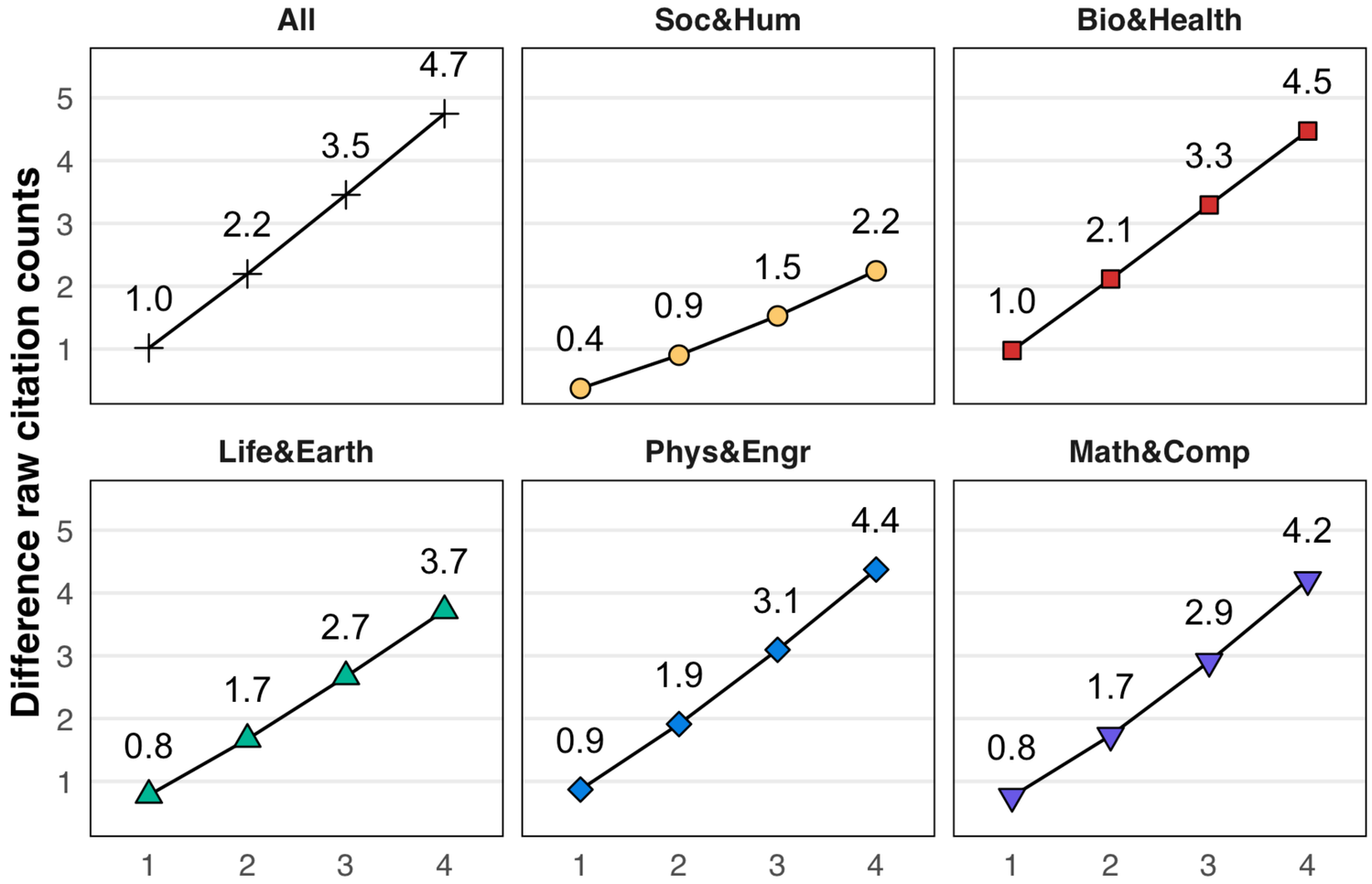
Preliminary results on the relationship between disagreement and citation

Compared papers with a disagreement citation to those without



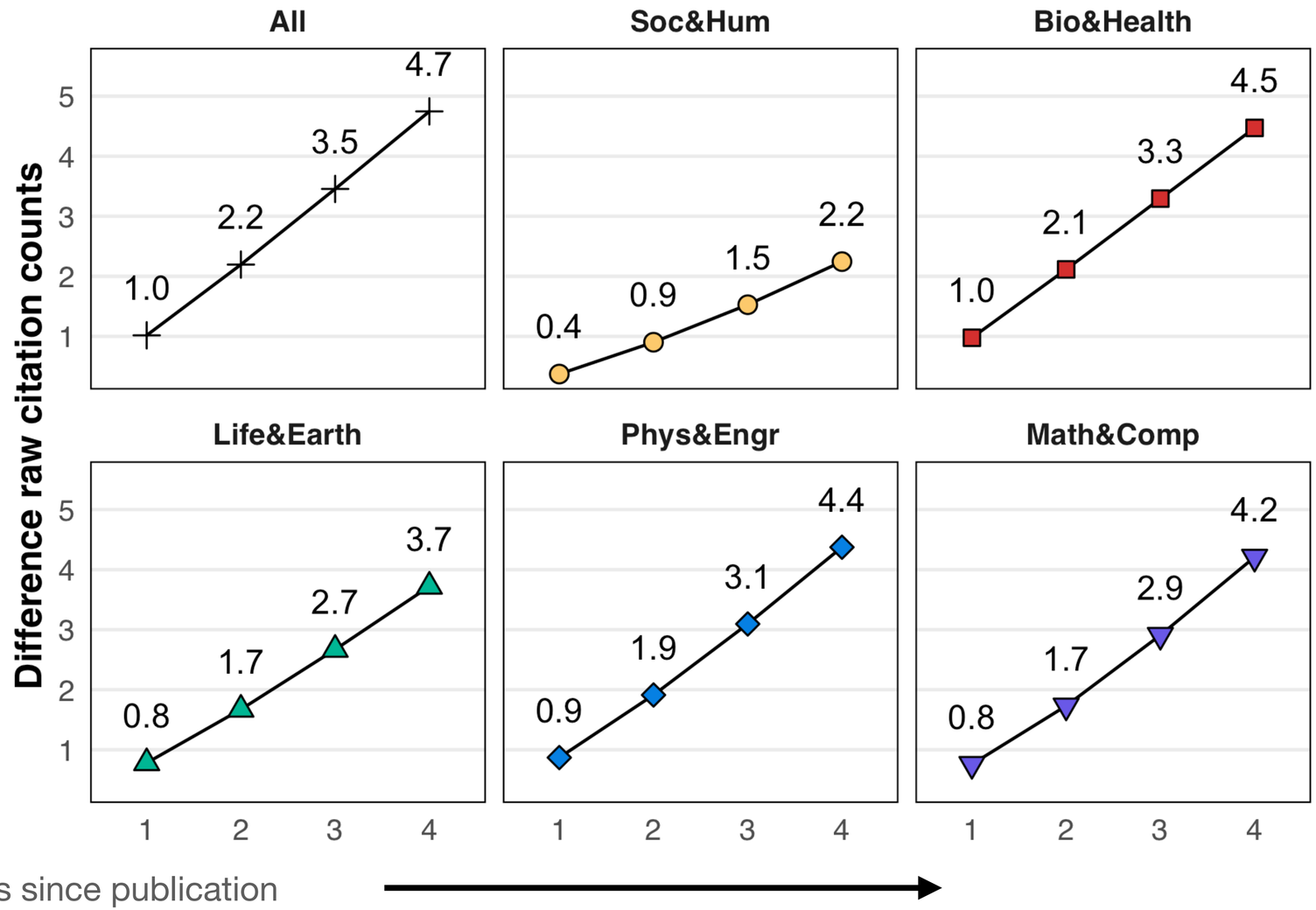
Compared papers with a disagreement citation to those without

Higher values indicate that papers with disagreement sentences received more citations



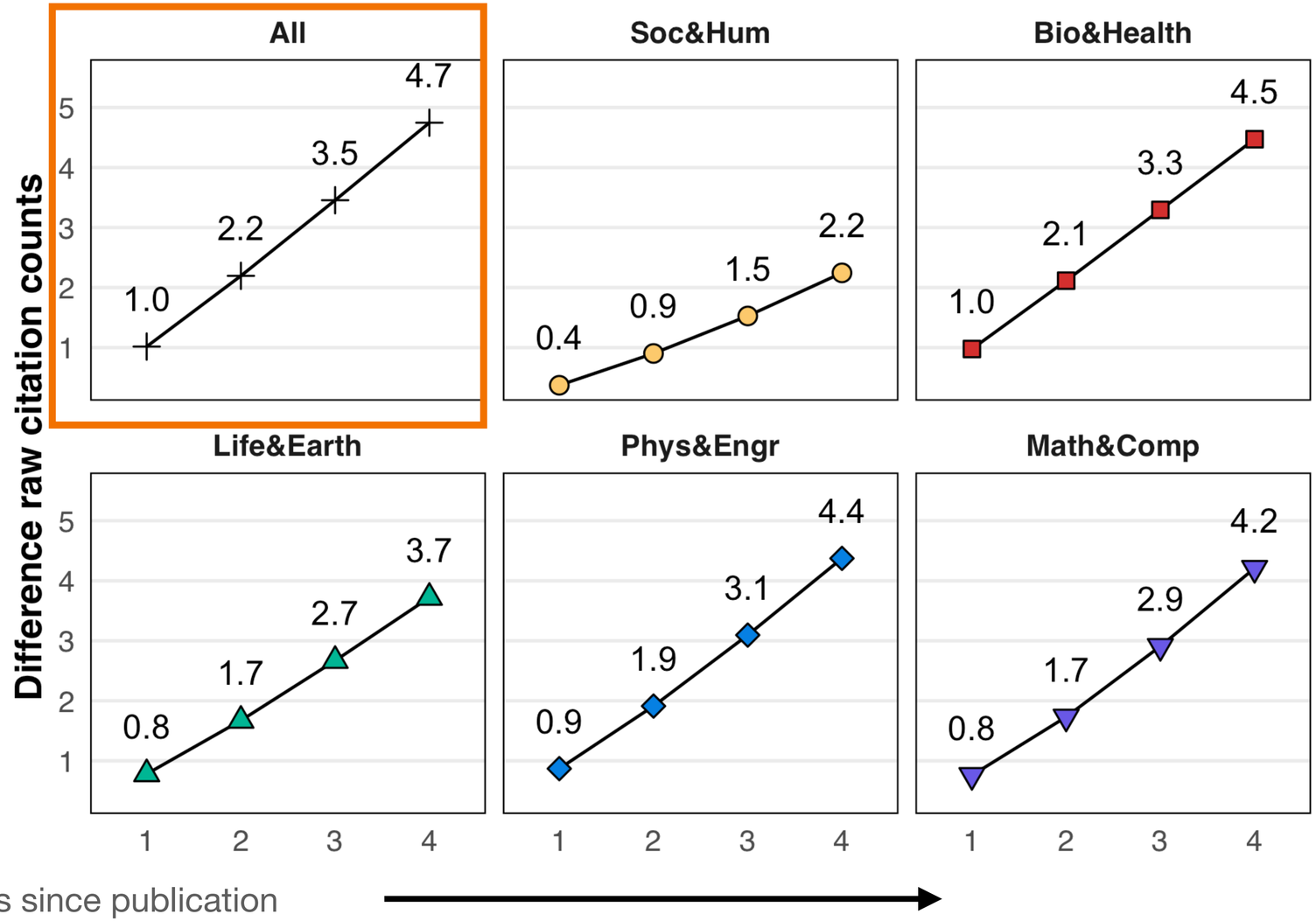
Compared papers with a disagreement citation to those without

Higher values indicate that papers with disagreement sentences received more citations



Papers that disagree have more citations

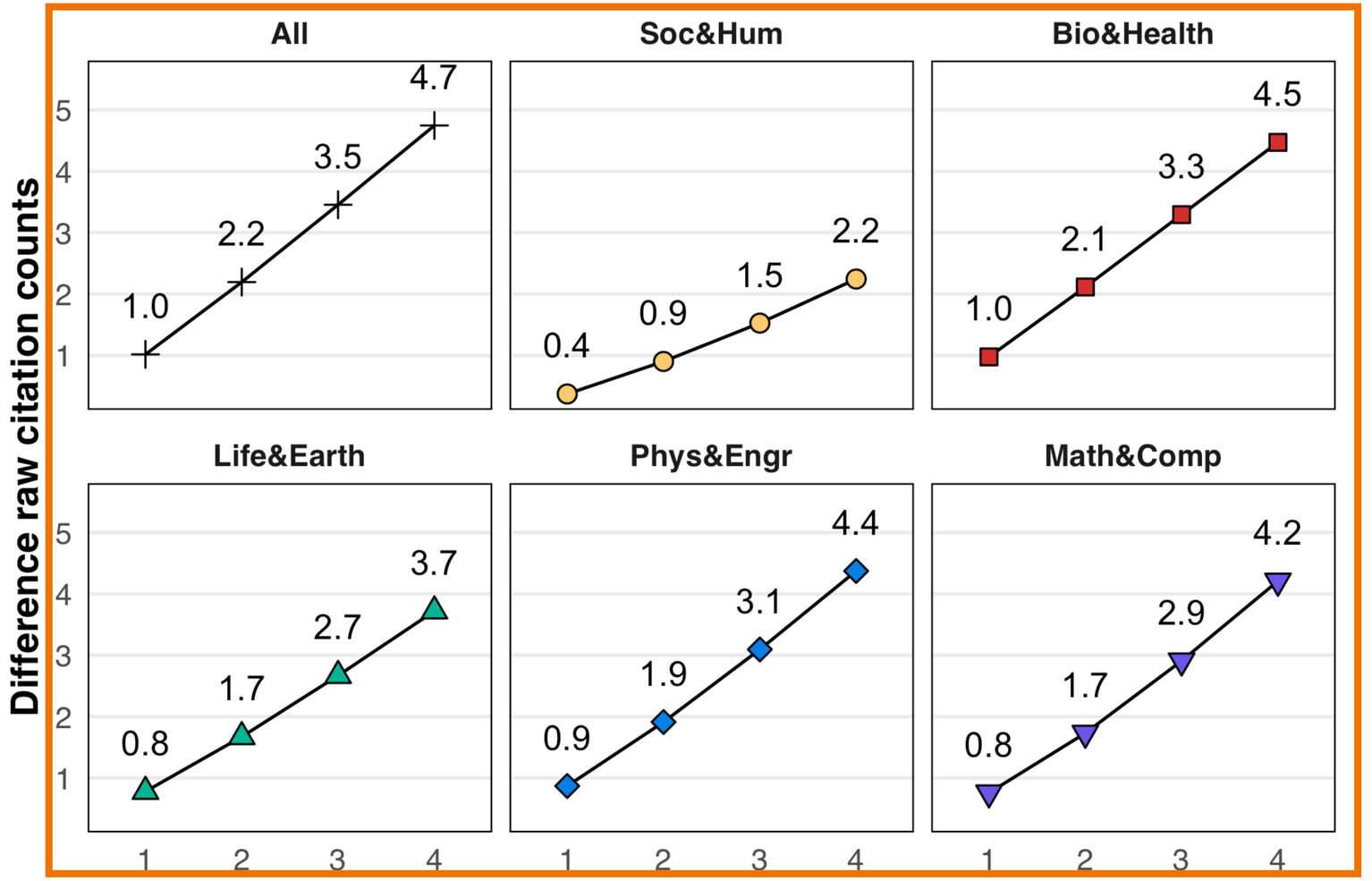
Higher values indicate that papers with disagreement sentences received more citations



Papers with disagreement distances have more citations

True across every major field

Higher values indicate that papers with disagreement distances received more citations



What about being disagreed with?

Compare papers that were similar, when one received a disagreement citation

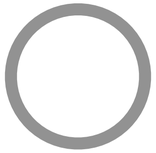
Disagreement

Received 10 citations in 5 years

Received **?** Citations in following year

No Disagreement

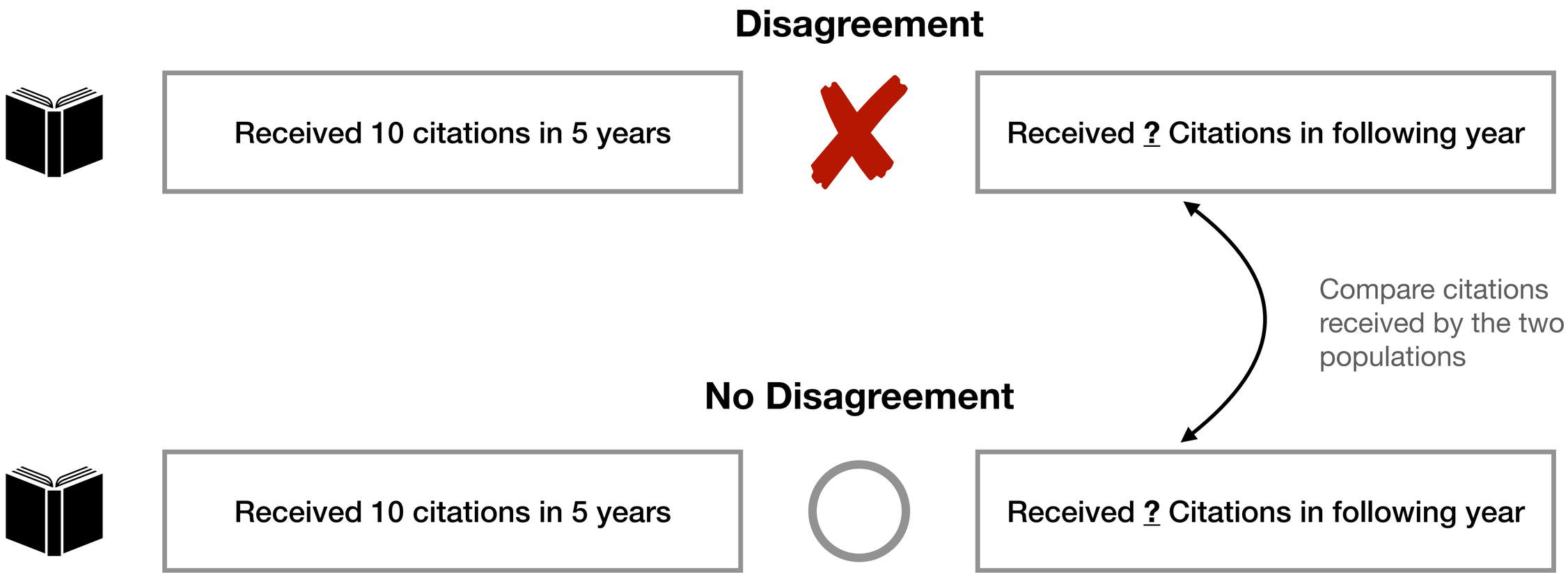
Received 10 citations in 5 years



Received **?** Citations in following year

What about being disagreed with?

Compare papers that were similar, when one received a disagreement citation



Being disagreed with has little effect

Field	Avg. citations in year following disagreement	Expected citations in year following disagreement	Difference
All	3.03	3.08	-0.05
Bio & Health	2.73	2.81	-0.08
Life & Earth	3.43	3.35	+0.08
Math & Comp	3.36	3.34	+0.02
Phys & Engr	3.55	3.52	+0.03
Soc & Hum	3.04	3.11	-0.07

Being disagreed with has little effect

Less than one tenth of a citation difference, in all cases

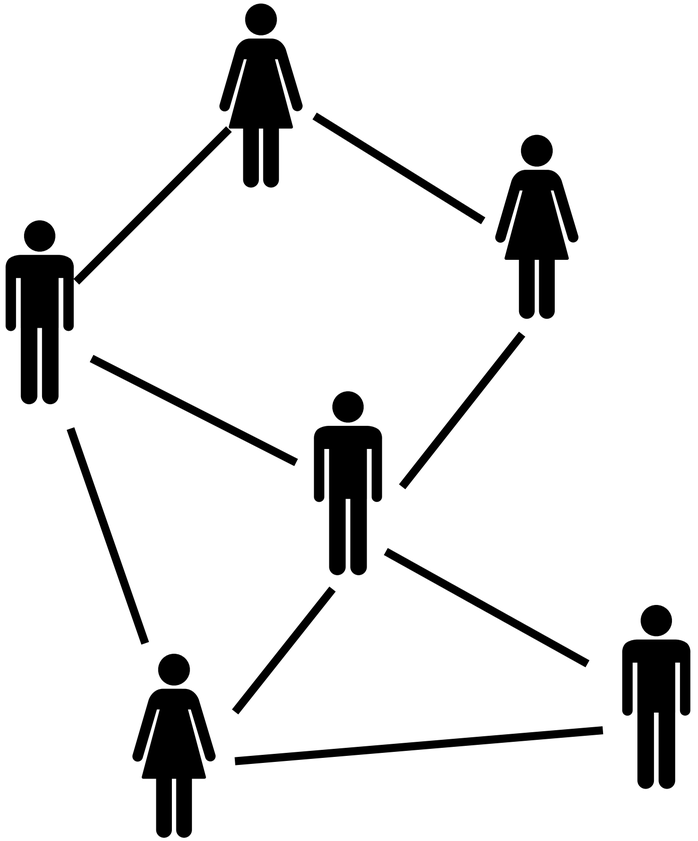
Field	Avg. citations in year following disagreement	Expected citations in year following disagreement	Difference
All	3.03	3.08	-0.05
Bio & Health	2.73	2.81	-0.08
Life & Earth	3.43	3.35	+0.08
Math & Comp	3.36	3.34	+0.02
Phys & Engr	3.55	3.52	+0.03
Soc & Hum	3.04	3.11	-0.07

Difference in citations between those that received a disagreement citation, and those that didn't

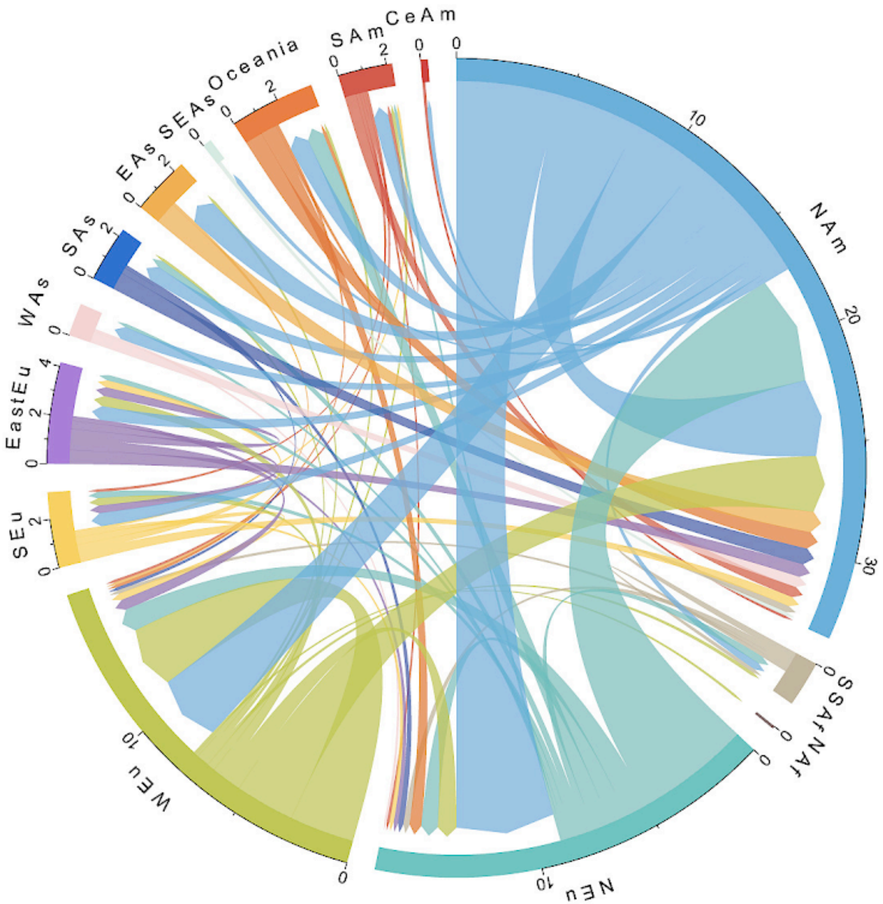
Stay tuned for more on this!

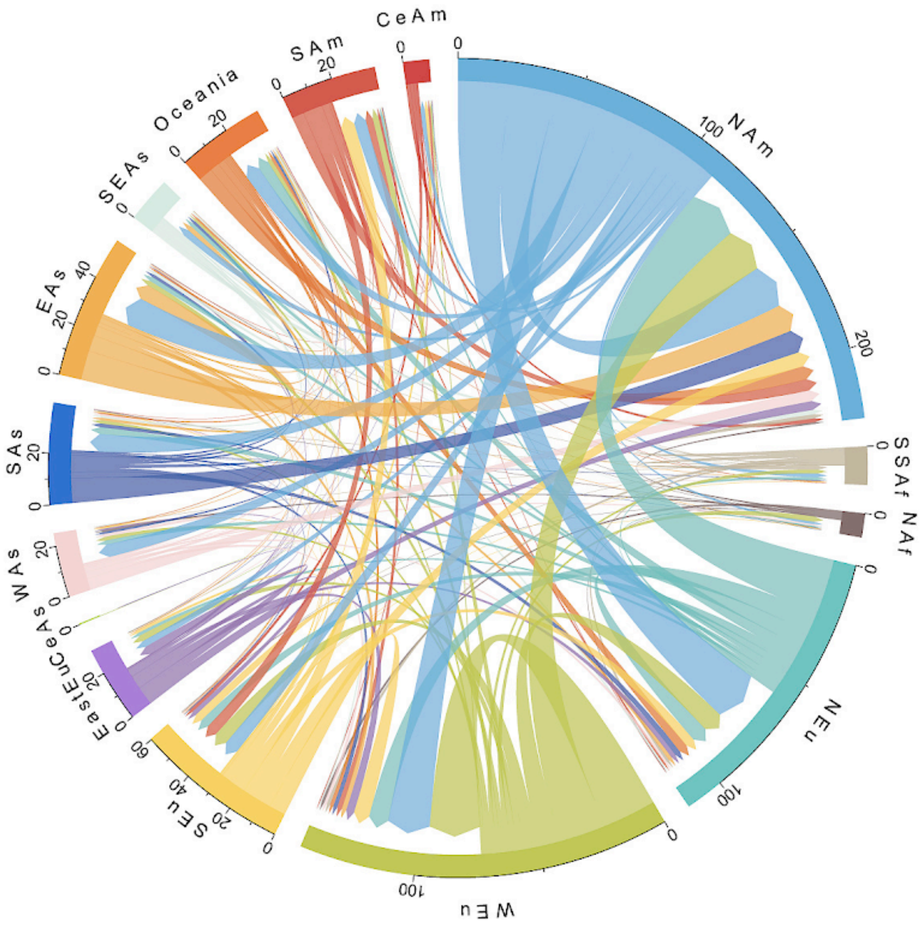
Also, we welcome ideas

Reputation
(networks)



Scientific mobility



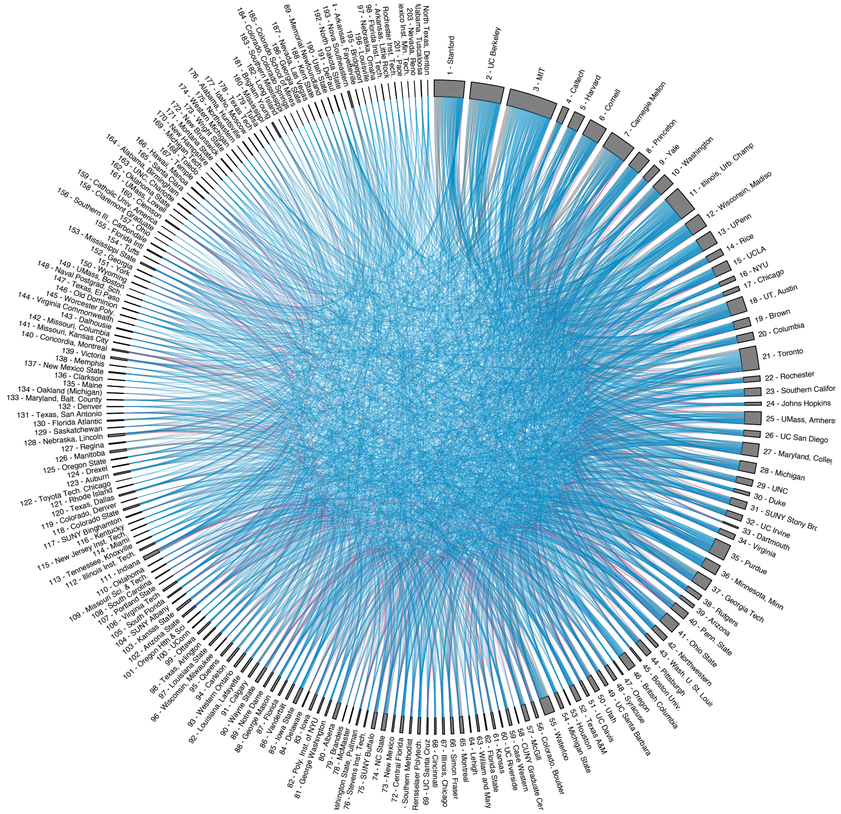


Czaika, M., & Orazbayev, S. (2018). The globalisation of scientific mobility, 1970–2014. *Applied Geography*, *96*, 1–10.

Mobility is intimately tied to evaluation and performance

Mobility is intimately tied to evaluation and performance

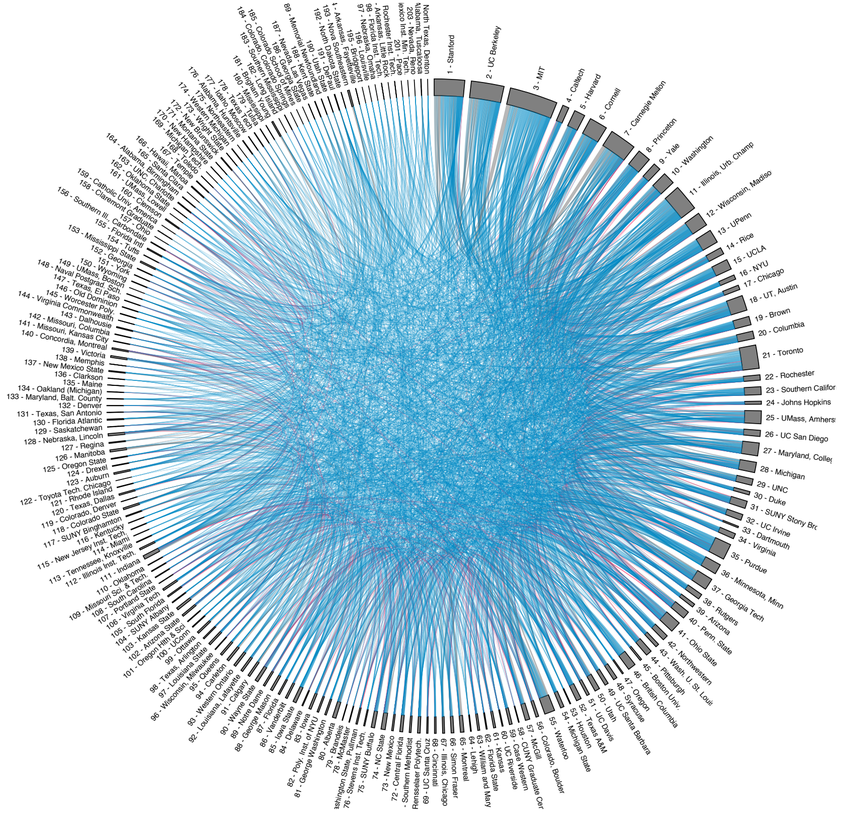
Mobility an output of evaluation in faculty hiring



Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*(1), e1400005.

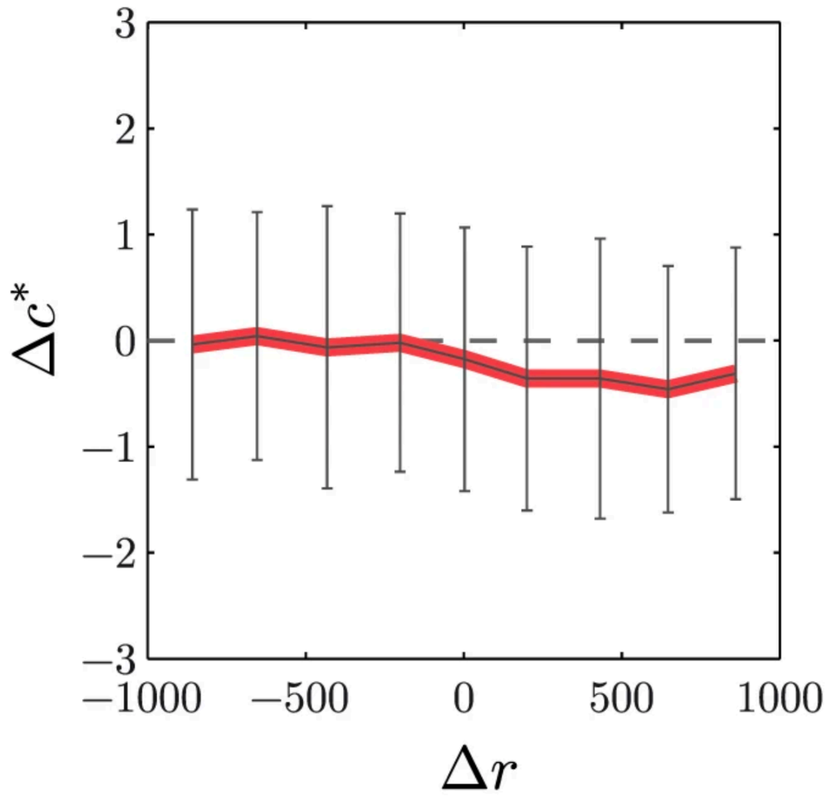
Mobility is intimately tied to evaluation and performance

Mobility an output of evaluation in faculty hiring



Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*(1), e1400005.

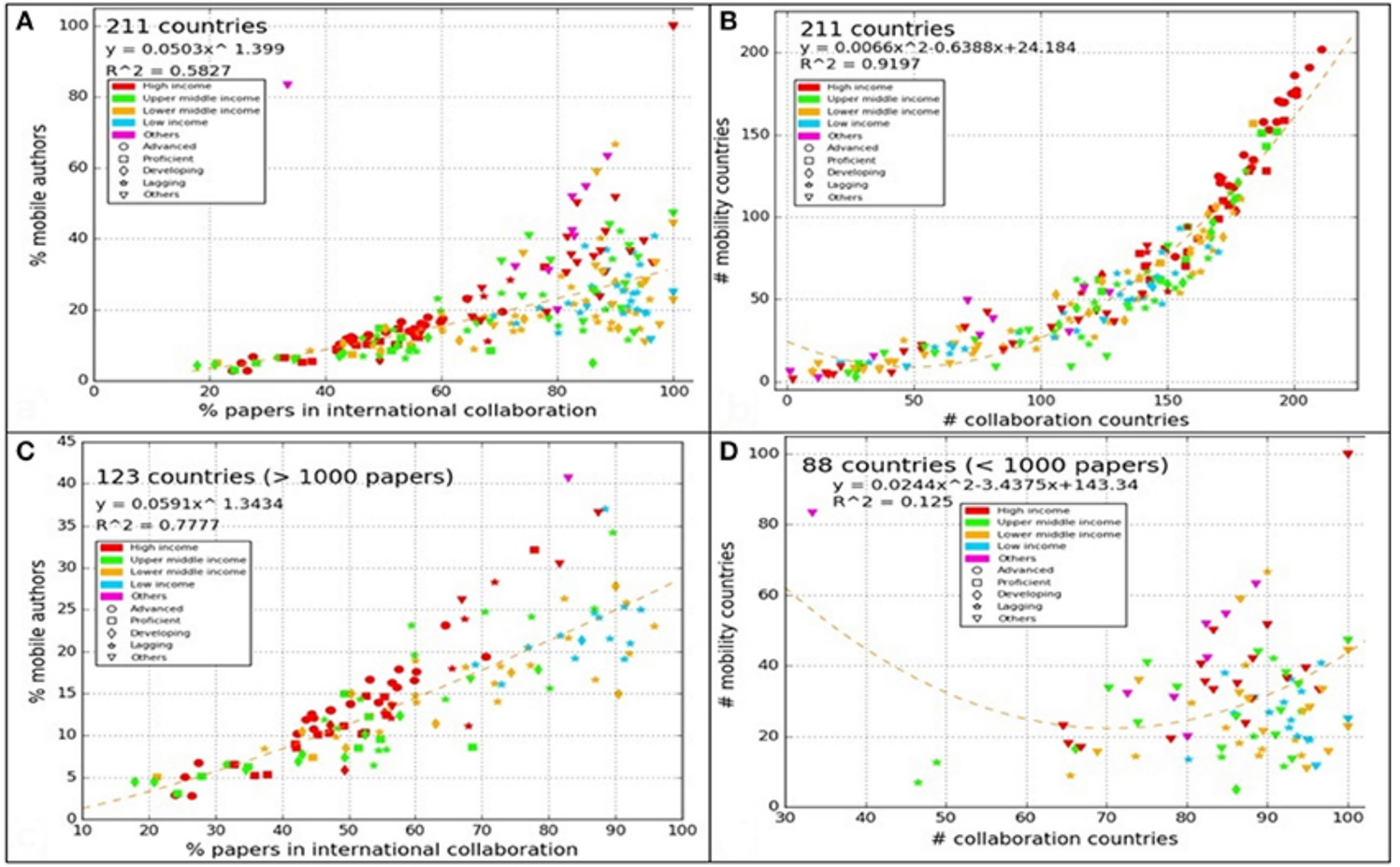
Moving to a low-rank institution lowers impact



Deville, P., Wang, D., Sinatra, R., Song, C., Blondel, V. D., & Barabási, A.-L. (2014). Career on the Move: Geography, Stratification and Scientific Impact. *Scientific Reports*, *4*(1), 1–7.

Mobility expands our professional networks

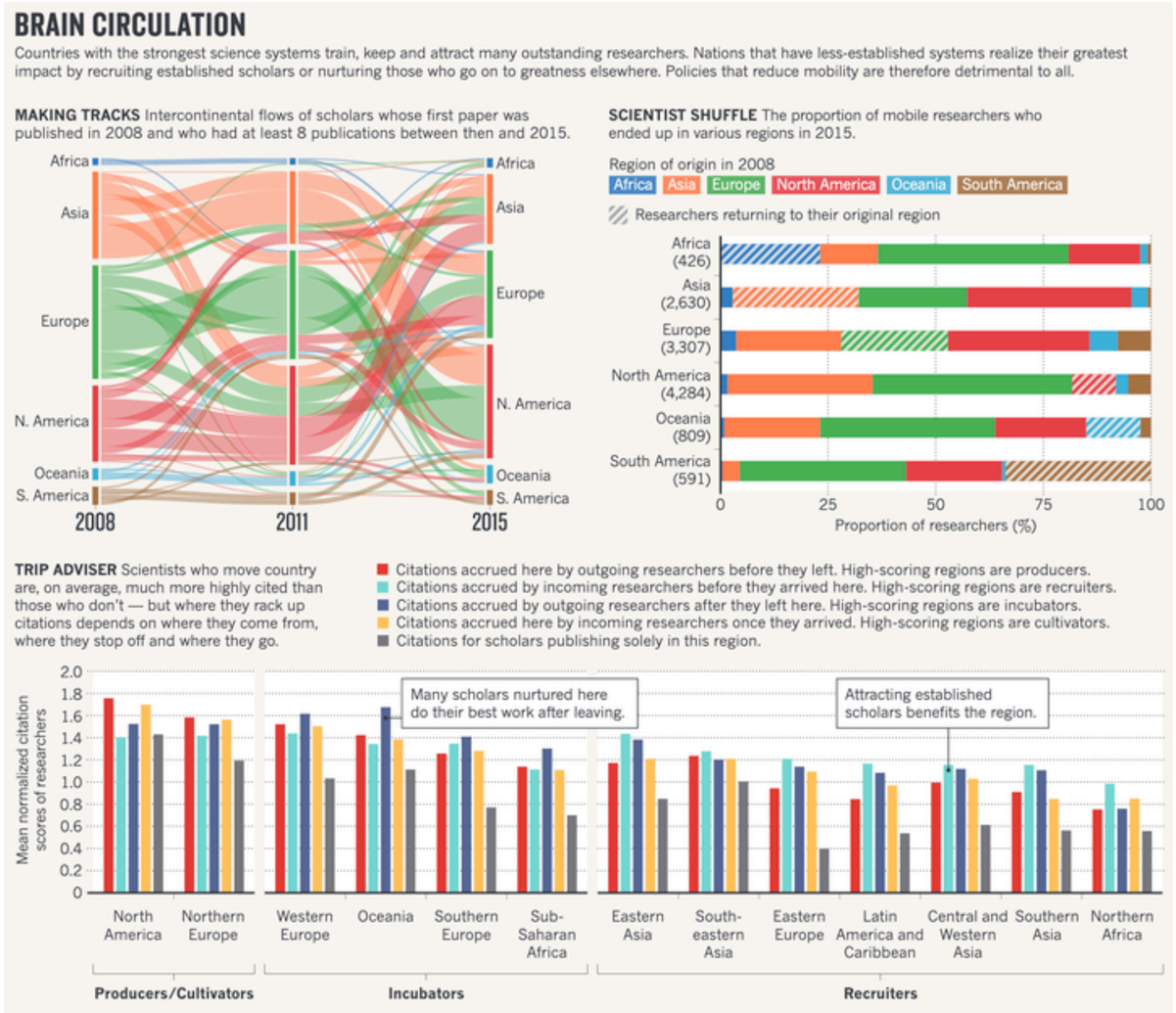
Countries with more mobility between them, also have more collaboration!



Chinchilla-Rodríguez, Z., Miao, L., **Murray, D.**, Robinson-García, N., Costas, R., & Sugimoto, C. R. (2018). A Global Comparison of Scientific Mobility and Collaboration According to National Scientific Capacities. *Frontiers in Research Metrics and Analytics*, *3*.

And mobility fosters improved performance

Scholars who are mobile have higher citation impact!



Sugimoto, C. R., Robinson-Garcia, N., **Murray, D. S.**, Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, *550*(7674), 29–31.

Mobility is deeply contextual

Not all mobility is the same: how and where you move depends on context

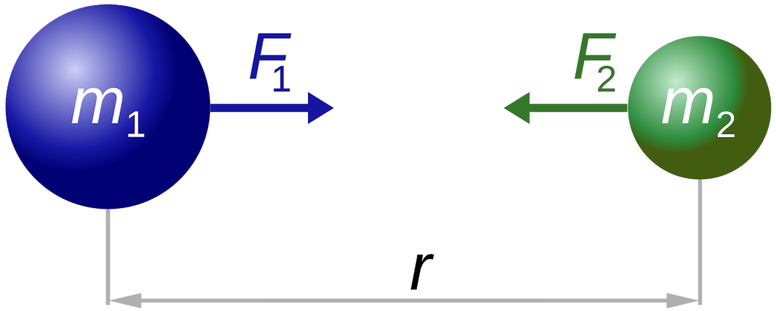
Mobility is deeply contextual

Not all mobility is the same: how and where you move depends on context

Models and techniques do exist for understanding mobility

Gravity model

Ubiquitous and intuitive



$$F_1 = F_2 = G\frac{m_1 \times m_2}{r^2}$$

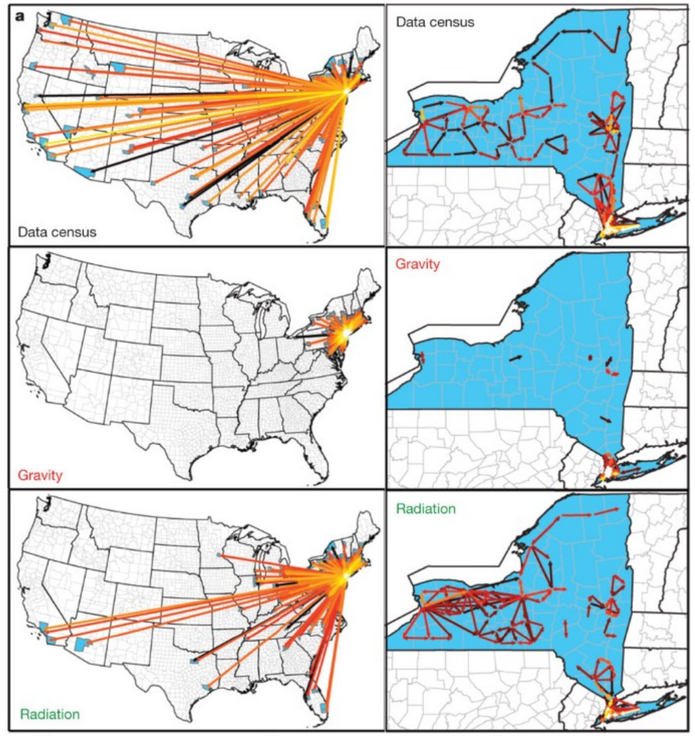
Wikipedia user Dennis Nilsson

$$\hat{T}_{ij} = C m_i m_j f(r_{ij})$$

Flu *Populati* *Decay*

Radiation model

Improve upon the gravity model

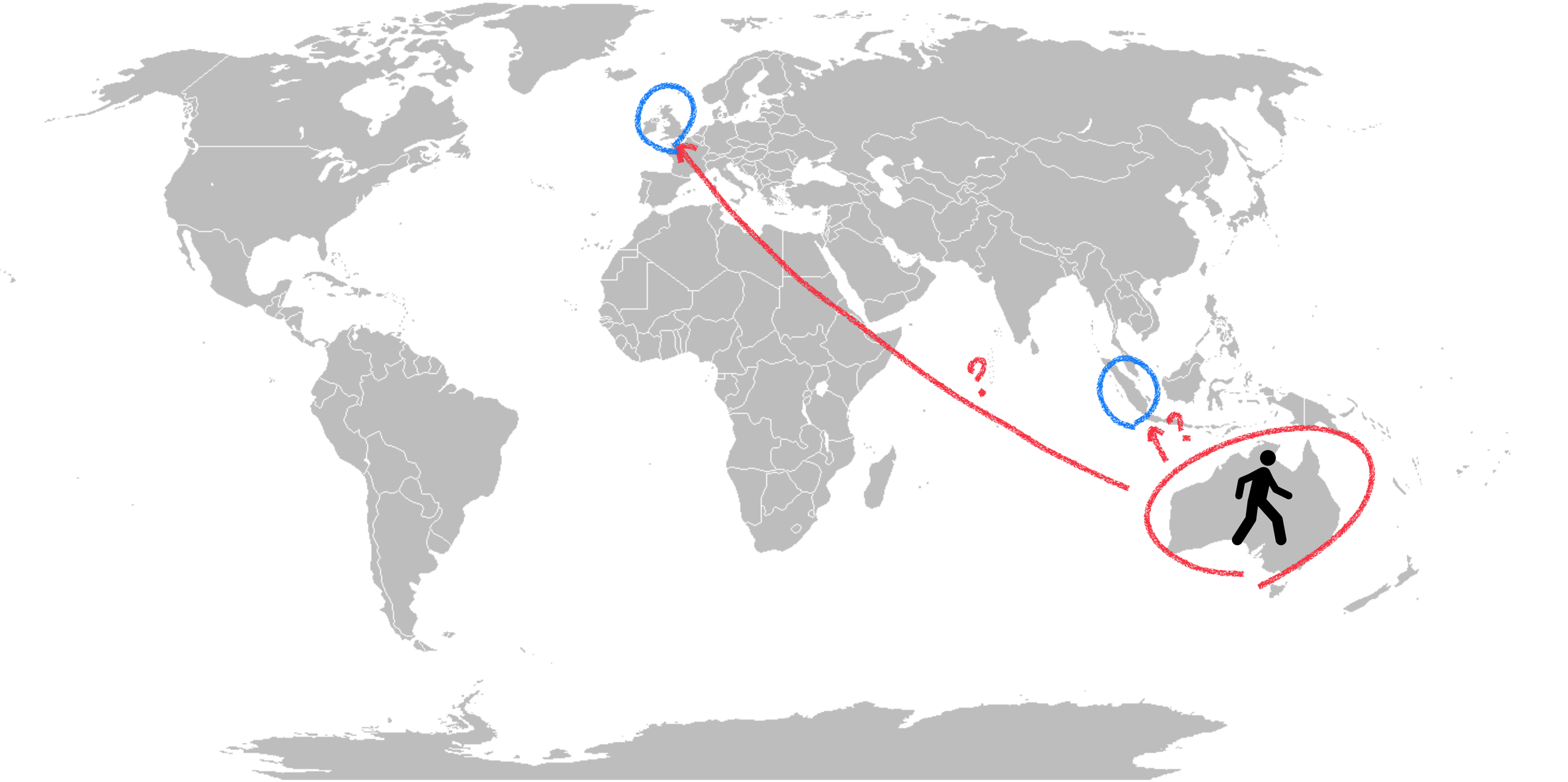


$$[T_{ij}] = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})}$$

Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, *484*(7392), 96–100.

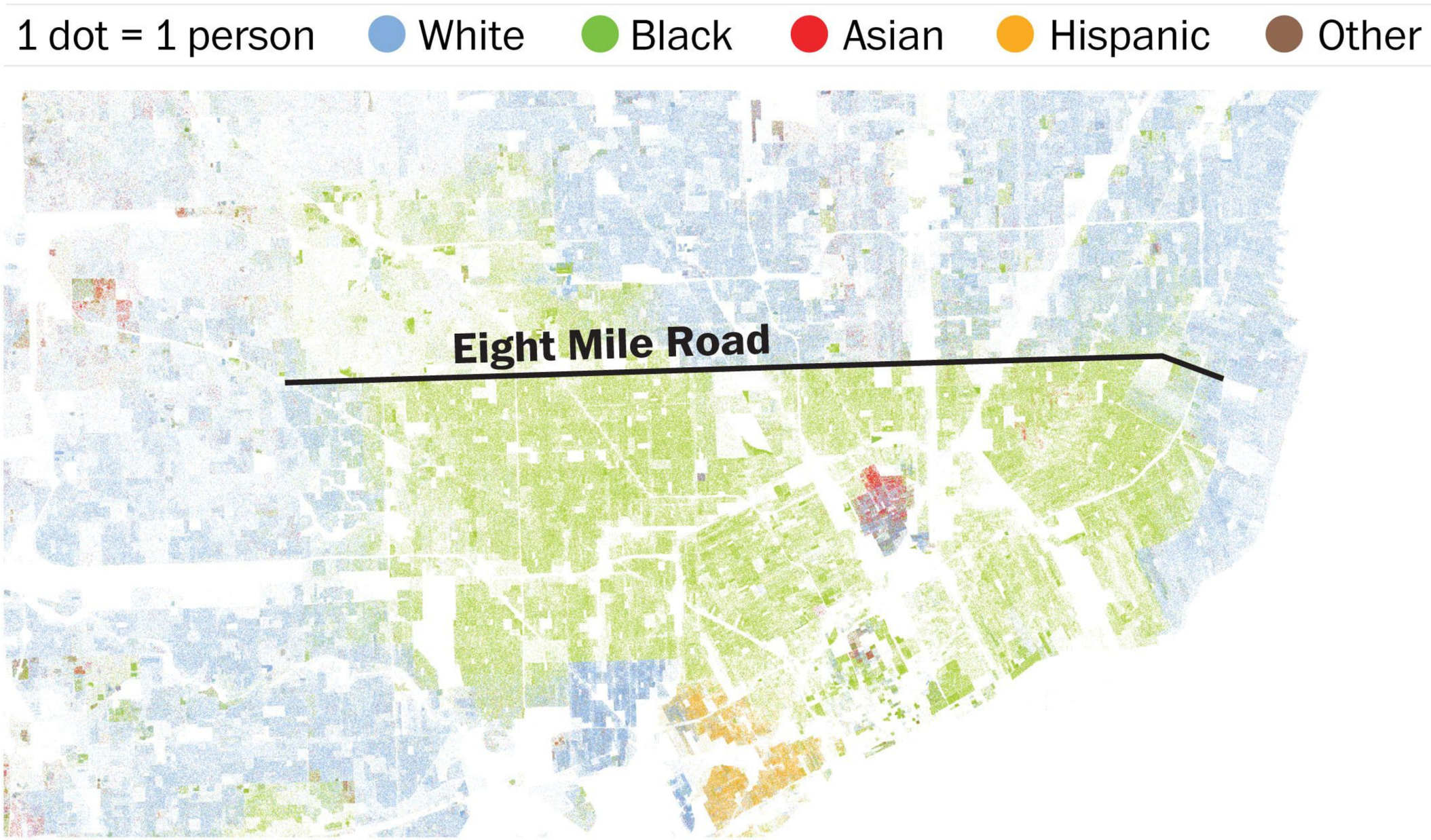
Mobility models are effective…

But geographic distance is not always appropriate



Lines of segregation in Detroit

Economic opportunity and demographic makeup can wildly diverse in just a few city blocks



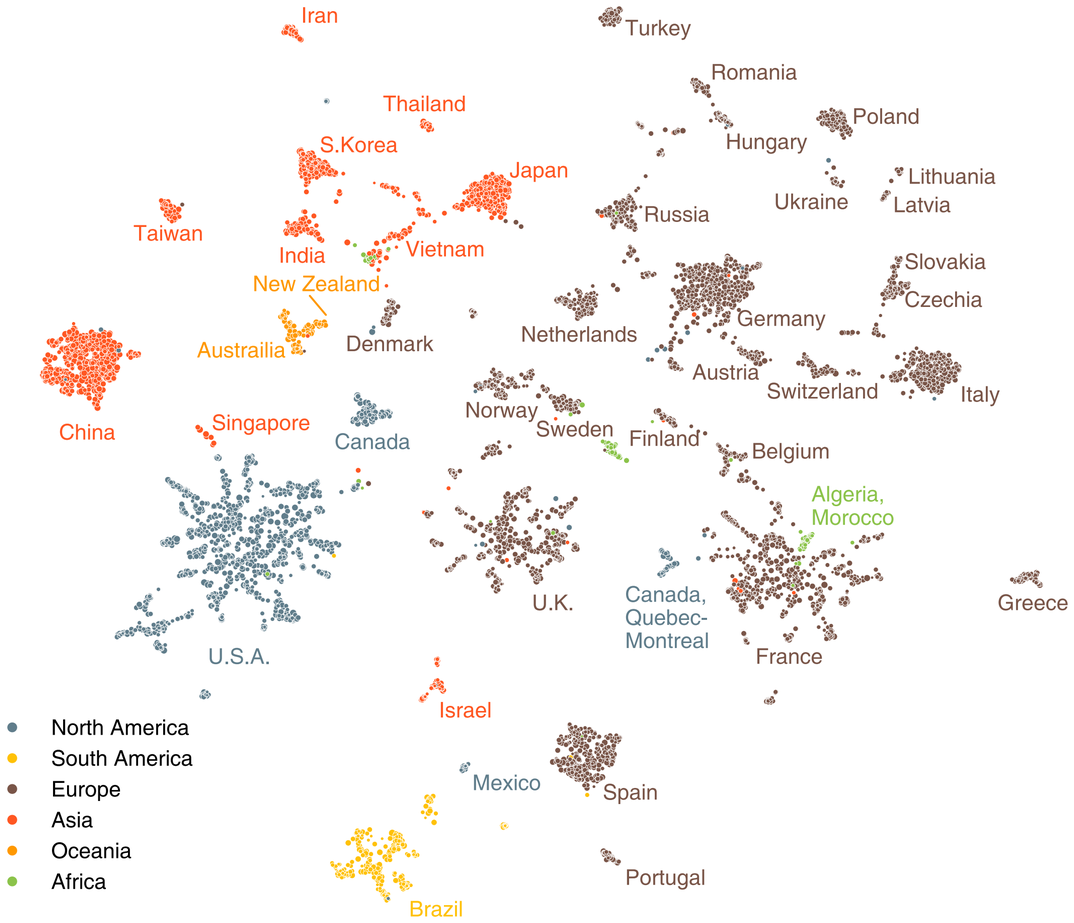
Source: U-Va. Cooper Center analysis of 2010 Census data

THE WASHINGTO

Geography matters, but…

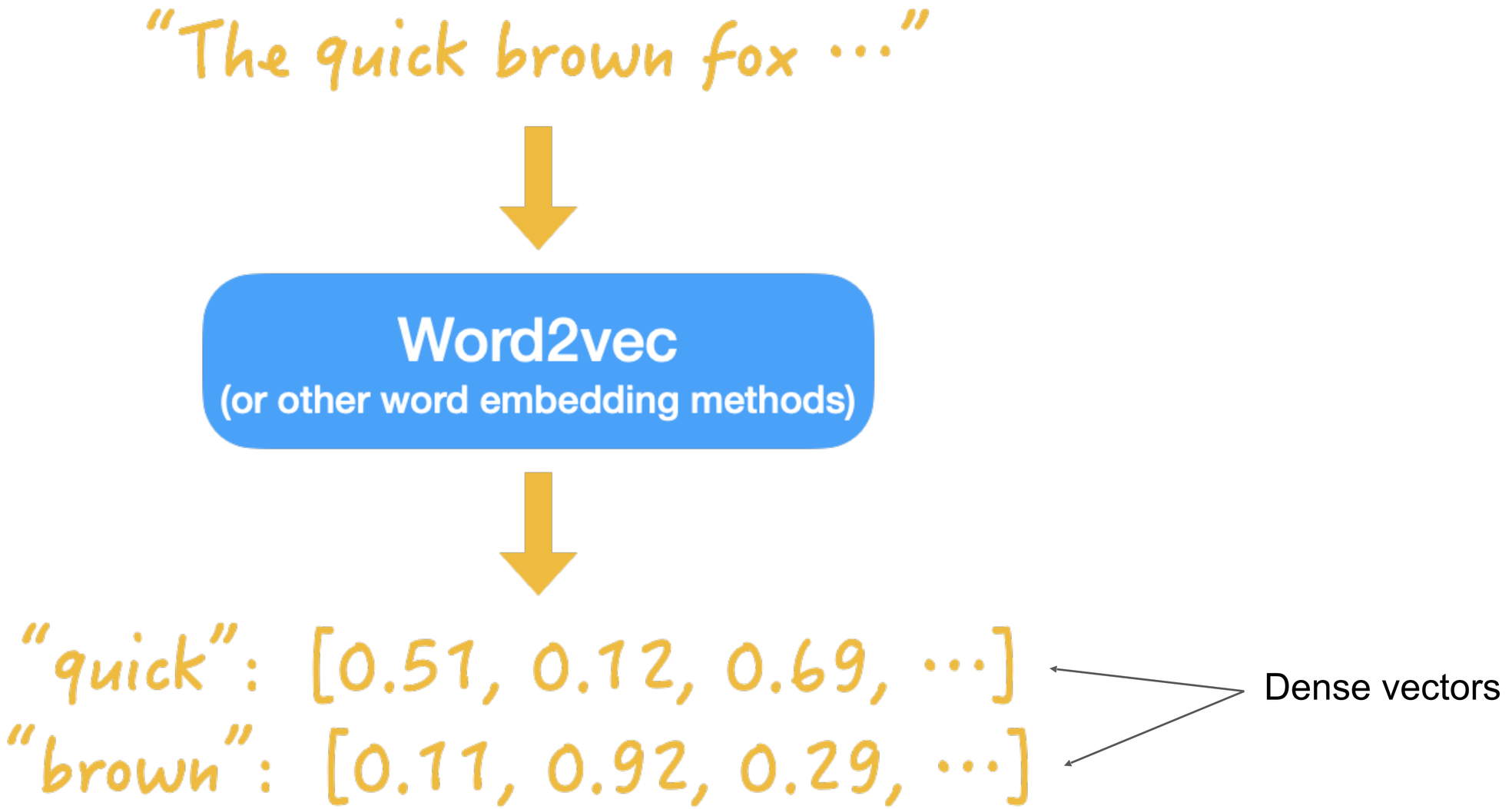
Cultural, linguistic, economic, and political distance are also important!

Can we instead learn an “embedding” that captures the distance between places?



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

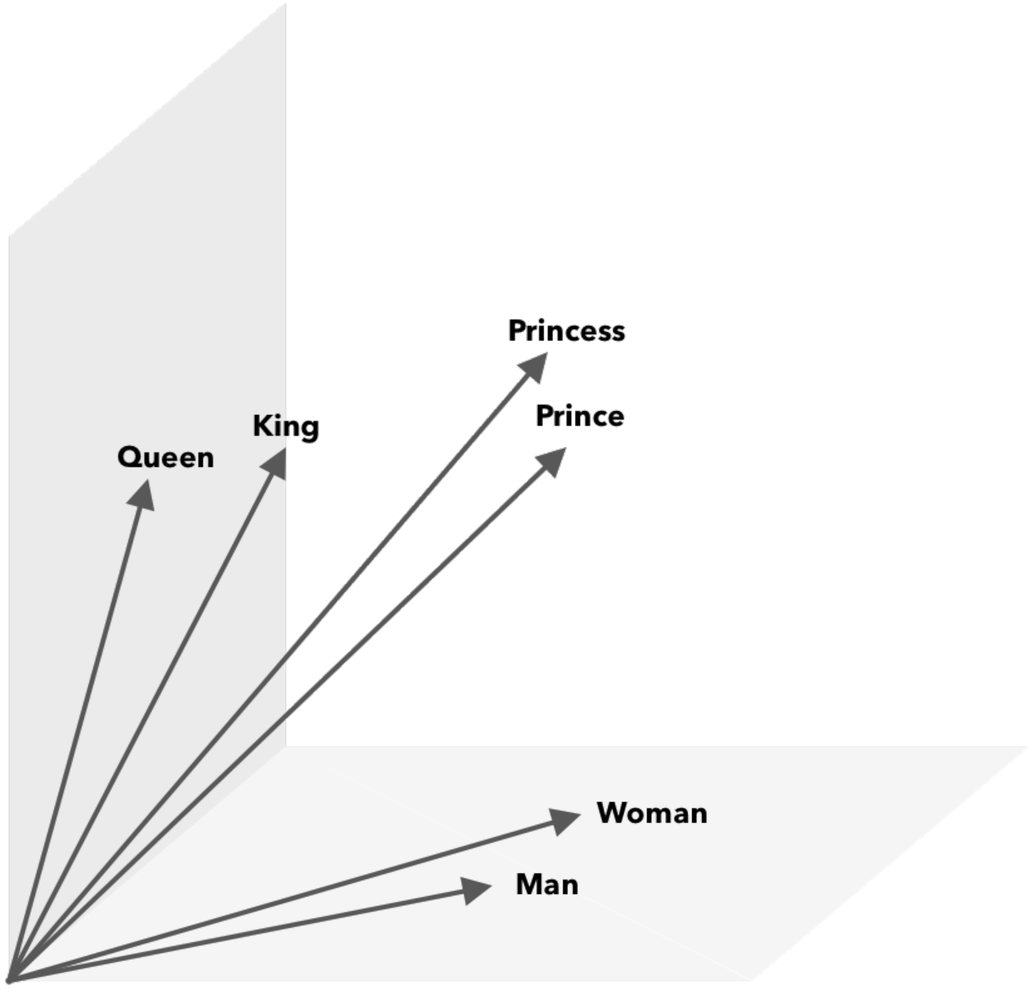
Neural embedding



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The geometry of the vector space encodes semantic relationships

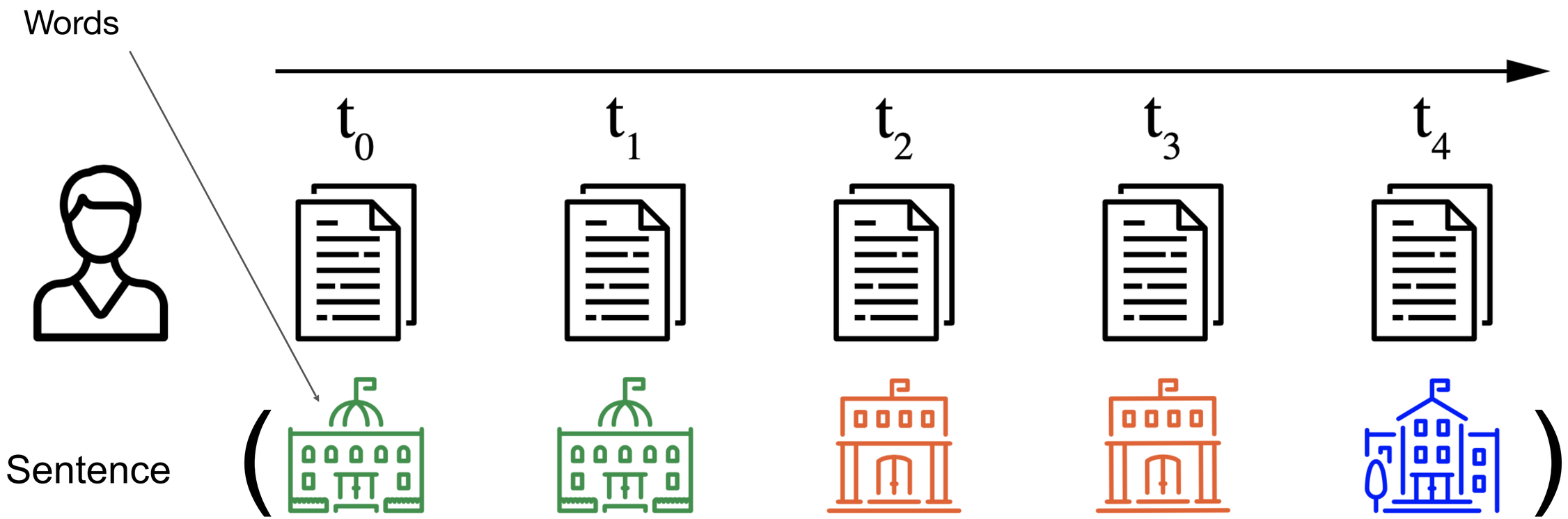
Using cosine distance



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

They don’t have to be “real” words or sentences

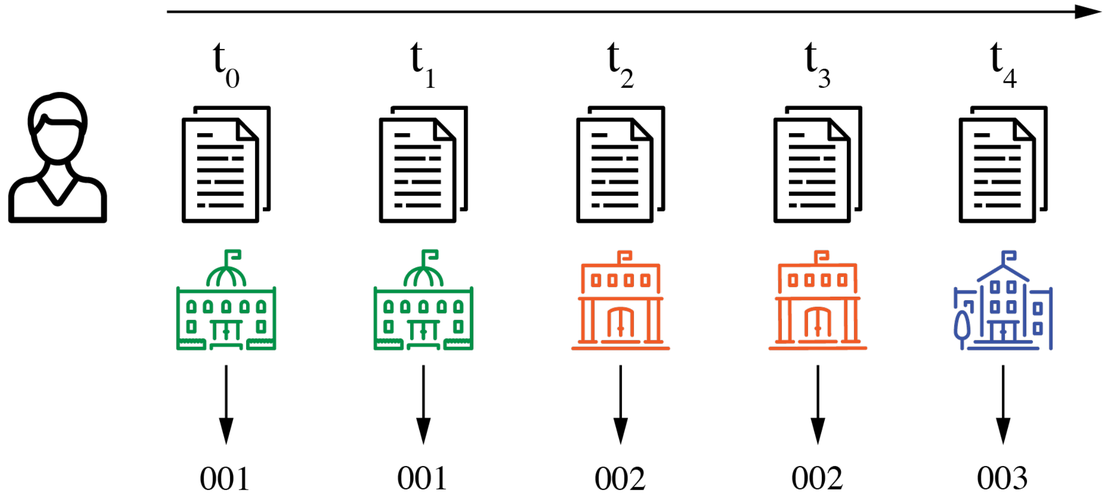
Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use **trajectories of scientists** as **sentences** and **organizations** as **words**.



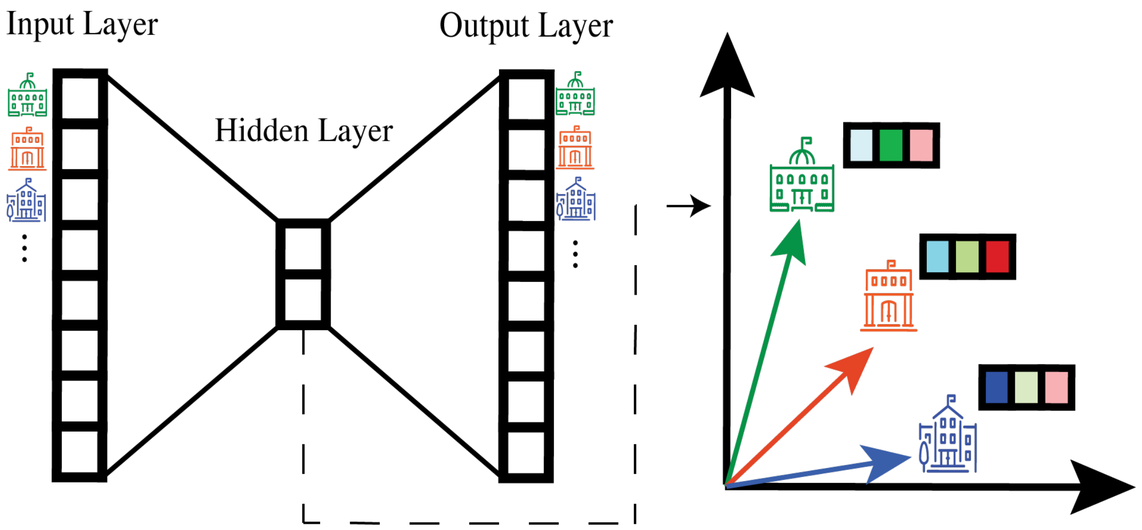
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Modelling mobility

- Career trajectories of 3 million scientists derived from publications
- Give as input to *word2vec*
- Can measure embedding distance between any pair of organizations



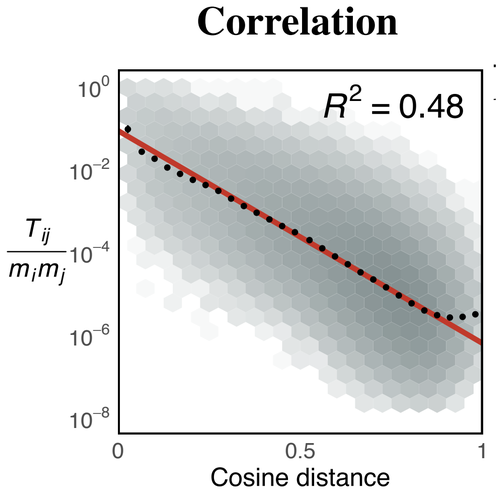
Trajectory: "001 — 001 — 002 — 002 — 003"



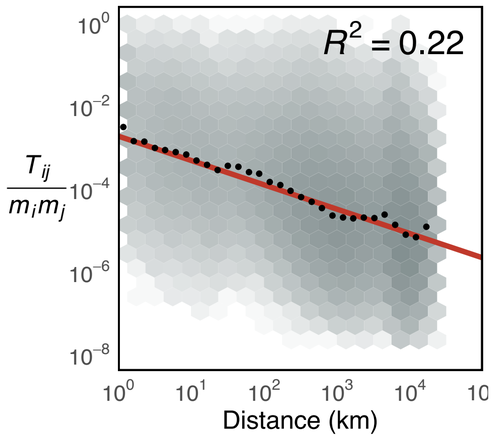
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Embedding Distance



Geographic Distance

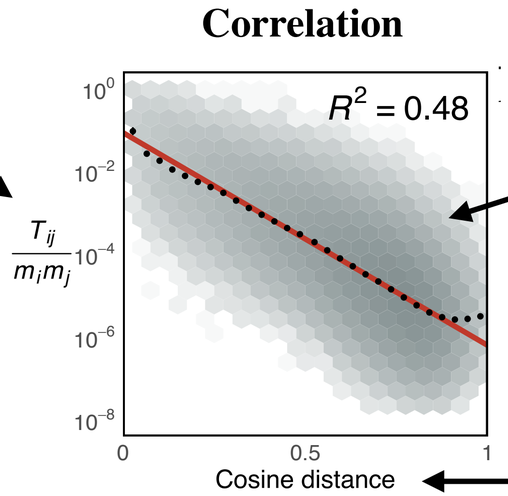


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Flux, given the organizations sizes

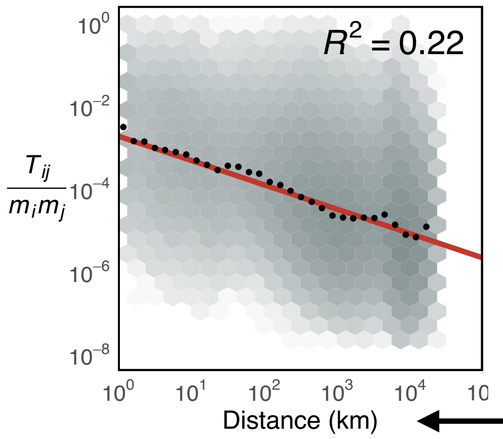
Embedding Distance



Every point is a pair of organizations (binned)

Cosine distance between organization vectors

Geographic Distance



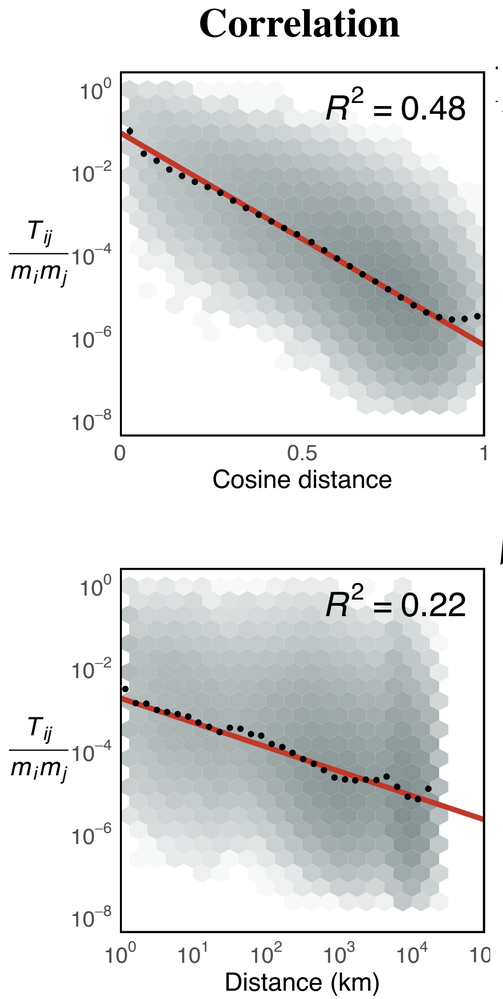
Geographic distance between organizations

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Embedding Distance

Geographic Distance



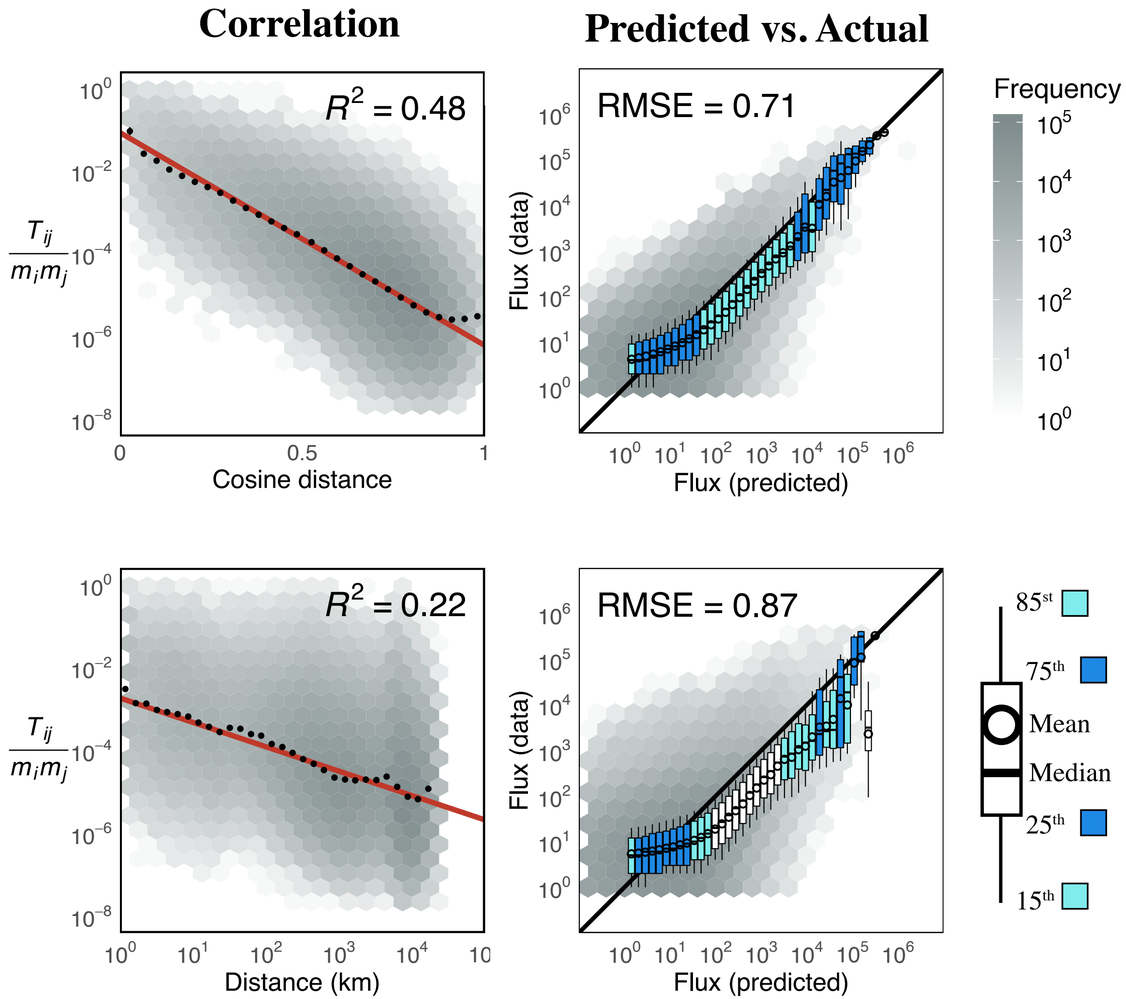
Flux more strongly correlates with embedding distance than geographic distance

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

Embedding distance outperforms geographic distance

Embedding Distance

Geographic Distance

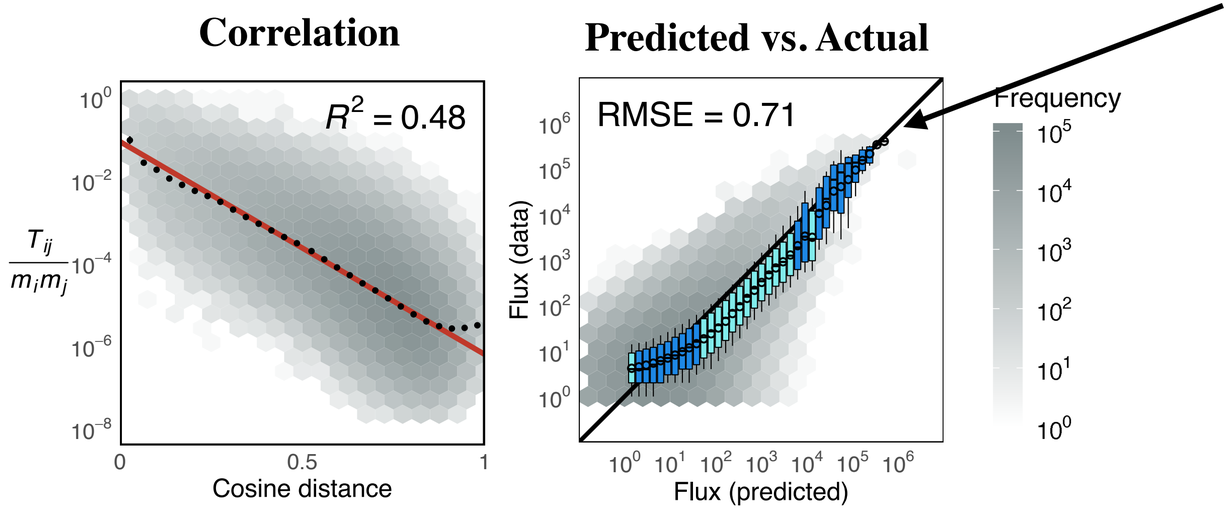


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

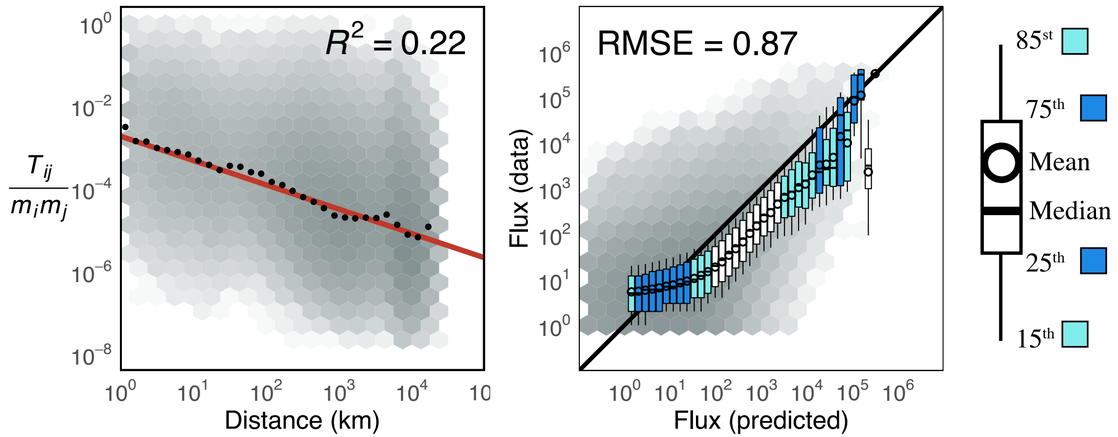
Embedding distance outperforms geographic distance

Distribution of predicted vs. actual flux using gravity model (binned)

Embedding Distance



Geographic Distance



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

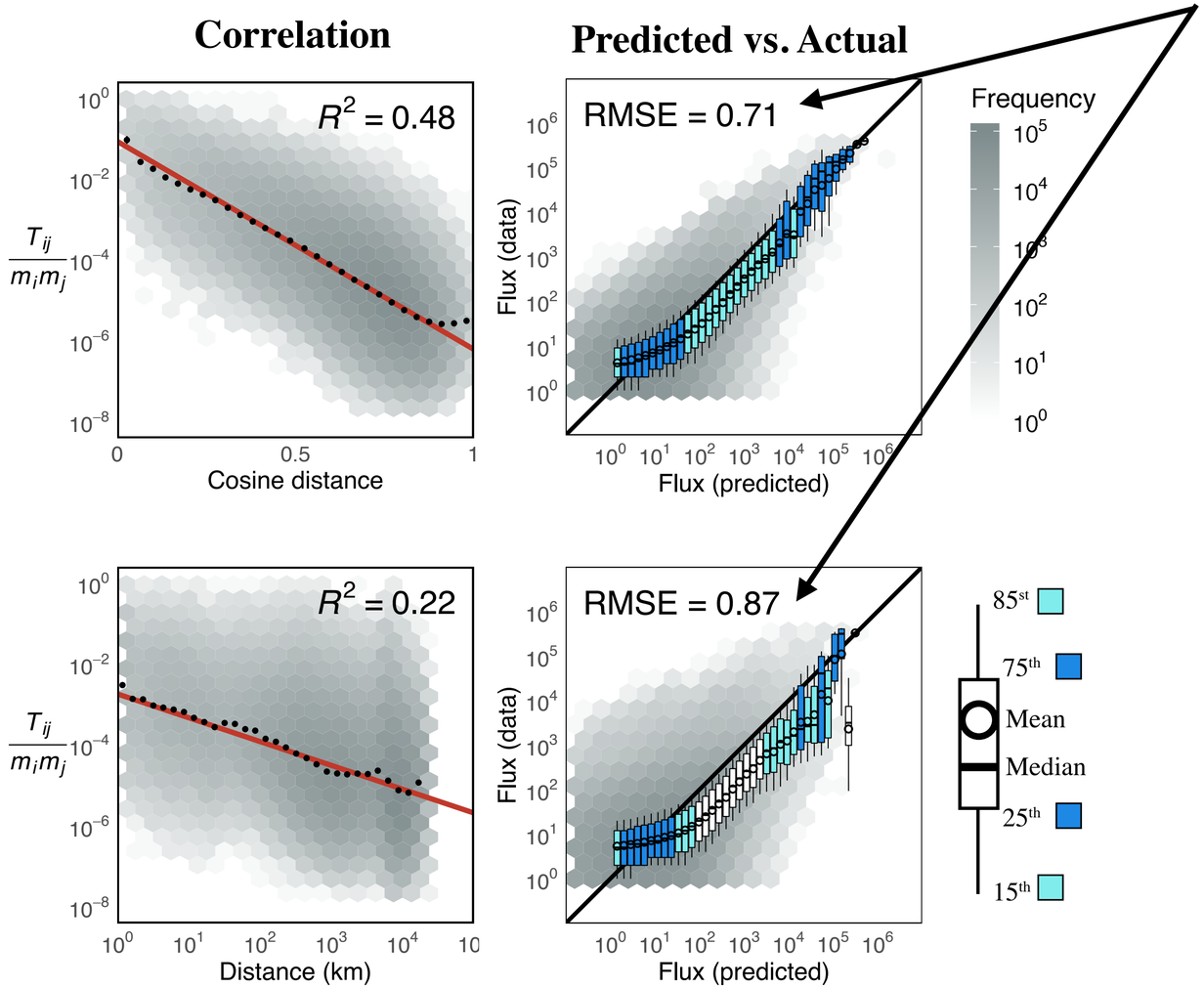
Embedding distance outperforms geographic distance

Root mean squared error of prediction

Embedding distance leads to better predictions

Embedding Distance

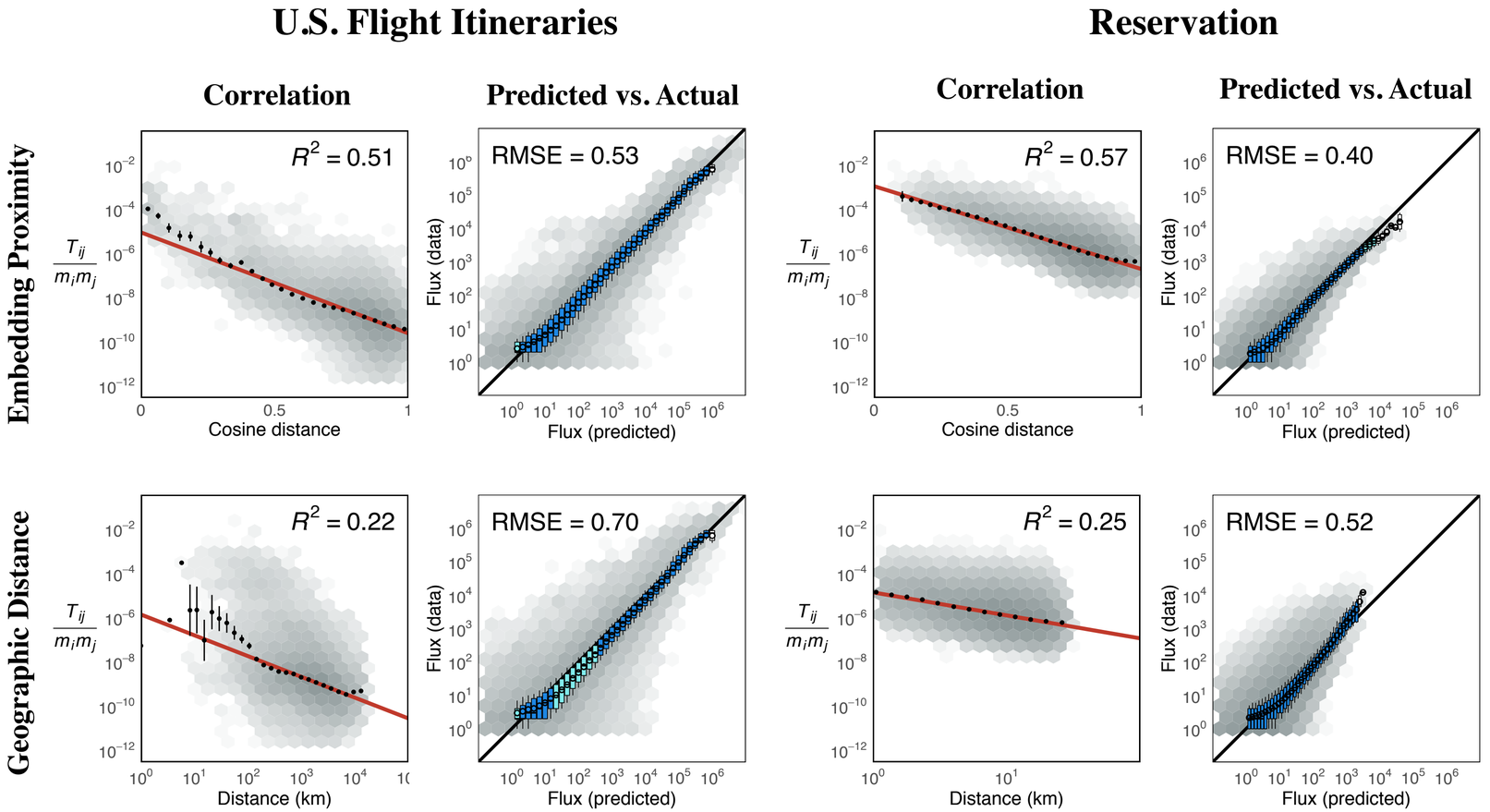
Geographic Distance



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The embedding performs well in other domains!

U.S. domestic flight itineraries and South Korean hotel reservation trajectories

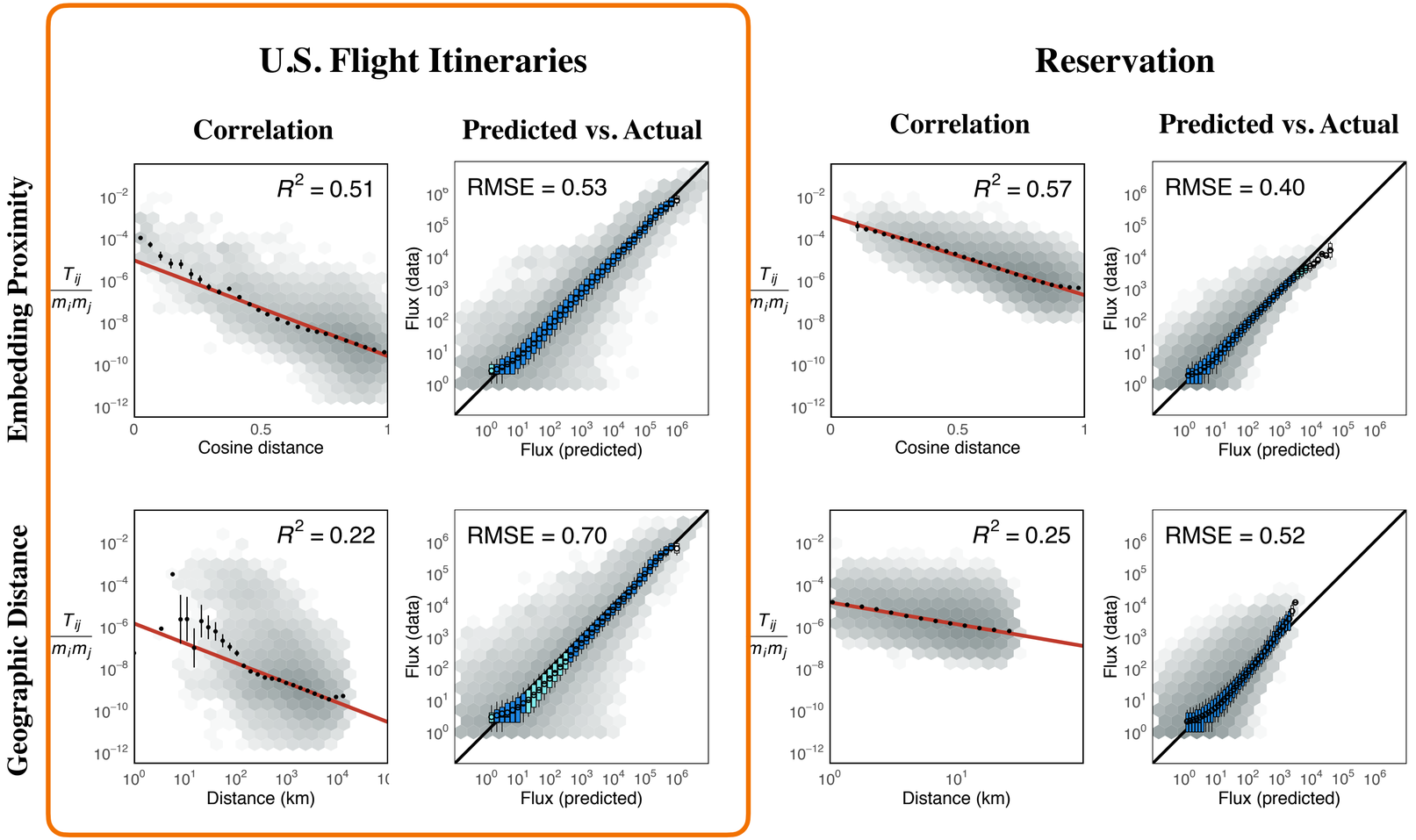


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The embedding performs well in other domains!

U.S. domestic flight itineraries and South Korean hotel reservation trajectories

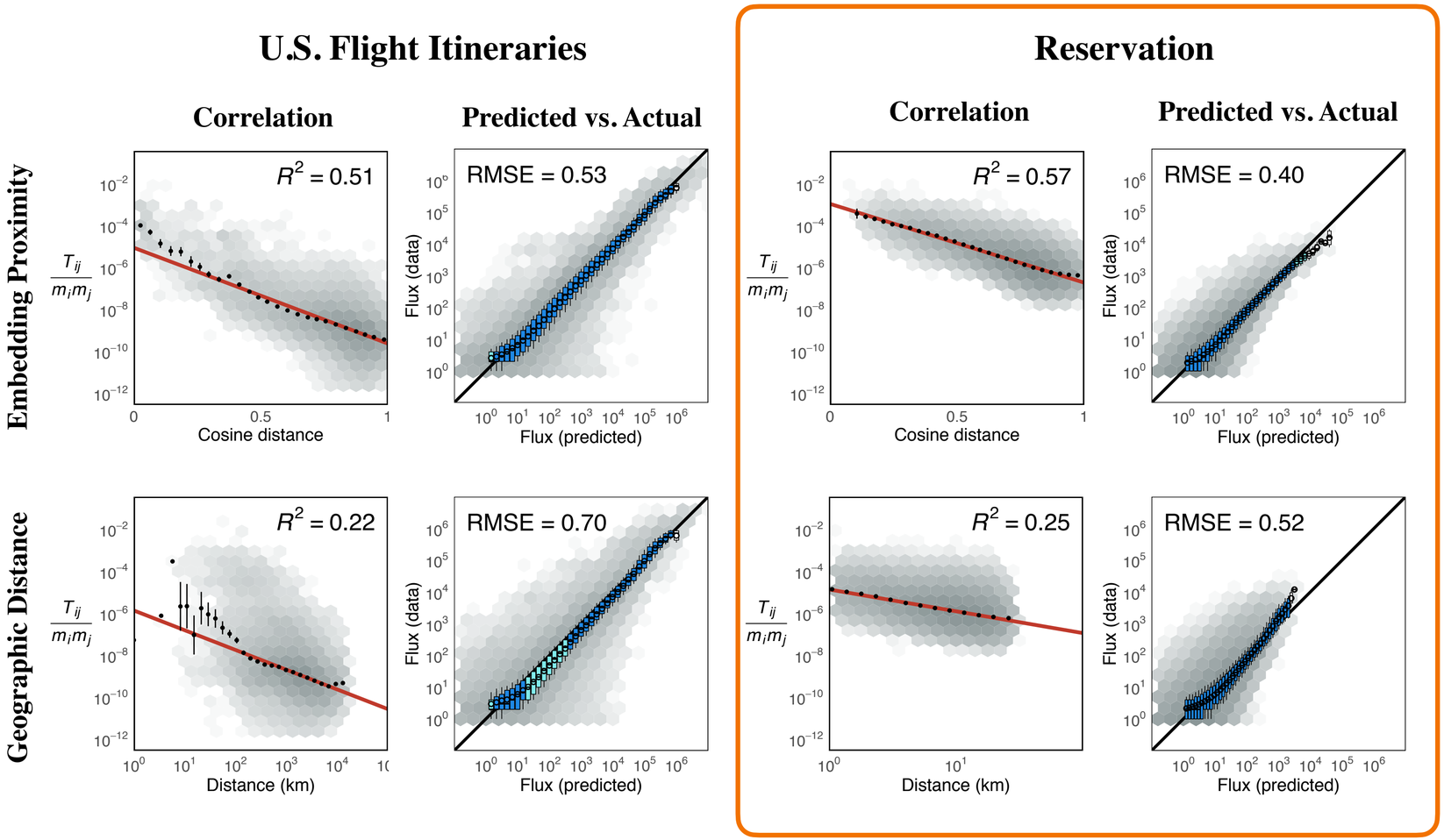
Better performance using embedding distance on trajectories derived from U.S. domestic flights



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

The embedding performs well in other domains!

U.S. domestic flight itineraries and South Korean hotel reservation trajectories



And again for South Korean hotel reservation trajectories

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(In revision)*

word2vec captures the latent structure of mobility

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

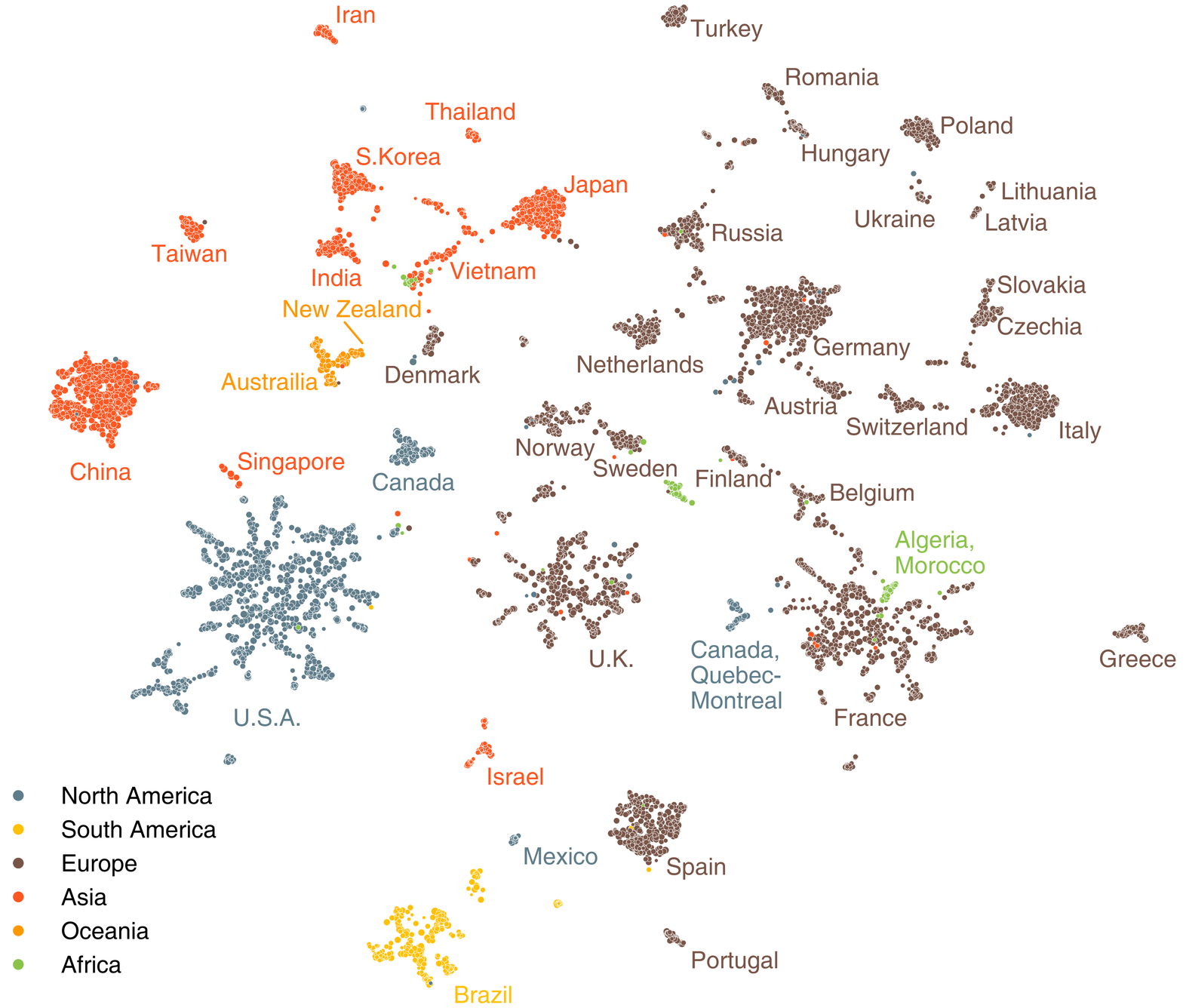
word2vec captures the latent structure of mobility

What is that structure?

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

Visualizing the embedding space

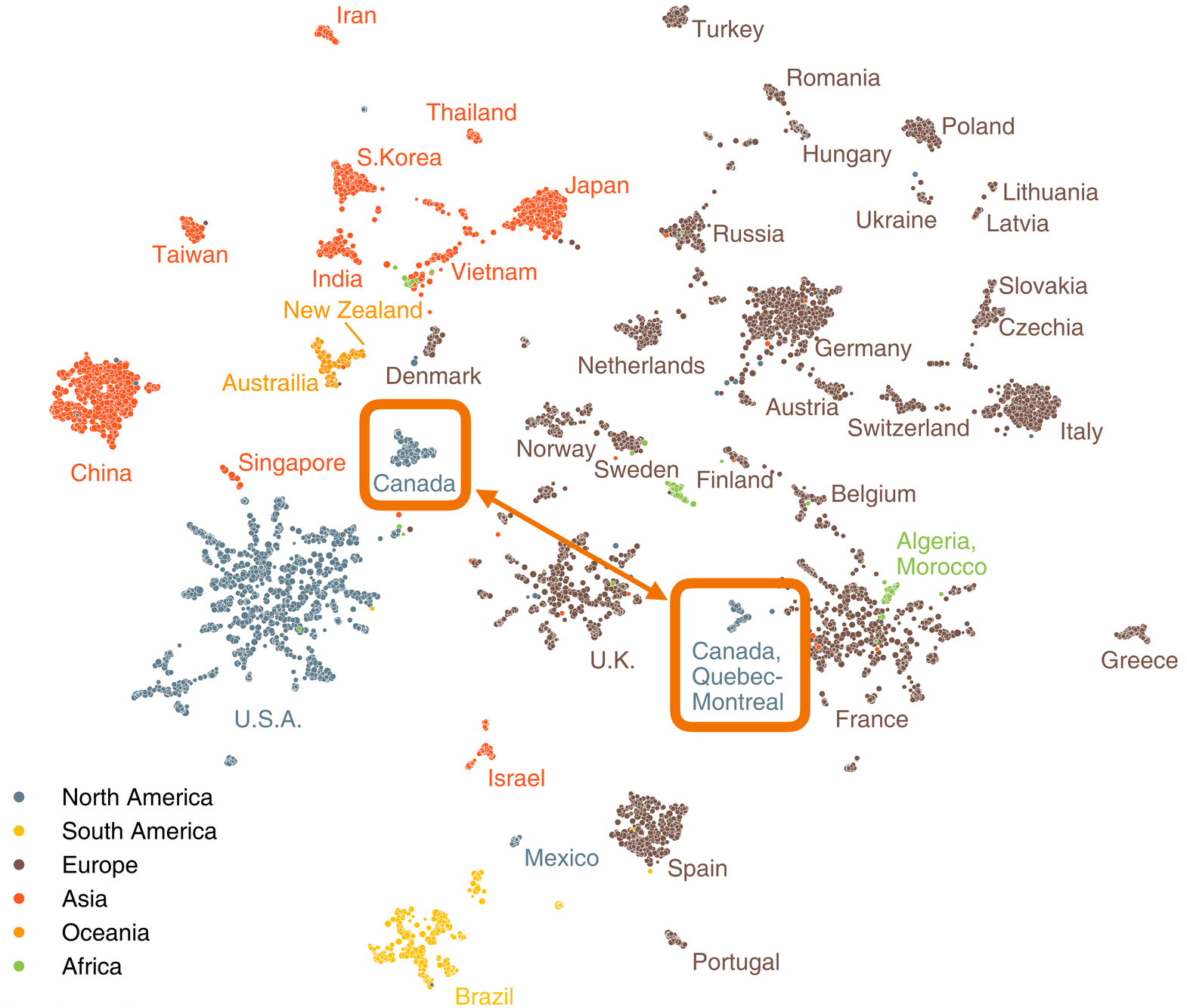
UMAP projection of organizations



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Visualizing the embedding space

Canada, Quebec, & French



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

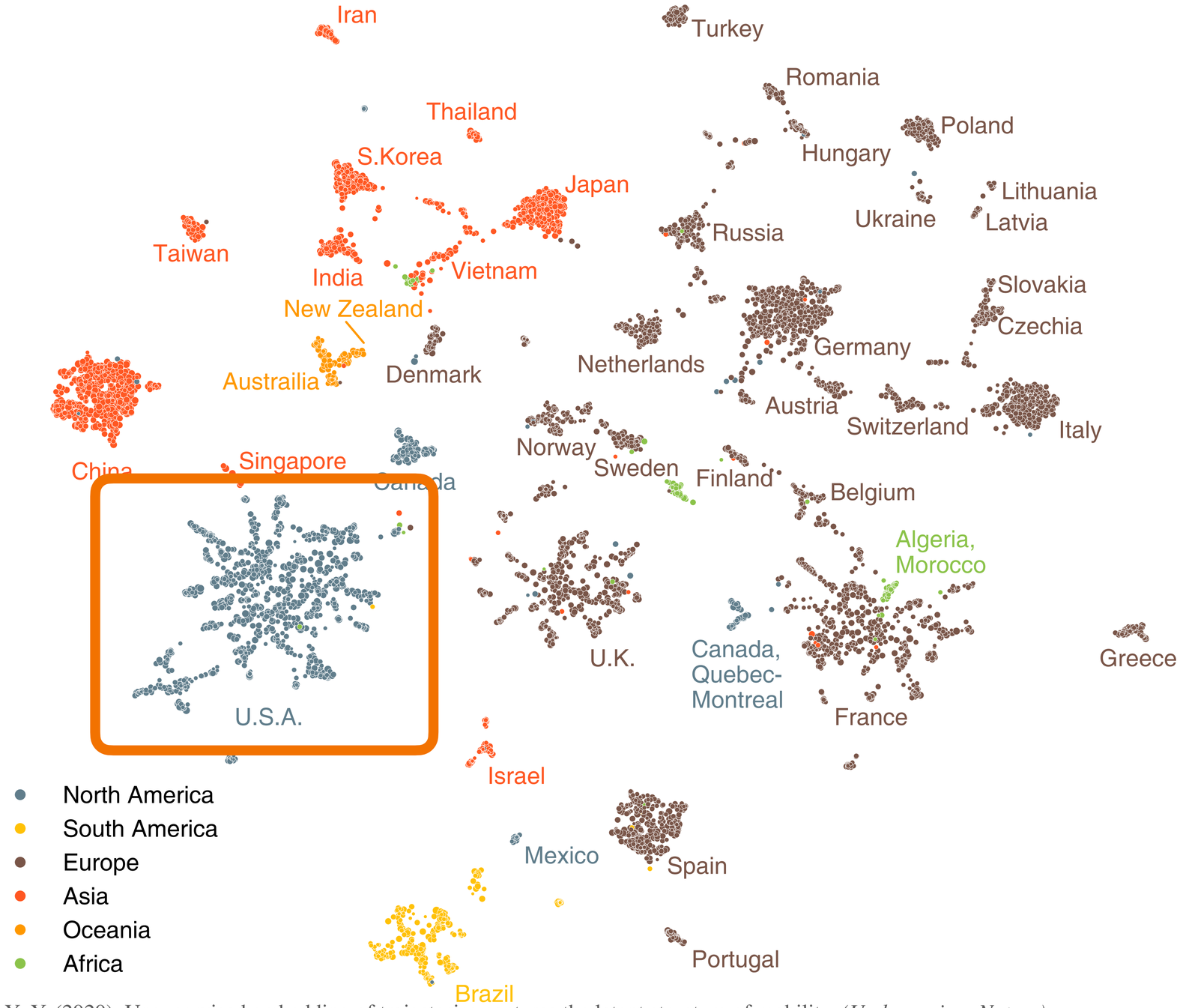
Visualizing the embedding space

South America & the Iberian countries

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

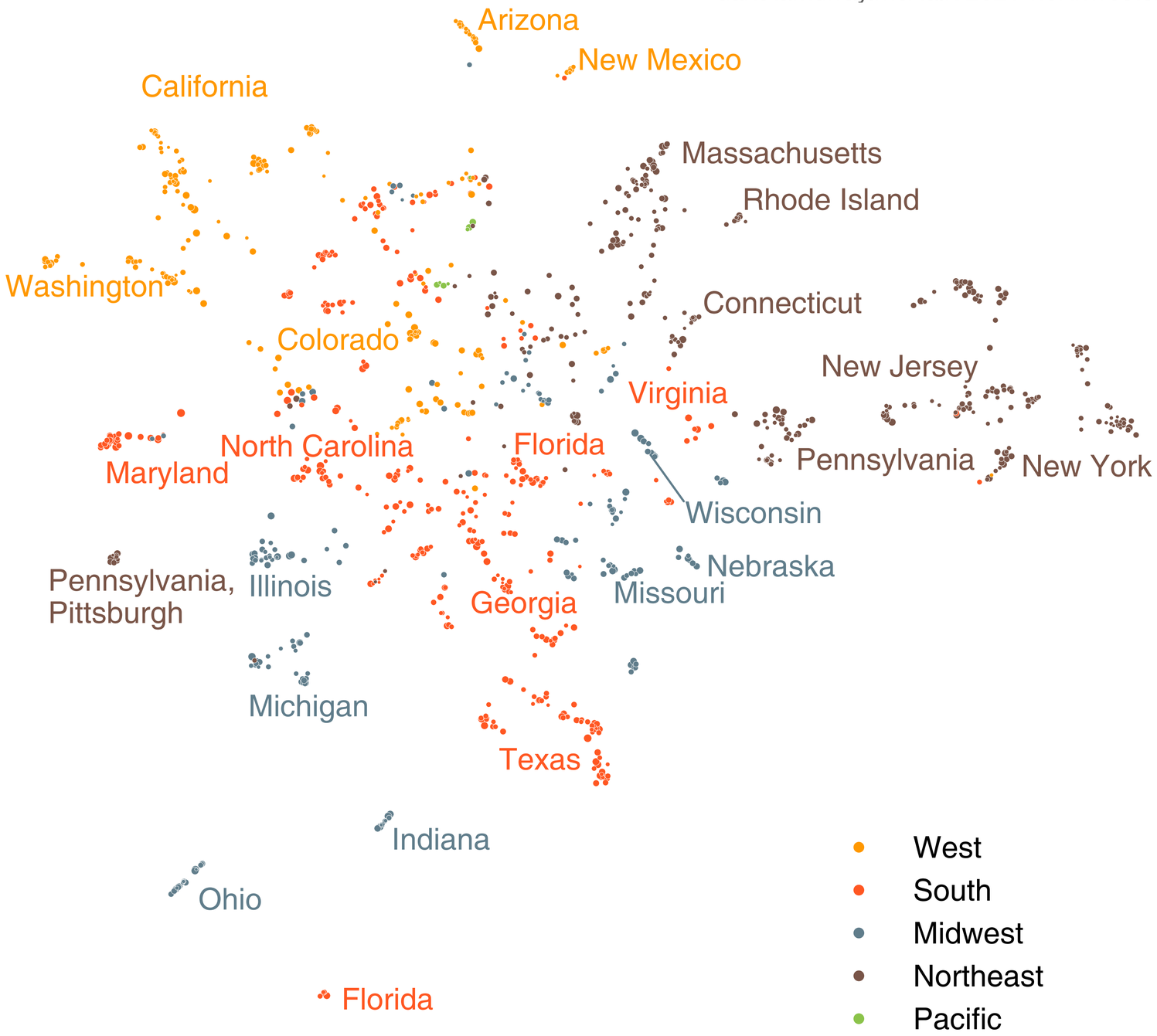
Visualizing the embedding space

We can “zoom in”



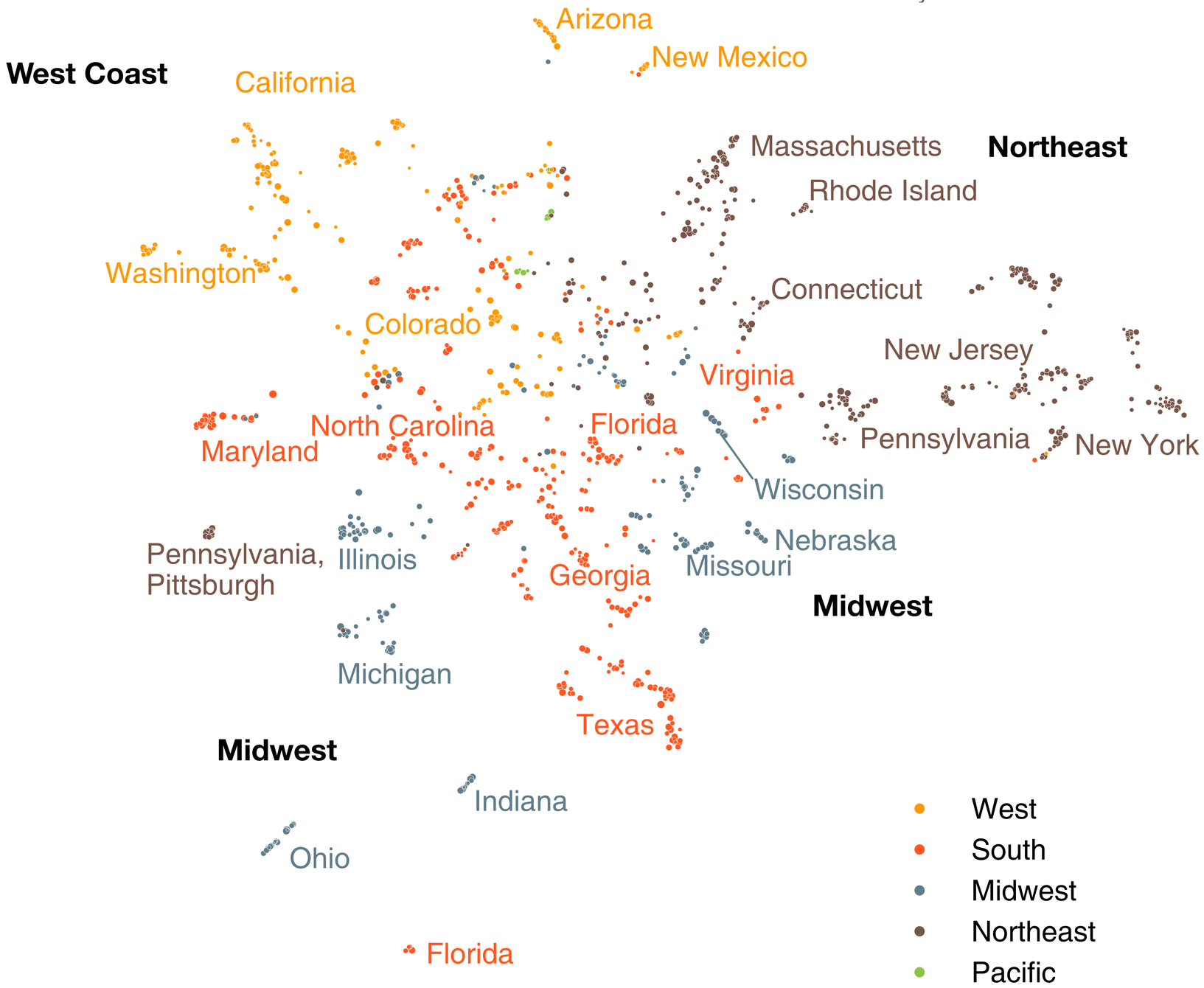
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

United States



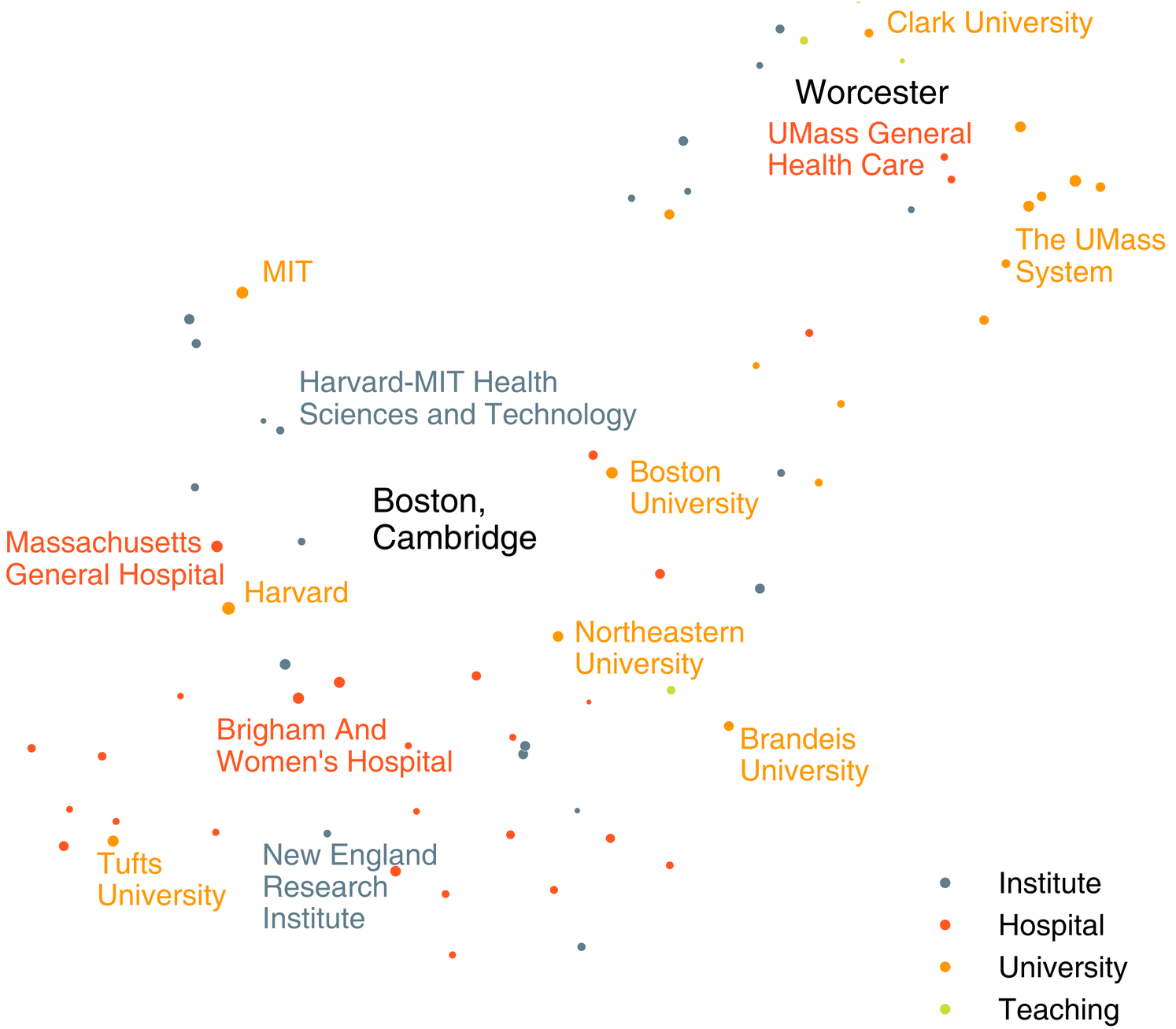
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

United States



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

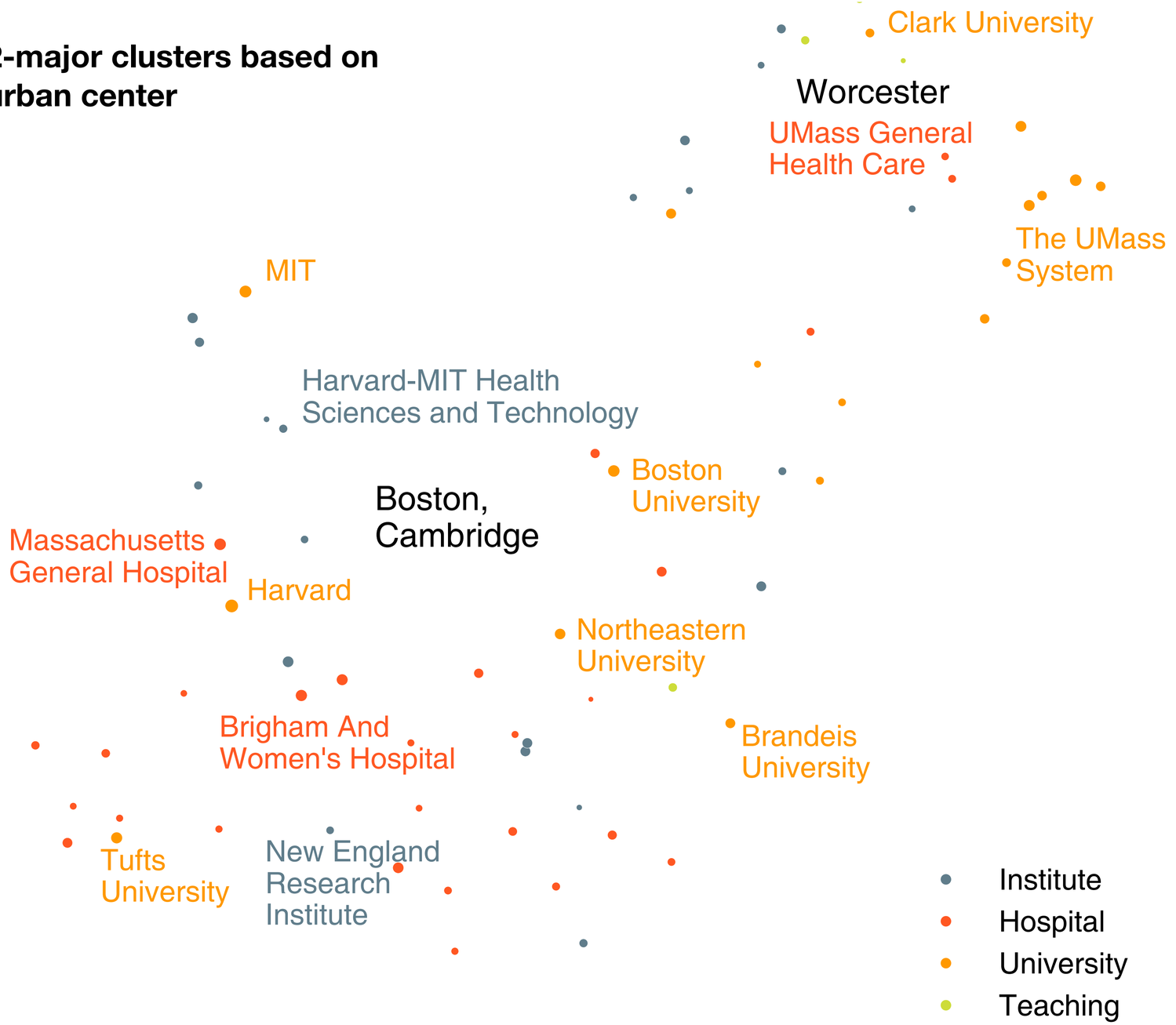
Massachusetts



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Massachusetts

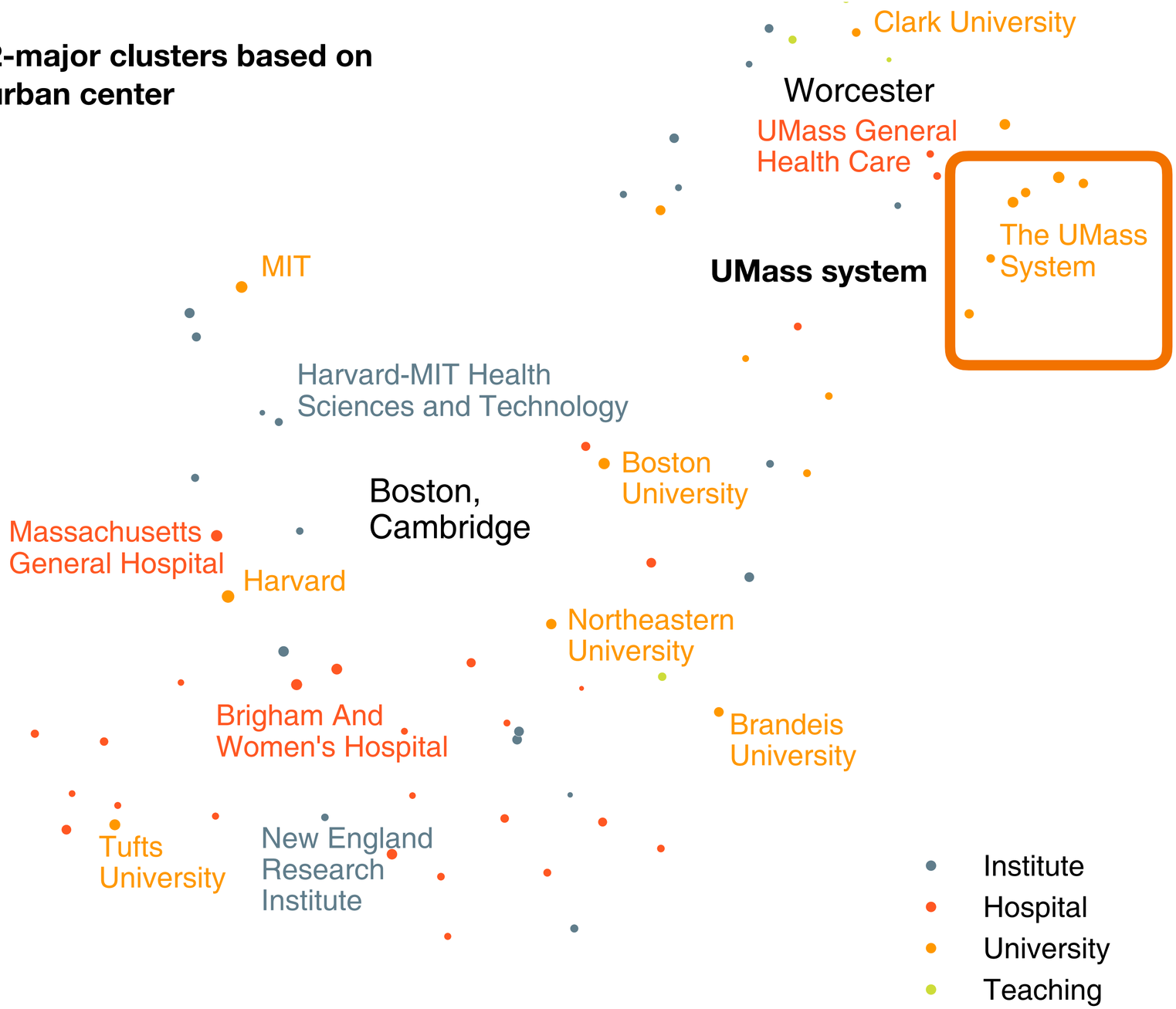
2-major clusters based on urban center



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

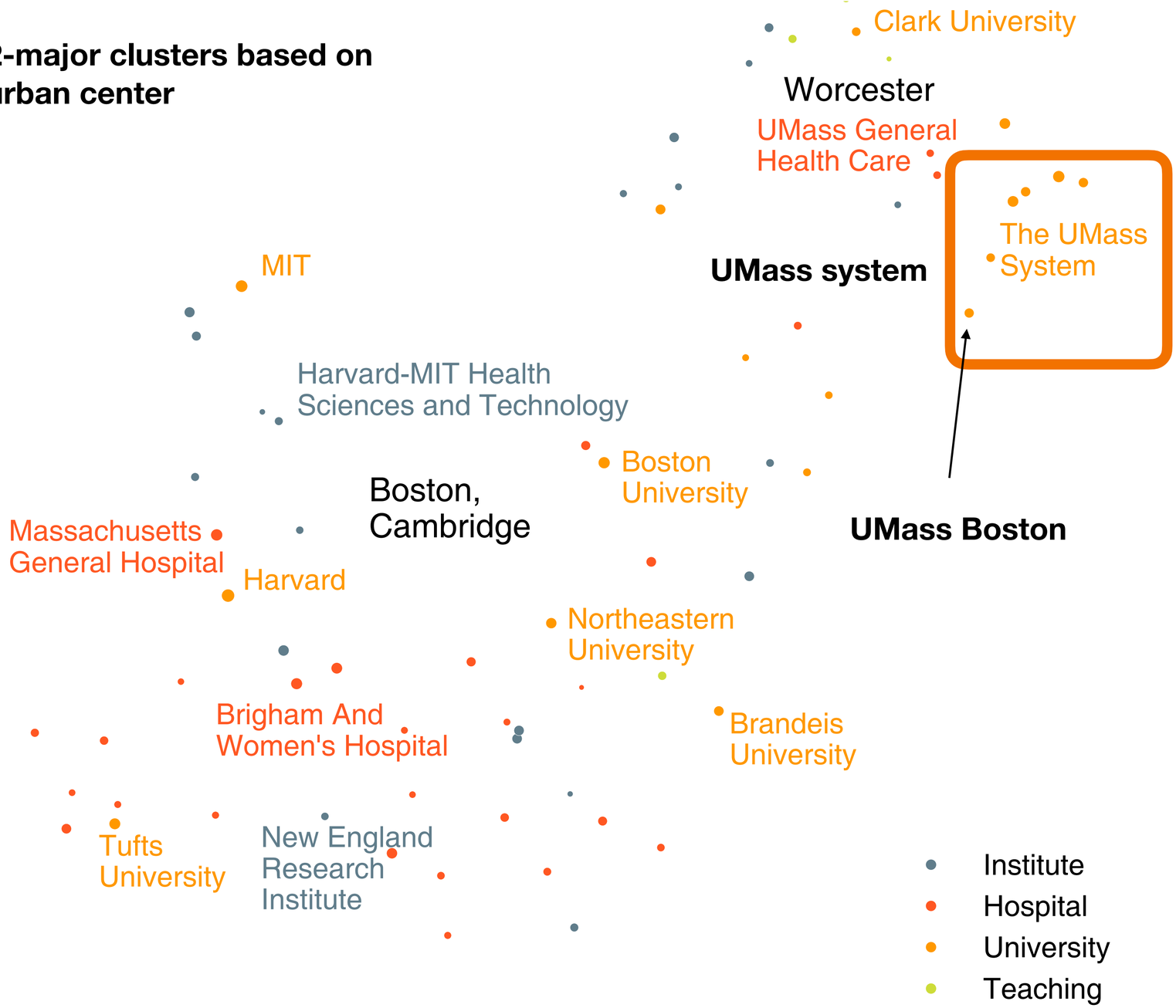
Massachusetts

2-major clusters based on urban center



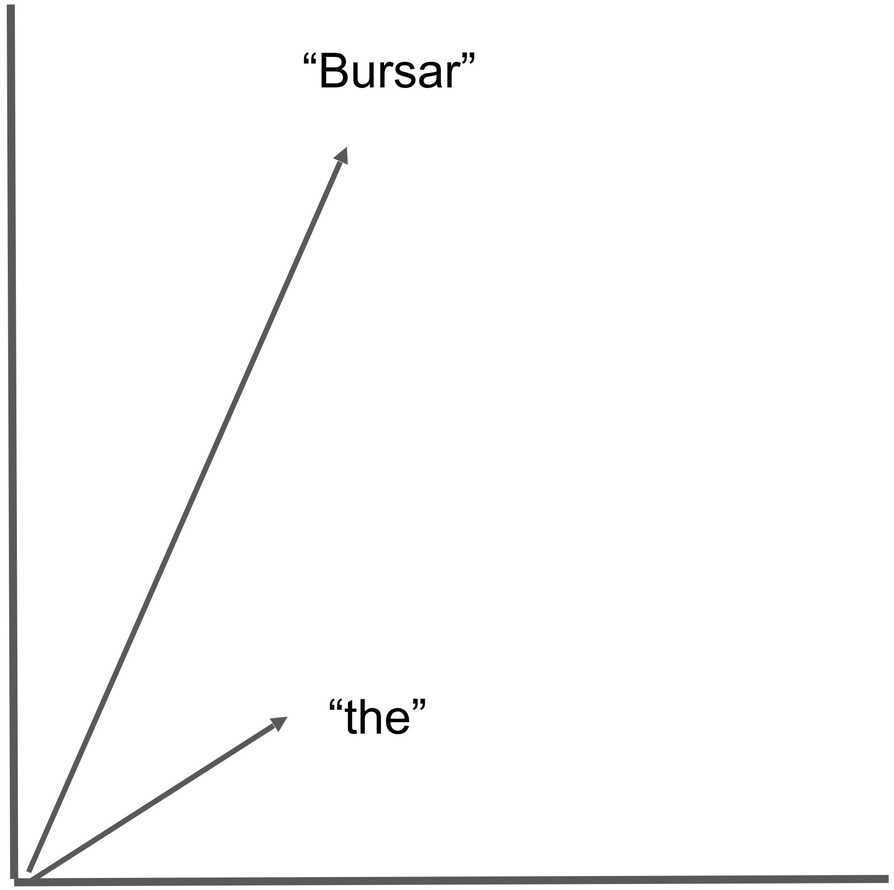
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Massachusetts



Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

What else is encoded? Vector length

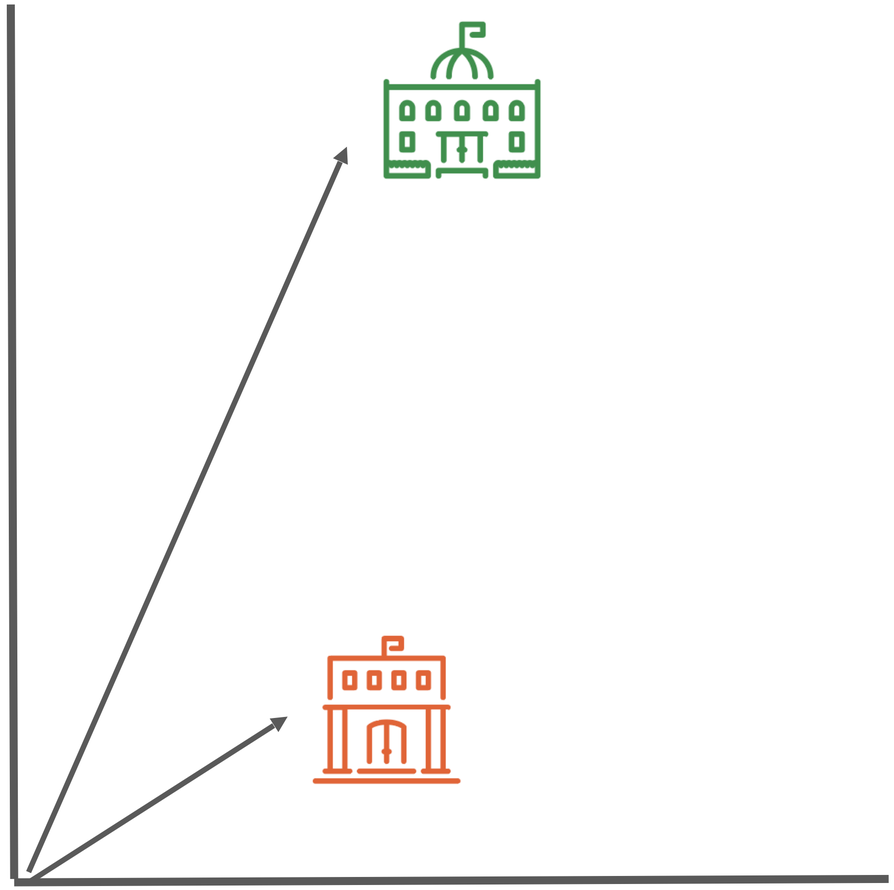


- In word embeddings, larger vectors (by magnitude) tend to appear in a single context

- Shorter vectors tend to appear in more and more different contexts—they are more universal. More *central*

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

What else is encoded? Vector length



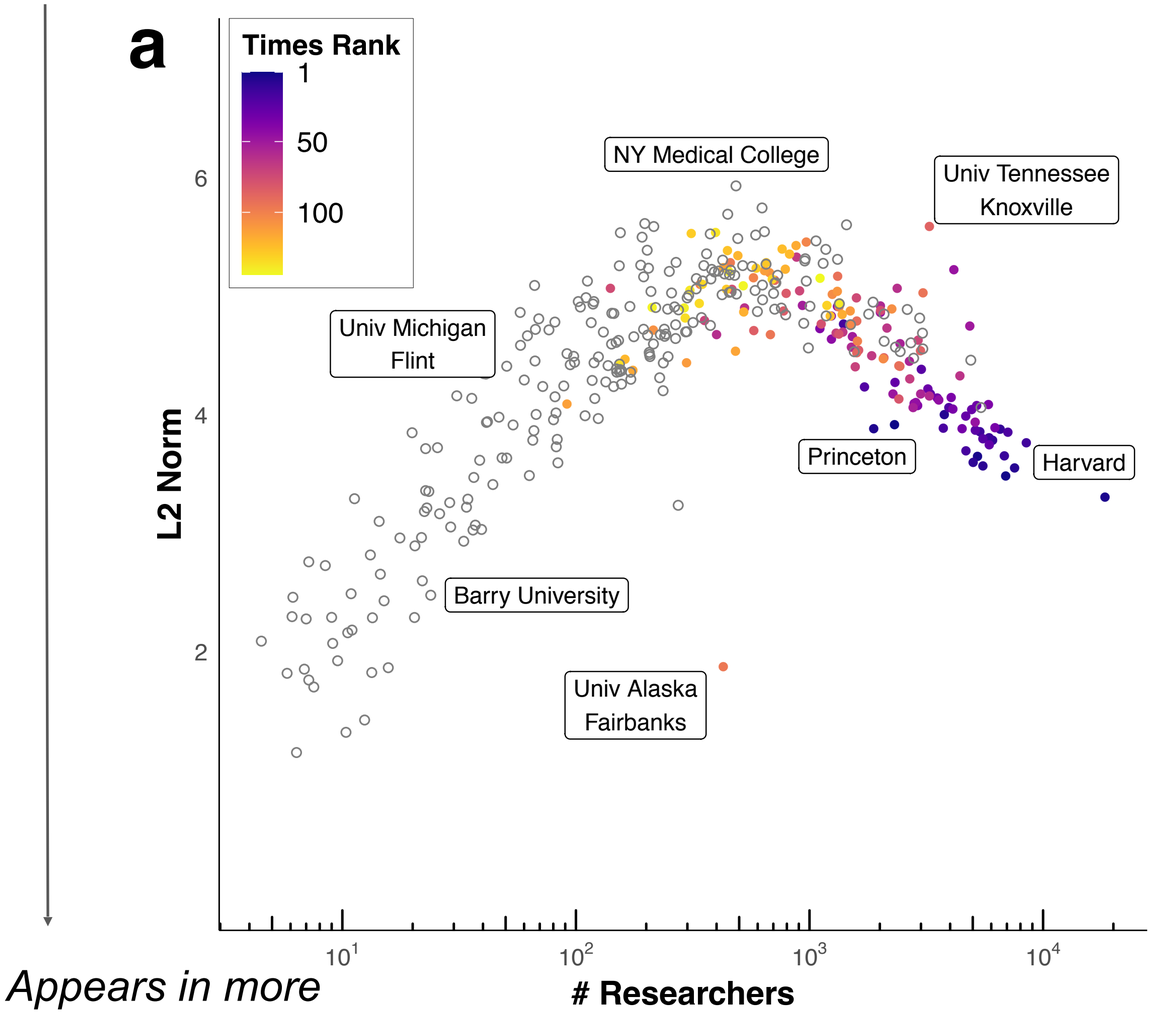
- In word embeddings, larger vectors (by magnitude) tend to appear in a single context

- Shorter vectors tend to appear in more and more different contexts—they are more universal. More *central*

- Also works for organizations

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

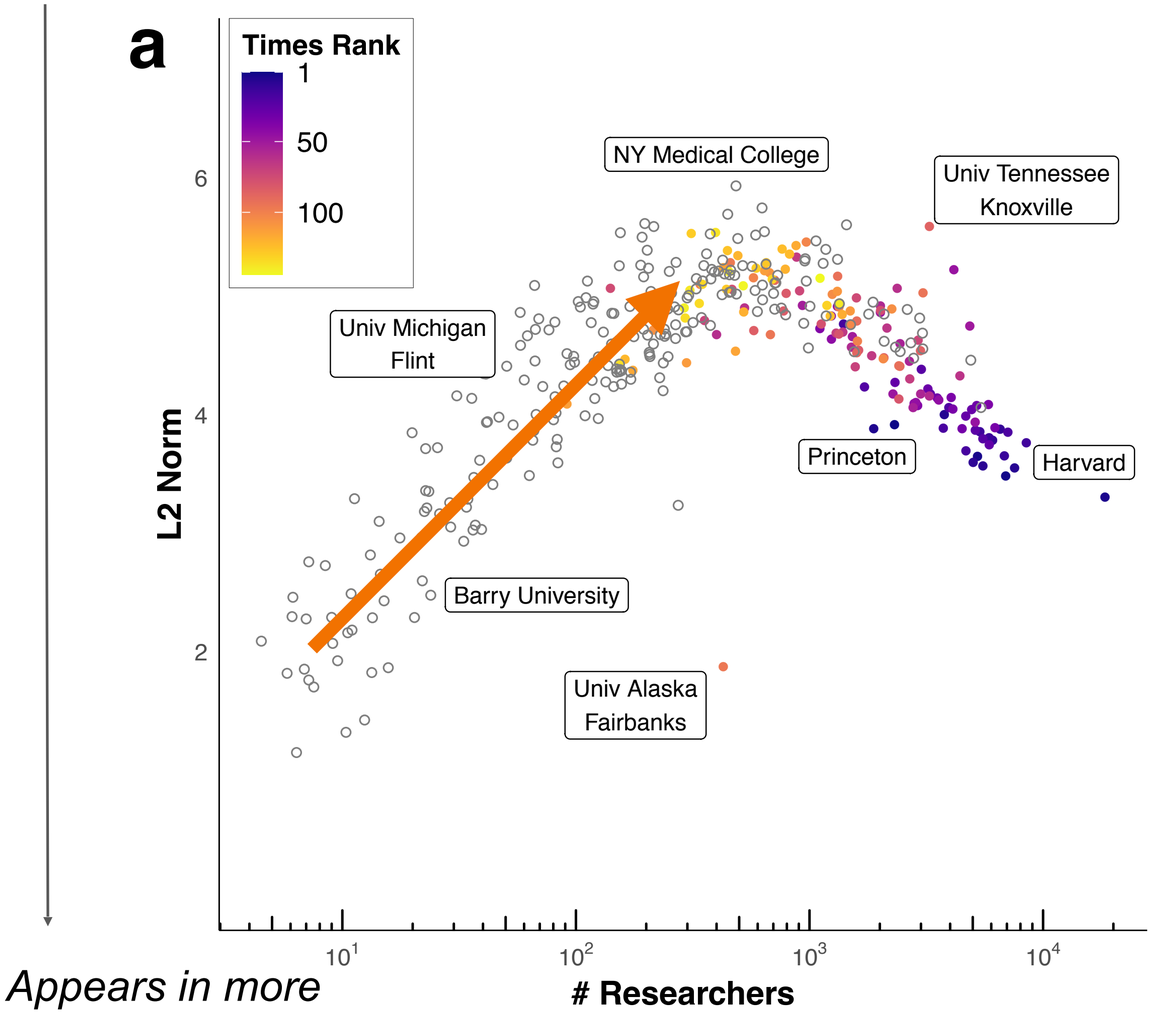
Prestigious U.S. universities appear in more diverse contexts



Appears in more diverse contexts

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Prestigious U.S. universities appear in more diverse contexts

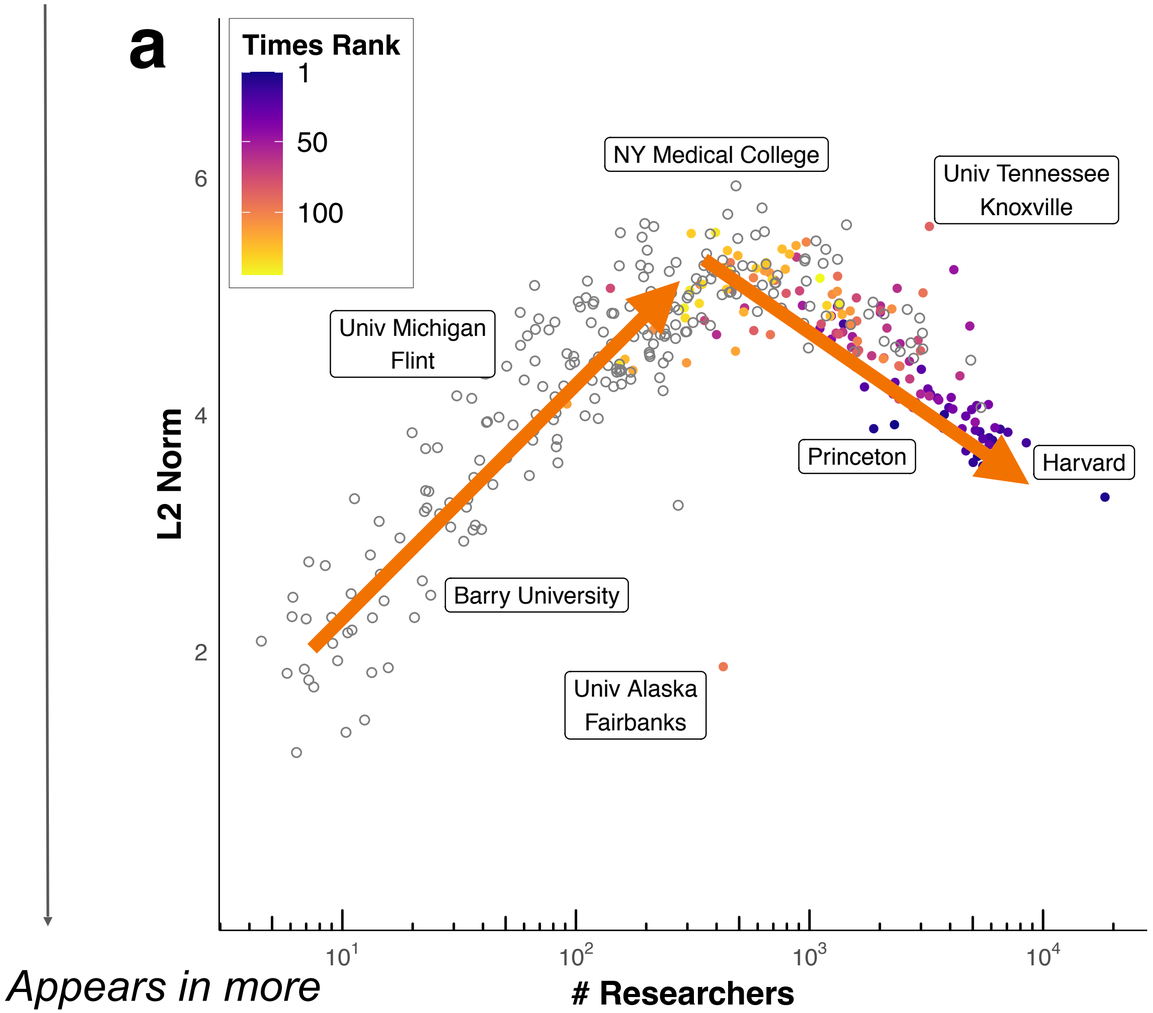


Appears in more diverse contexts

Bigger organizations are more isolated….unless they are prestigious

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Prestigious U.S. universities appear in more diverse contexts



Appears in more diverse contexts

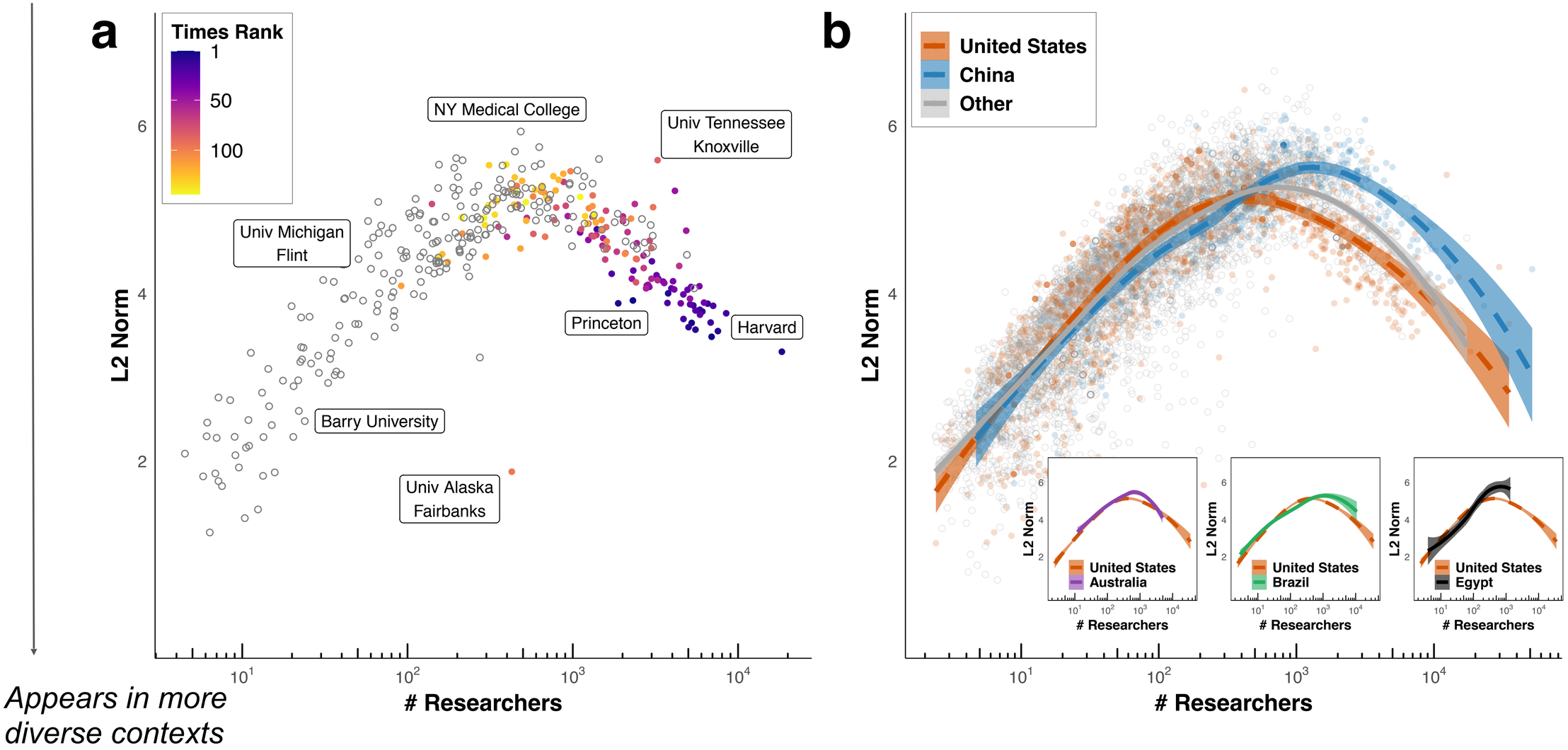
Bigger organizations are more isolated….unless they are prestigious

Those from prestigious universities are more central

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

Prestigious U.S. universities appear in more diverse contexts

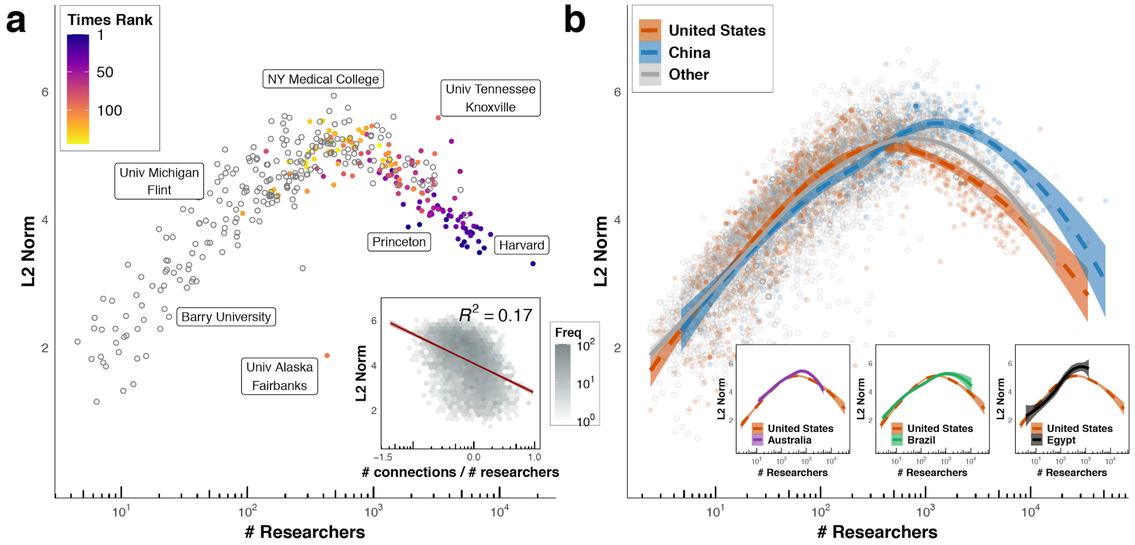
Repeats across many countries

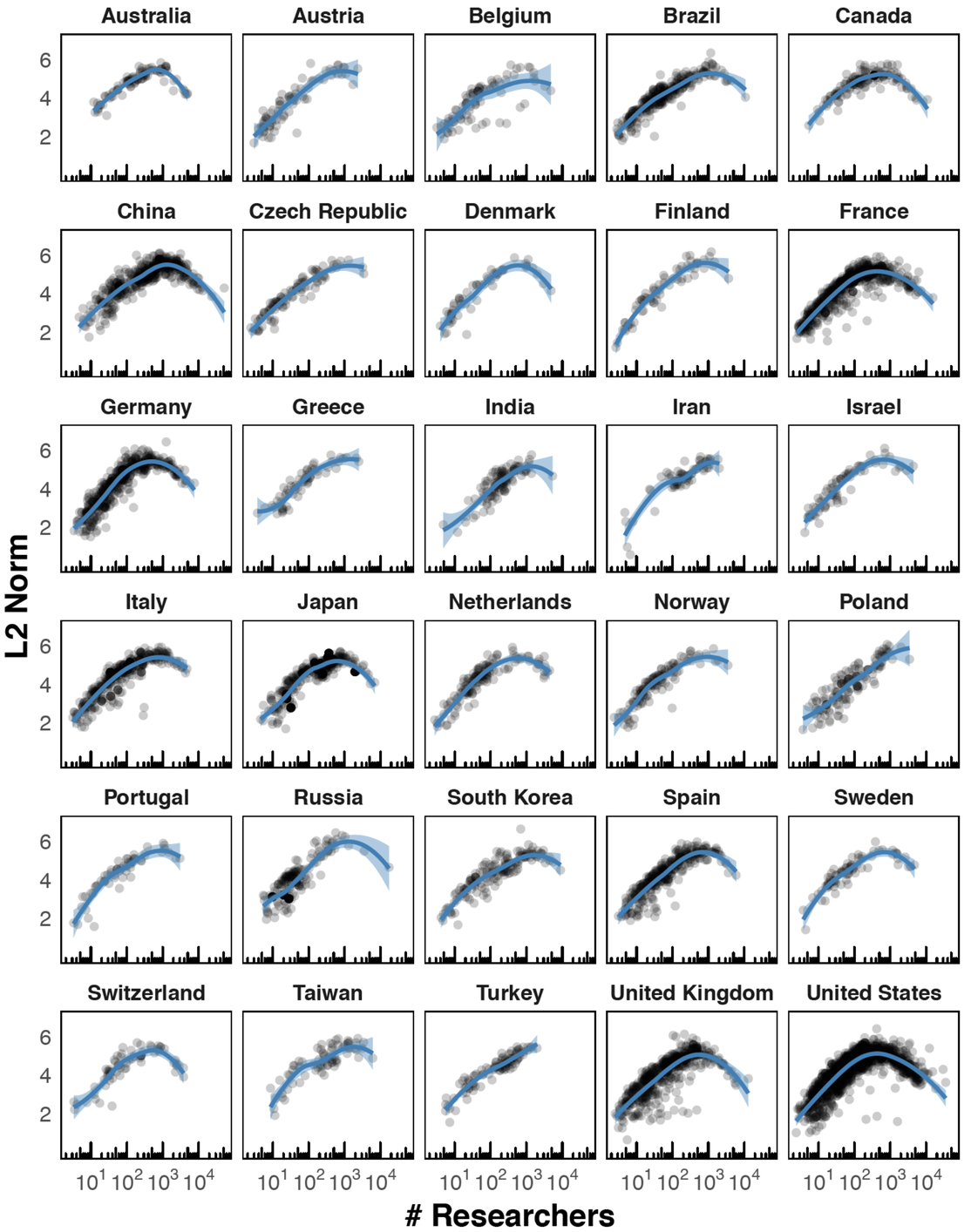


Appears in more diverse contexts

Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

The universal boomerang





Mobility occurs in a complex global context

Mobility occurs in a complex global context

Geography, language, culture, and prestige structure mobility

Mobility occurs in a complex global context

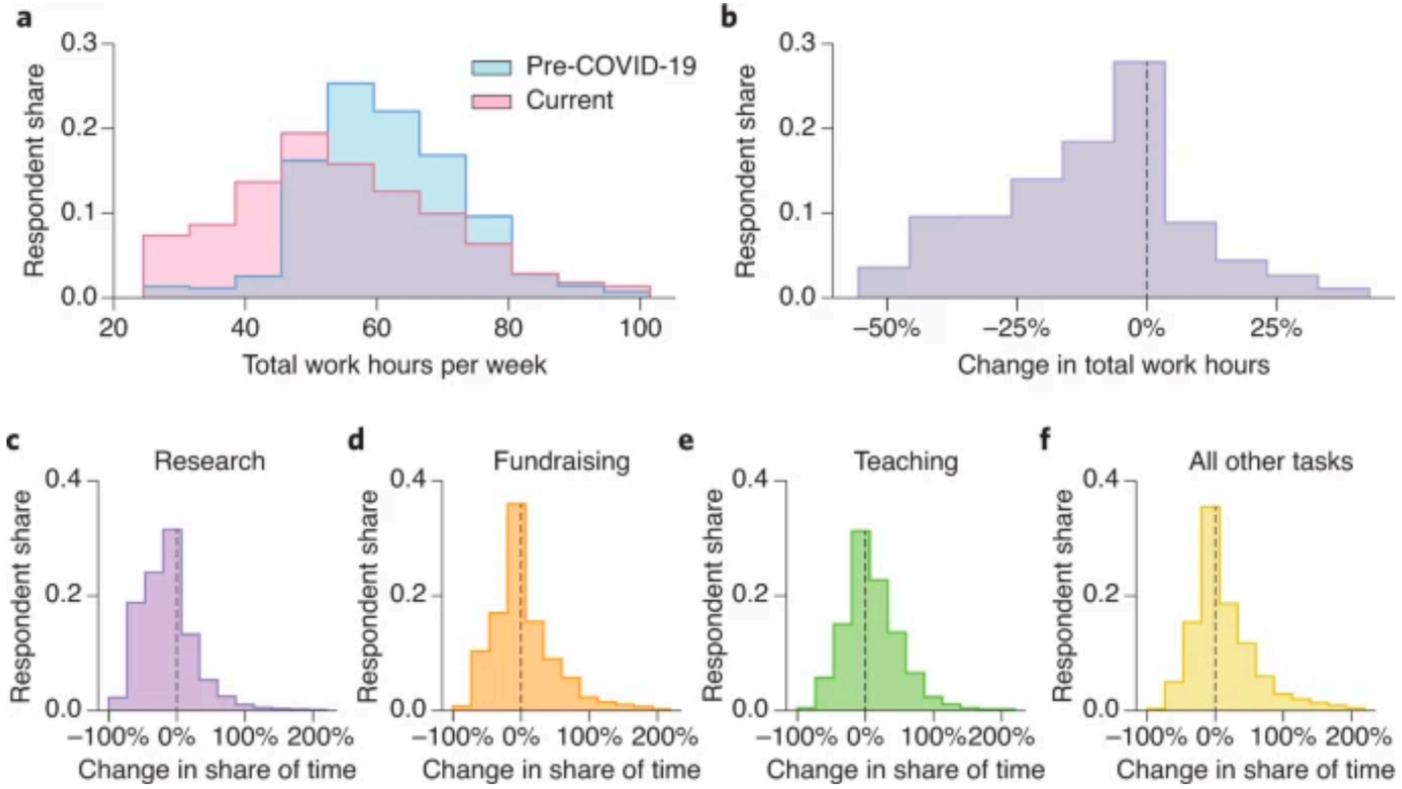
Geography, language, culture, and prestige structure mobility

And they structure success

Next steps: The effect of COVID

It has impacted scientific work, but what does it mean for mobility?

Fig. 1: Changes in levels and allocations of work time.



Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., Schiffer, P., Walsh, J. T., Lakhani, K. R., & Wang, D. (2020). Unequal effects of the COVID-19 pandemic on scientists. *Nature Human Behaviour*, *4*(9), 880–883.

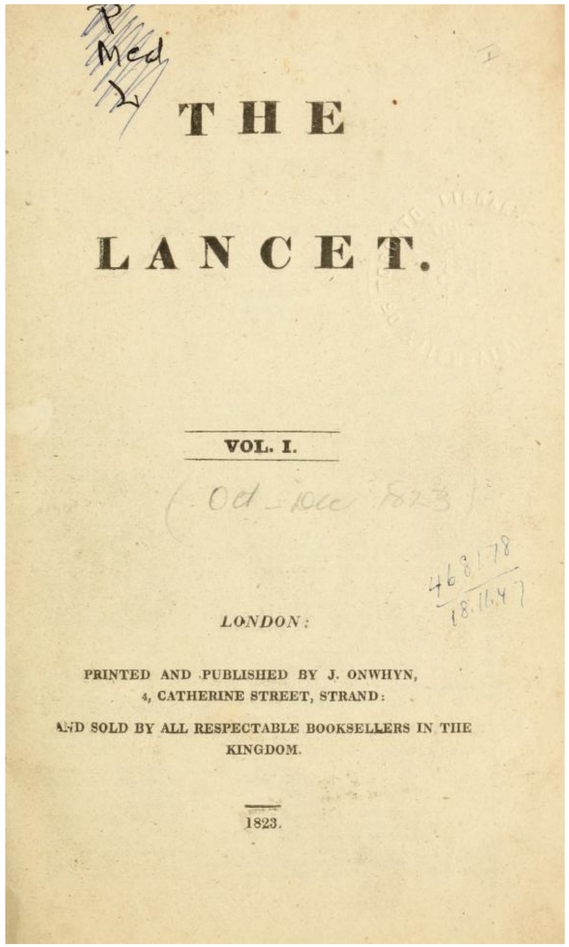
Jumping further ahead in a scientist's career…

Death

And how we are remembered

The final evaluation: the obituary

A window into how a scientific community recognizes its members



THE

LANCET.

VOL. I.

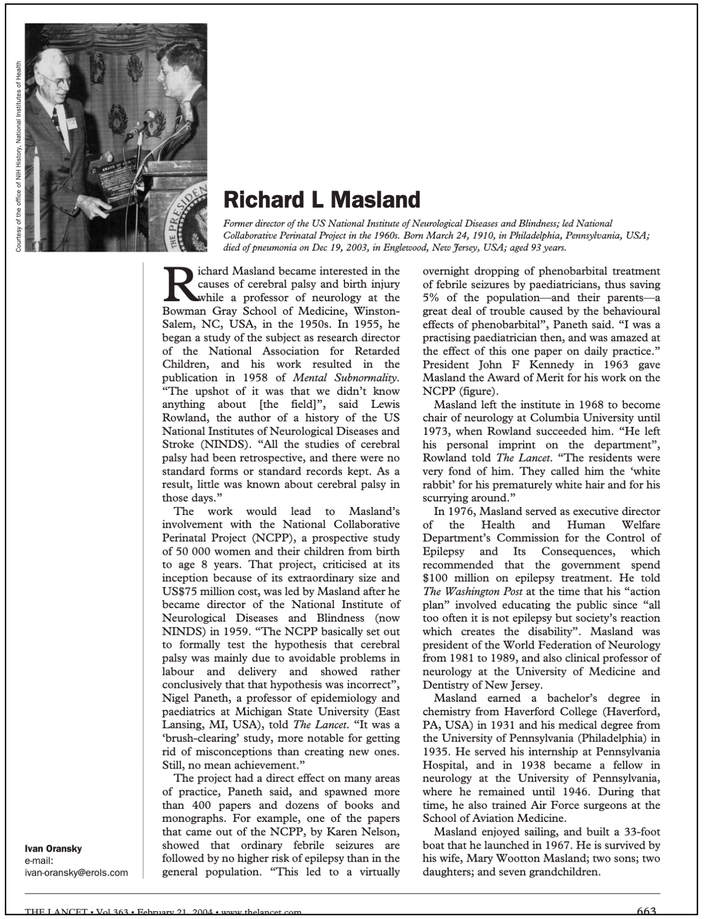
LONDON:

PRINTED AND PUBLISHED BY J. ONWHYN,
4, CATHERINE STREET, STRAND:

AND SOLD BY ALL RESPECTABLE BOOKSELLERS IN THE KINGDOM.

1823.

Extracted over 8,000 obituaries published in the Lancet between 1823 and 2019



Courtesy of the Office of NIH History, National Institutes of Health

Richard L Masland

Former director of the US National Institute of Neurological Diseases and Blindness; led National Collaborative Perinatal Project in the 1960s. Born March 24, 1910, in Philadelphia, Pennsylvania, USA; died of pneumonia on Dec 19, 2003, in Englewood, New Jersey, USA; aged 93 years.

Richard Masland became interested in the causes of cerebral palsy and birth injury while a professor of neurology at the Bowman Gray School of Medicine, Winston-Salem, NC, USA, in the 1950s. In 1955, he began a study of the subject as research director of the National Association for Retarded Children, and his work resulted in the publication in 1958 of *Mental Subnormality*. "The upshot of it was that we didn't know anything about [the field]", said Lewis Rowland, the author of a history of the US National Institutes of Neurological Diseases and Stroke (NINDS). "All the studies of cerebral palsy had been retrospective, and there were no standard forms or standard records kept. As a result, little was known about cerebral palsy in those days."

The work would lead to Masland's involvement with the National Collaborative Perinatal Project (NCPP), a prospective study of 50 000 women and their children from birth to age 8 years. That project, criticised at its inception because of its extraordinary size and US$75 million cost, was led by Masland after he became director of the National Institute of Neurological Diseases and Blindness (now NINDS) in 1959. "The NCPP basically set out to formally test the hypothesis that cerebral palsy was mainly due to avoidable problems in labour and delivery and showed rather conclusively that that hypothesis was incorrect", Nigel Paneth, a professor of epidemiology and paediatrics at Michigan State University (East Lansing, MI, USA), told *The Lancet*. "It was a 'brush-clearing' study, more notable for getting rid of misconceptions than creating new ones. Still, no mean achievement."

The project had a direct effect on many areas of practice, Paneth said, and spawned more than 400 papers and dozens of books and monographs. For example, one of the papers that came out of the NCPP, by Karen Nelson, showed that ordinary febrile seizures are followed by no higher risk of epilepsy than in the general population. "This led to a virtually overnight dropping of phenobarbital treatment of febrile seizures by paediatricians, thus saving 5% of the population—and their parents—a great deal of trouble caused by the behavioural effects of phenobarbital", Paneth said. "I was a practising paediatrician then, and was amazed at the effect of this one paper on daily practice." President John F Kennedy in 1963 gave Masland the Award of Merit for his work on the NCPP (figure).

Masland left the institute in 1968 to become chair of neurology at Columbia University until 1973, when Rowland succeeded him. "He left his personal imprint on the department", Rowland told *The Lancet*. "The residents were very fond of him. They called him the 'white rabbit' for his prematurely white hair and for his scurrying around."

In 1976, Masland served as executive director of the Health and Human Welfare Department's Commission for the Control of Epilepsy and Its Consequences, which recommended that the government spend $100 million on epilepsy treatment. He told *The Washington Post* at the time that his "action plan" involved educating the public since "all too often it is not epilepsy but society's reaction which creates the disability". Masland was president of the World Federation of Neurology from 1981 to 1989, and also clinical professor of neurology at the University of Medicine and Dentistry of New Jersey.

Masland earned a bachelor's degree in chemistry from Haverford College (Haverford, PA, USA) in 1931 and his medical degree from the University of Pennsylvania (Philadelphia) in 1935. He served his internship at Pennsylvania Hospital, and in 1938 became a fellow in neurology at the University of Pennsylvania, where he remained until 1946. During that time, he also trained Air Force surgeons at the School of Aviation Medicine.

Masland enjoyed sailing, and built a 33-foot boat that he launched in 1967. He is survived by his wife, Mary Wootton Masland; two sons; two daughters; and seven grandchildren.

Ivan Oransky
e-mail:
ivan-oransky@erols.com

Chart the history of *The Lancet*

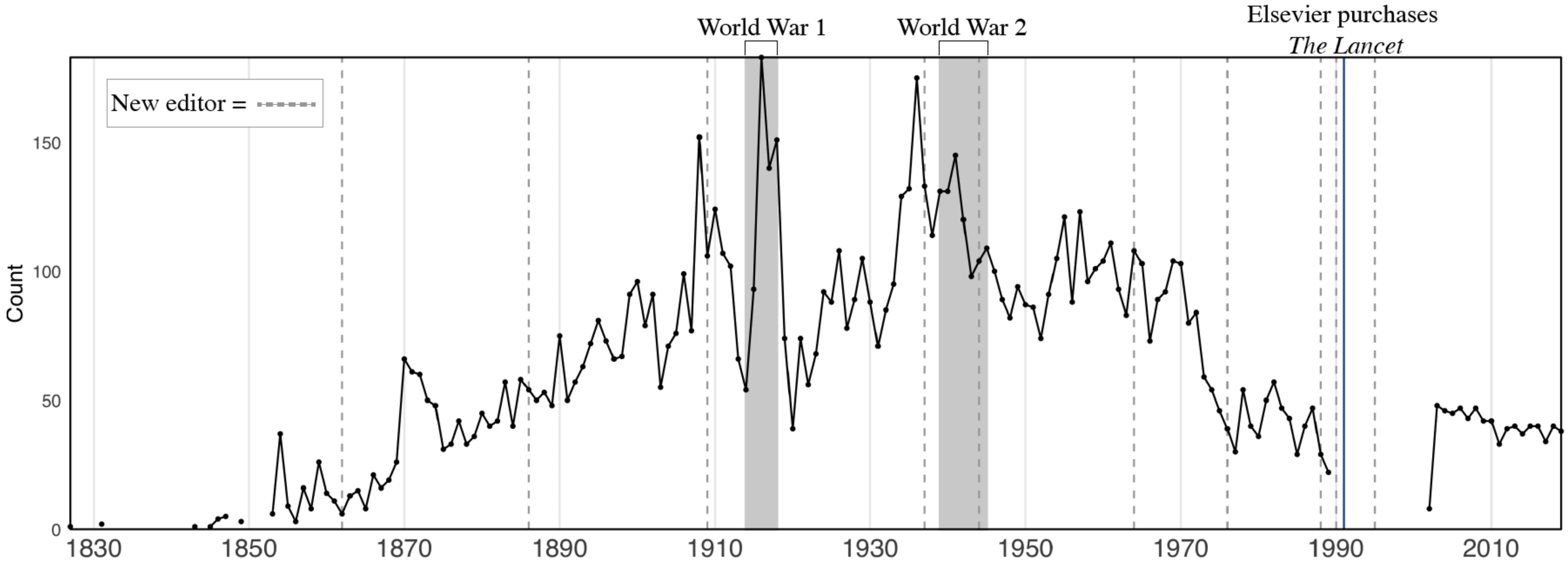


Chart the history of *The Lancet*

The Elsevier purchase, the death of the obituary, and its rebirth

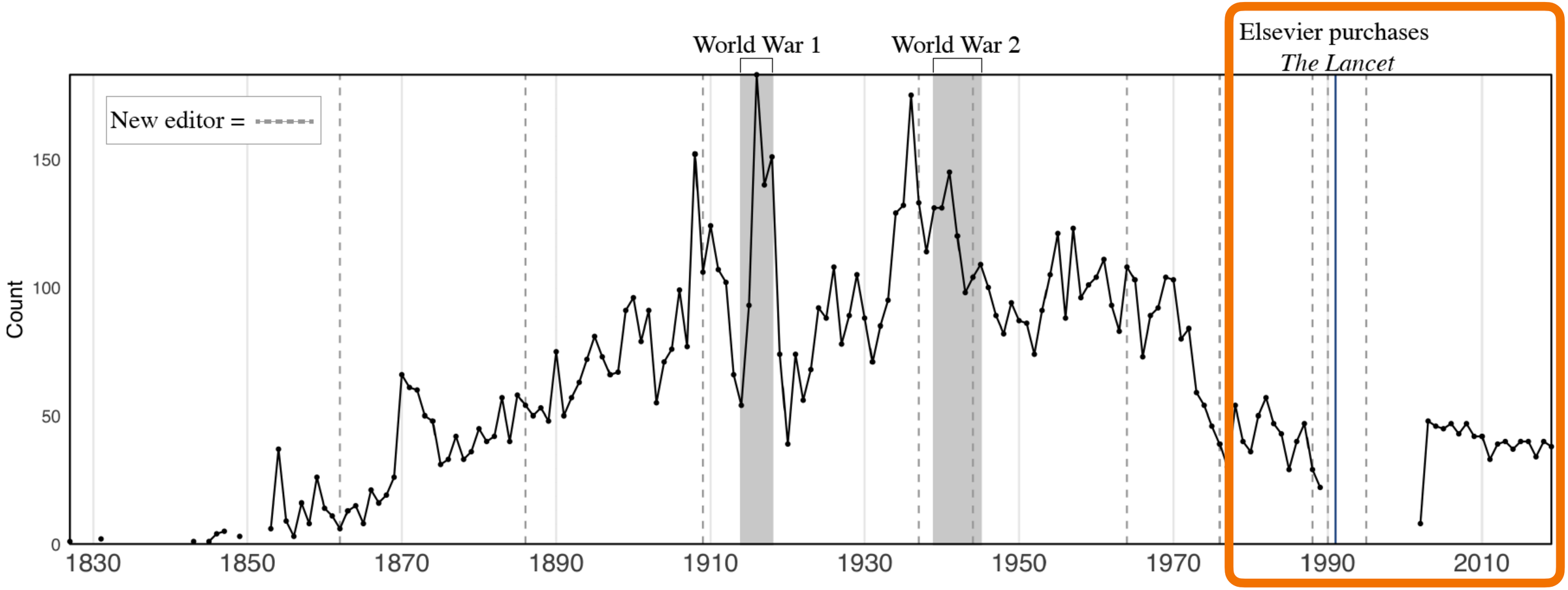


Chart the history of *The Lancet*

The deaths of the first world war

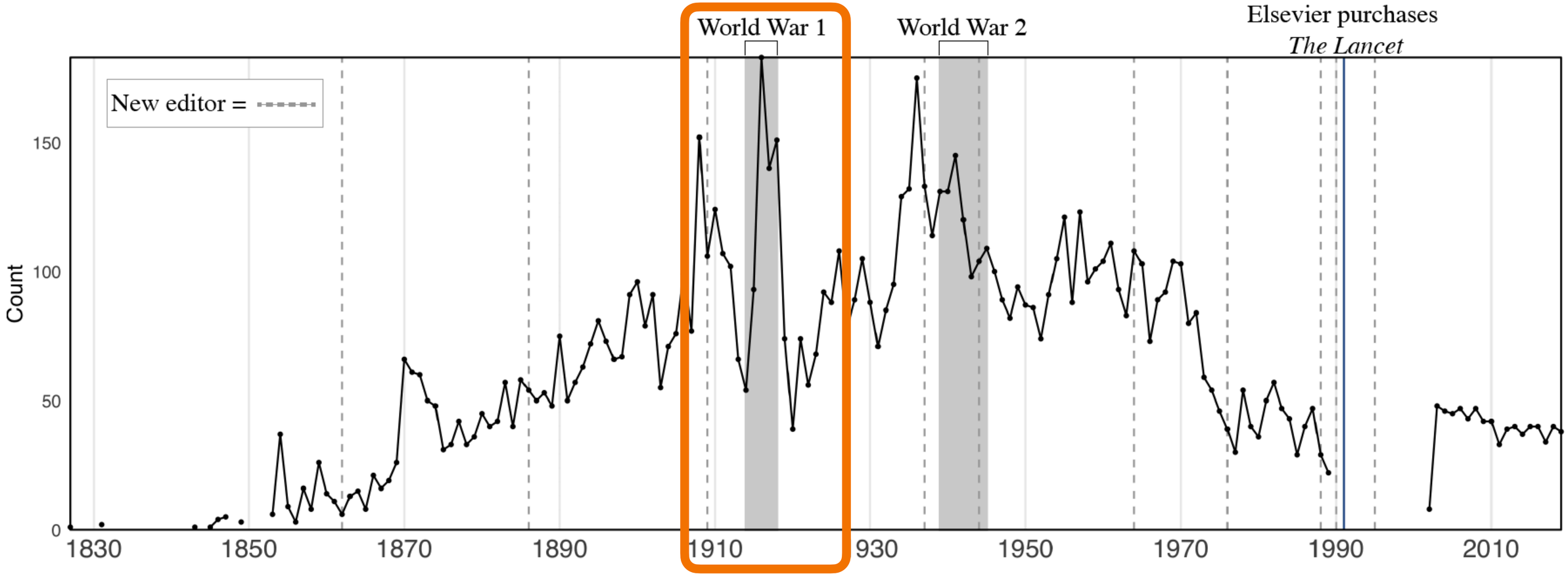
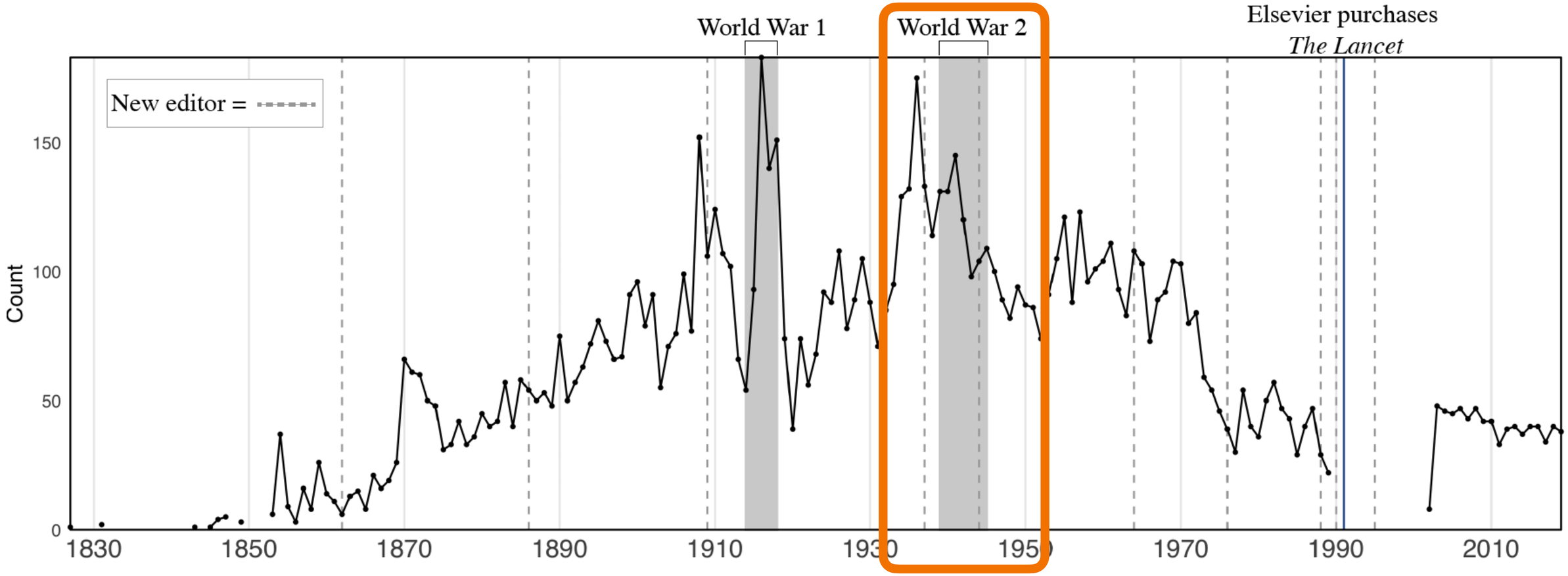


Chart the history of *The Lancet*

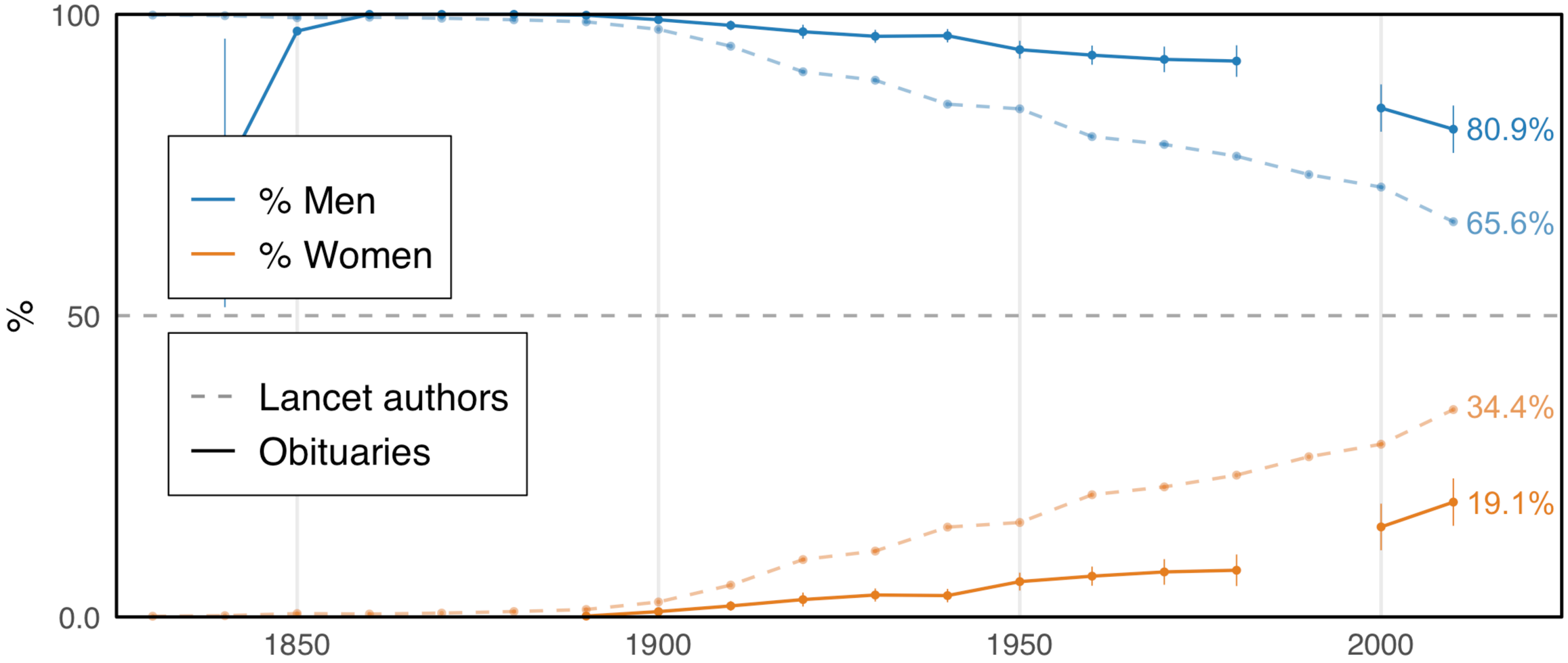
The death of WW1 veterans, and a similar, though smaller WW2 spike



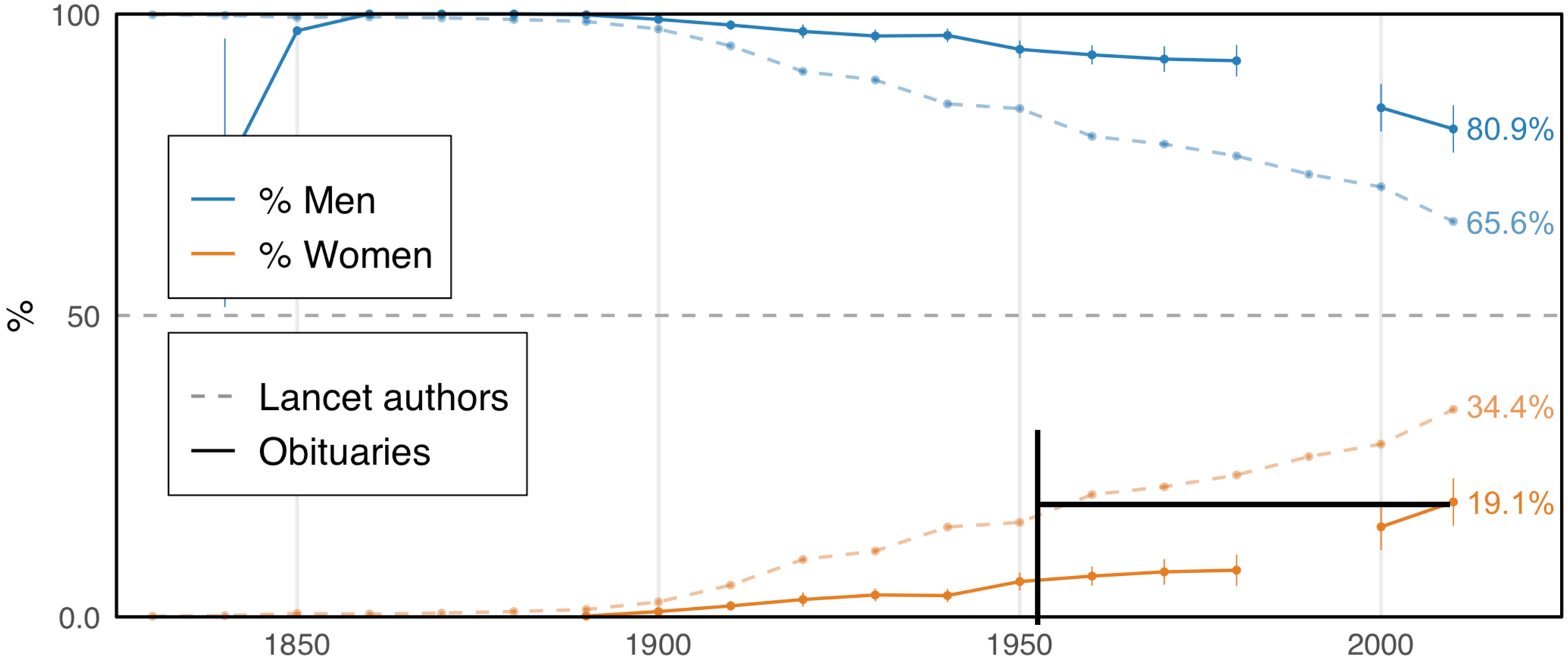
Who is in the medical profession, and who gets an obituary?

Gender representation has improved

Extracted gender using pronouns in text, from author names of all Lancet Papers

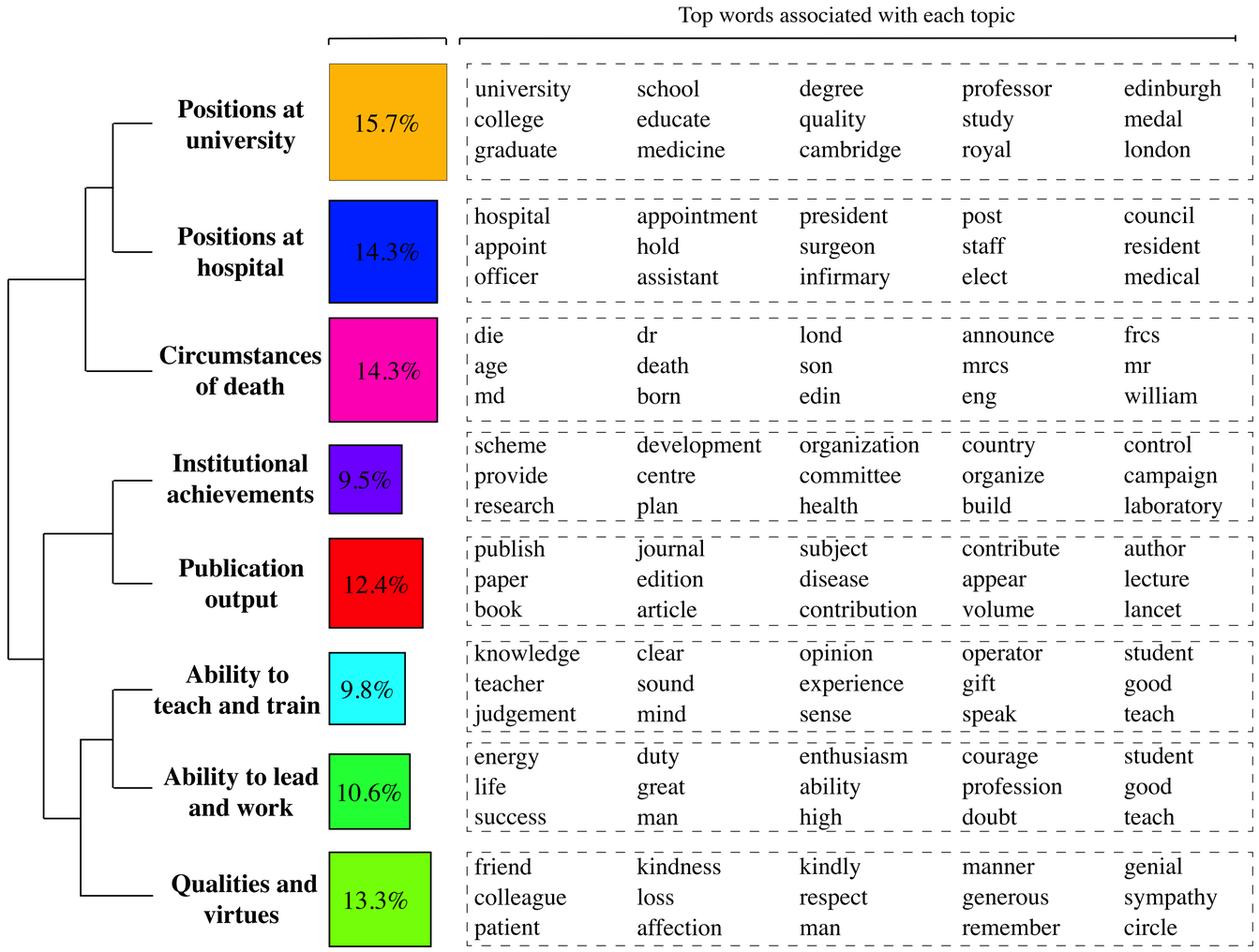


Women's obituaries in 2010 match the % of women's authors in the 1950s

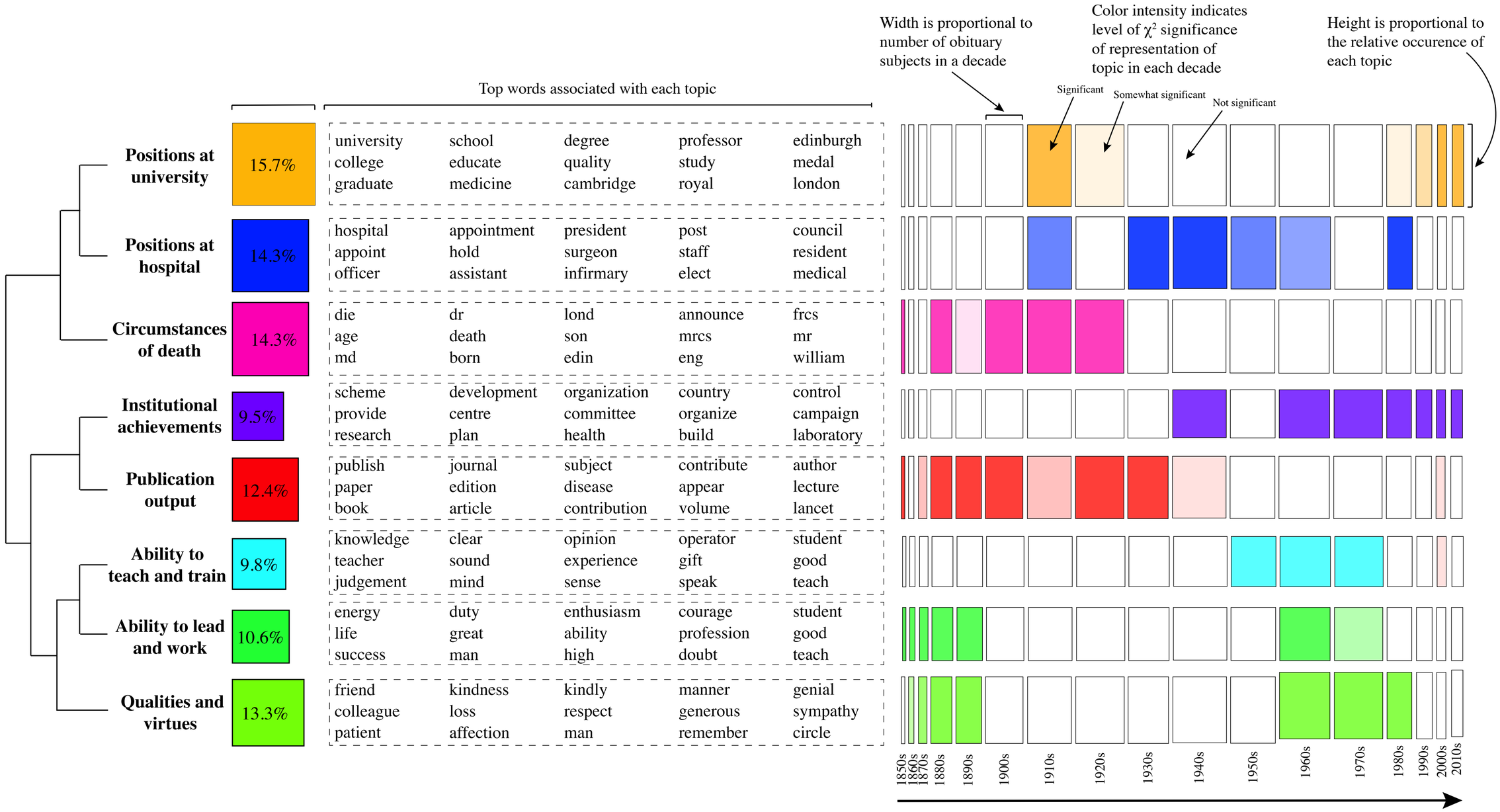


How do we talk about the dead?

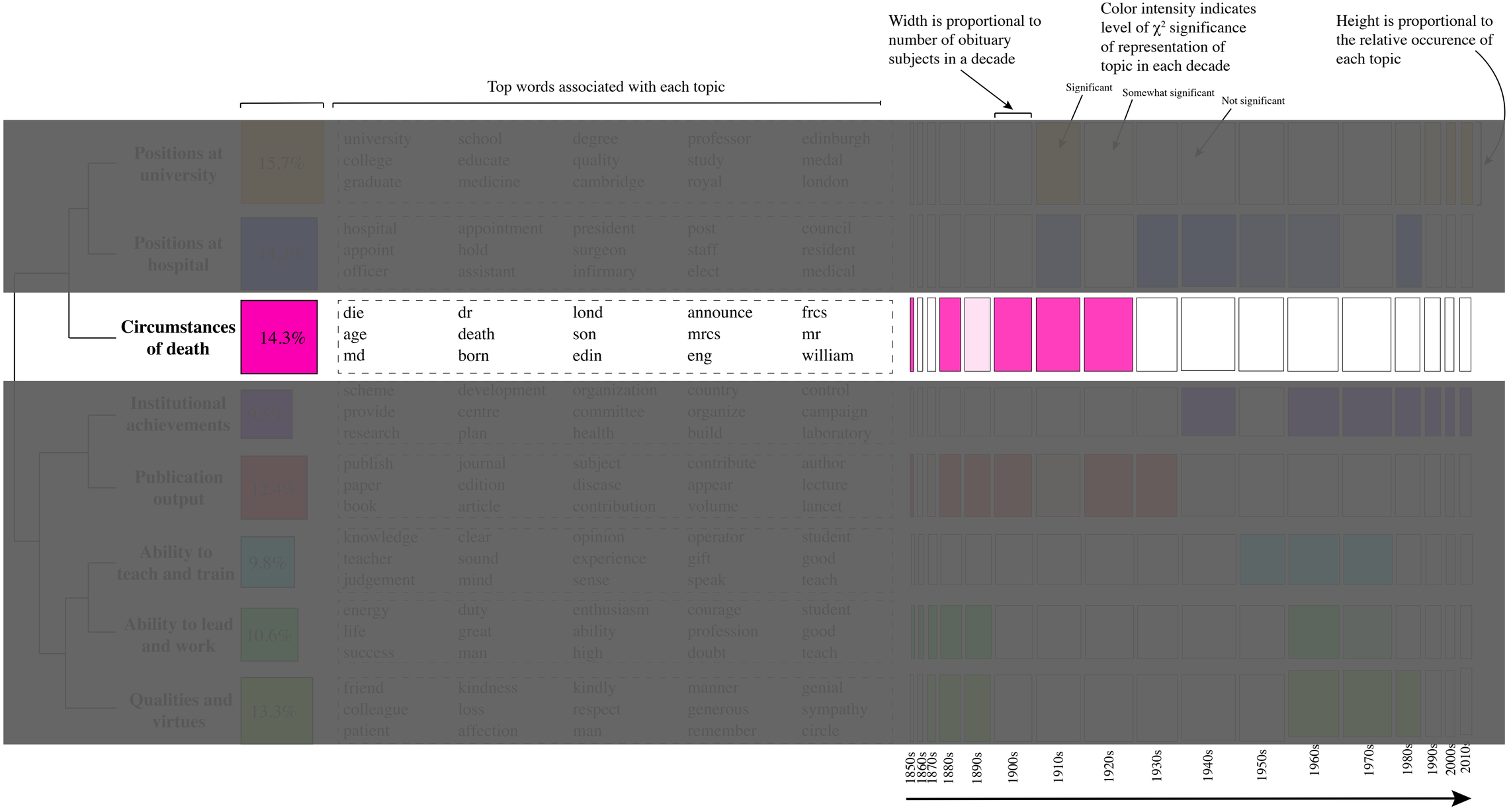
Identified 8 topics



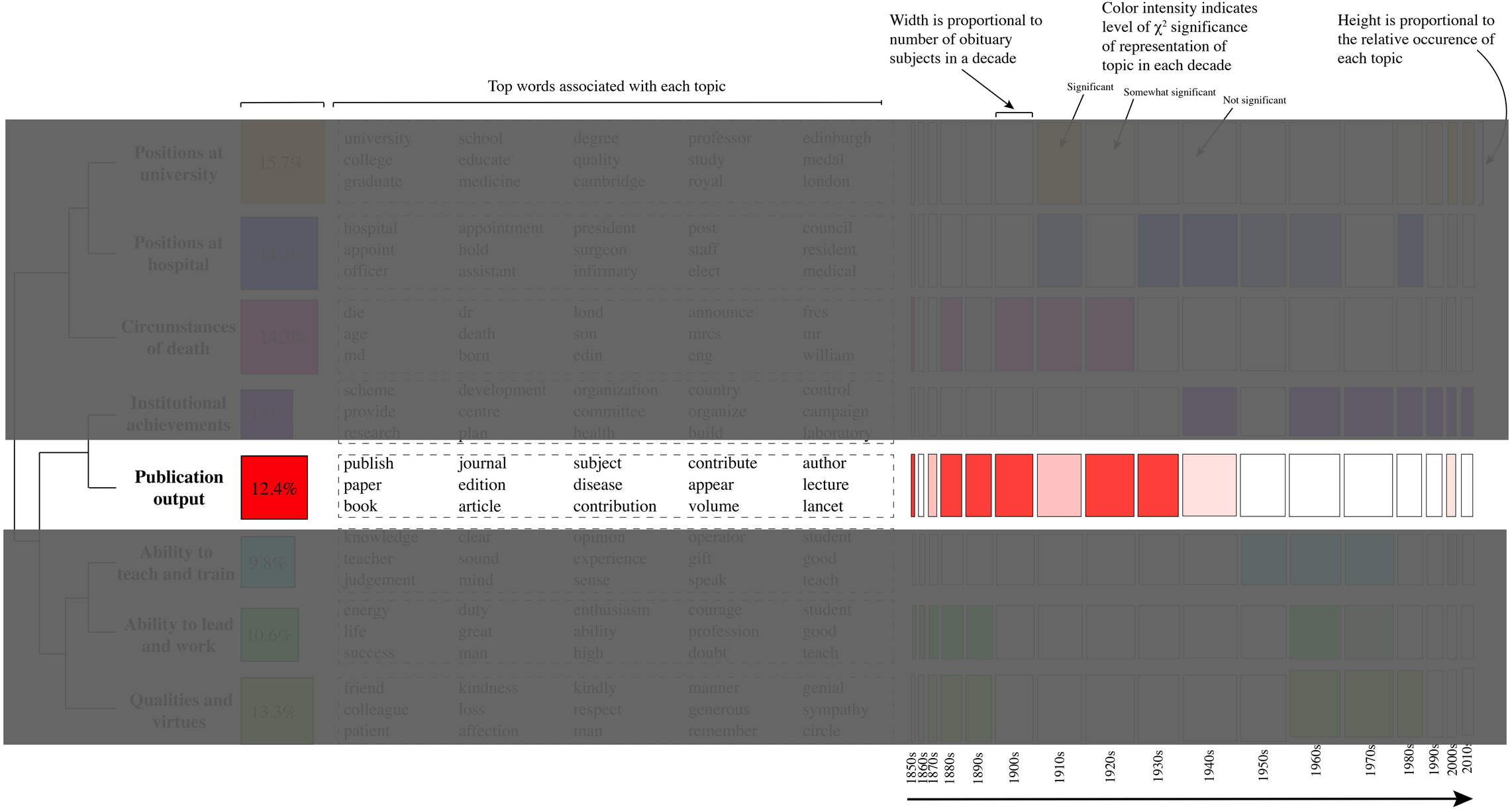
Tracked their prominence over time



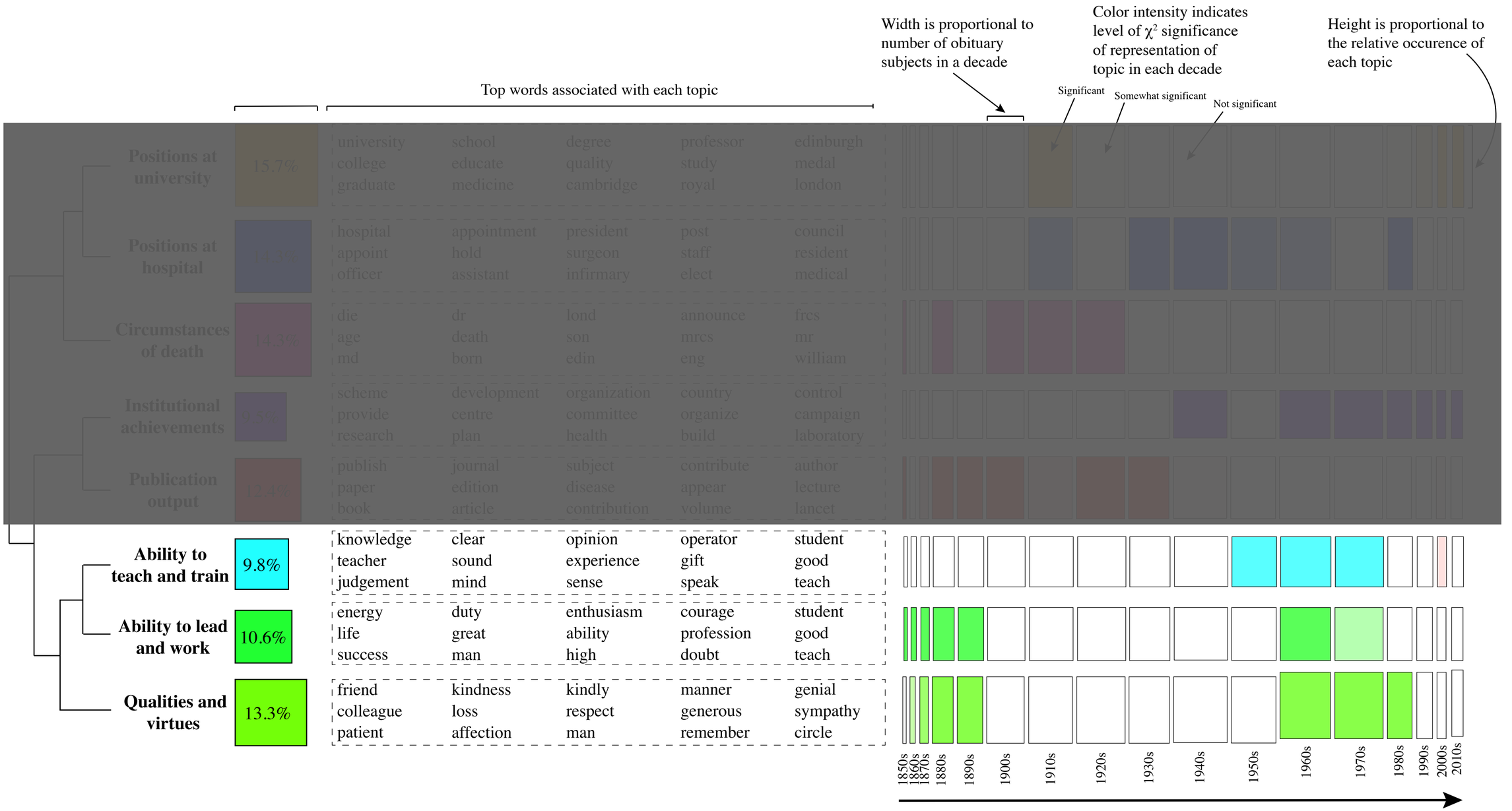
Details of death common until 1920s



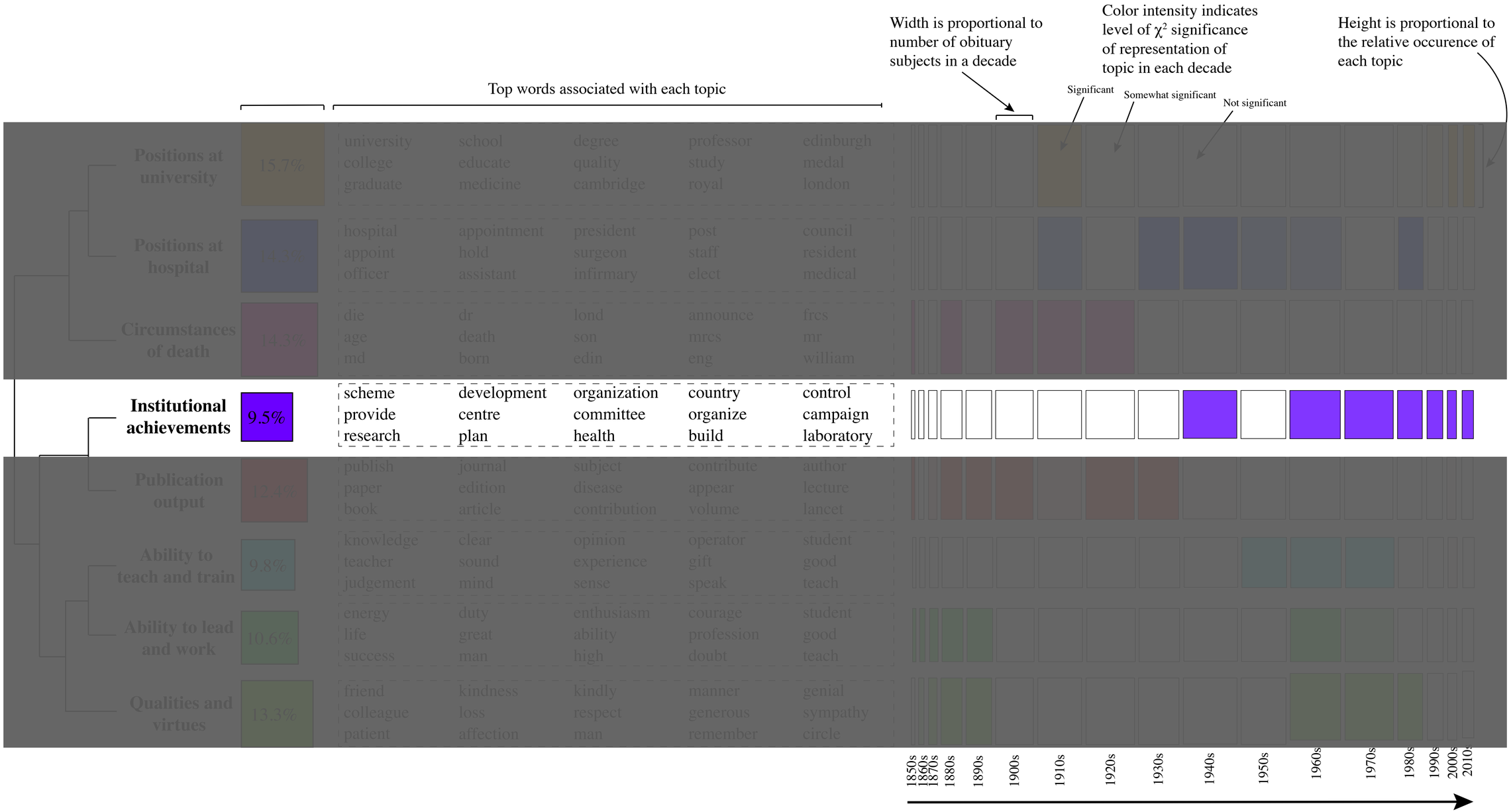
Publications prominent until 1940s



Teaching, leadership, and virtues common from 60s to 80s



Professional achievements prominent in recent years



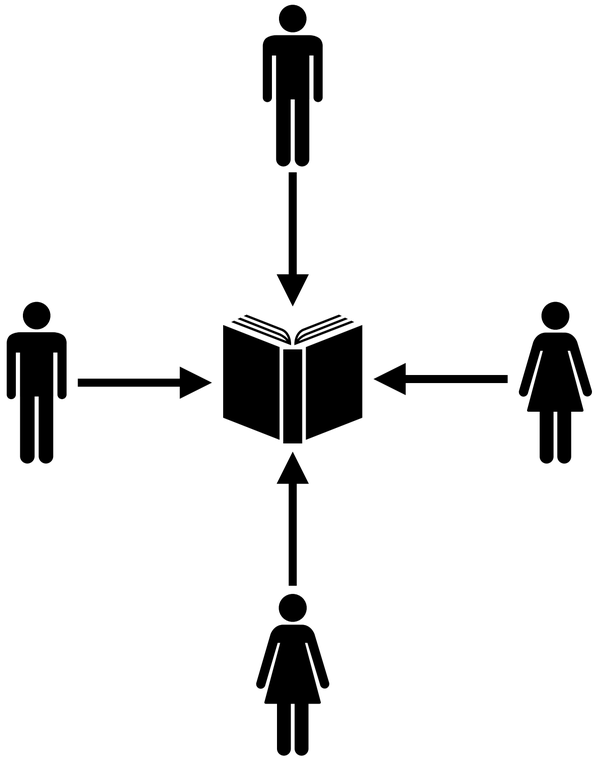
How we evaluate (and honor) the life of a researcher changes over time, in line with history, the field, and culture

Early results, thoughts appreciated!

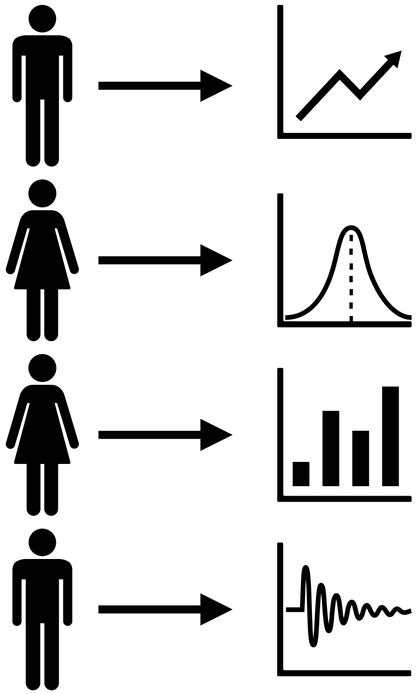
Conclusion

Ideally, evaluation should capture true merit

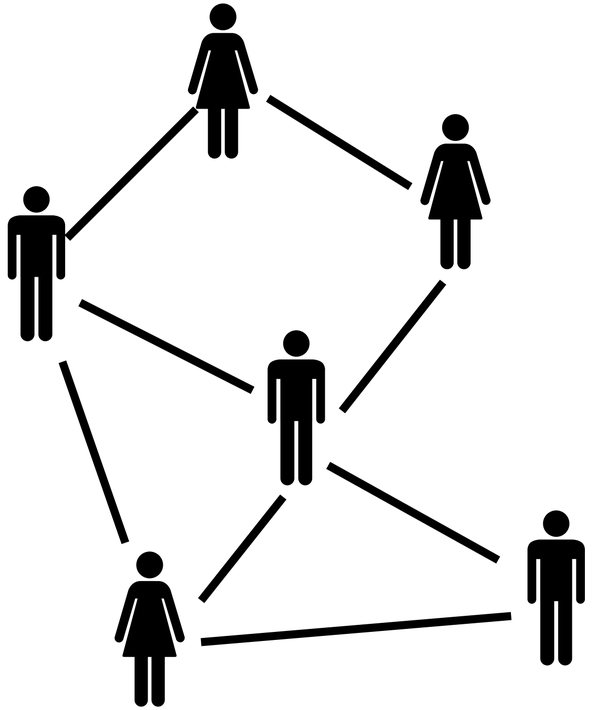
Peer review



Performance metrics



Reputation
(networks)



Evaluation is deeply contextual

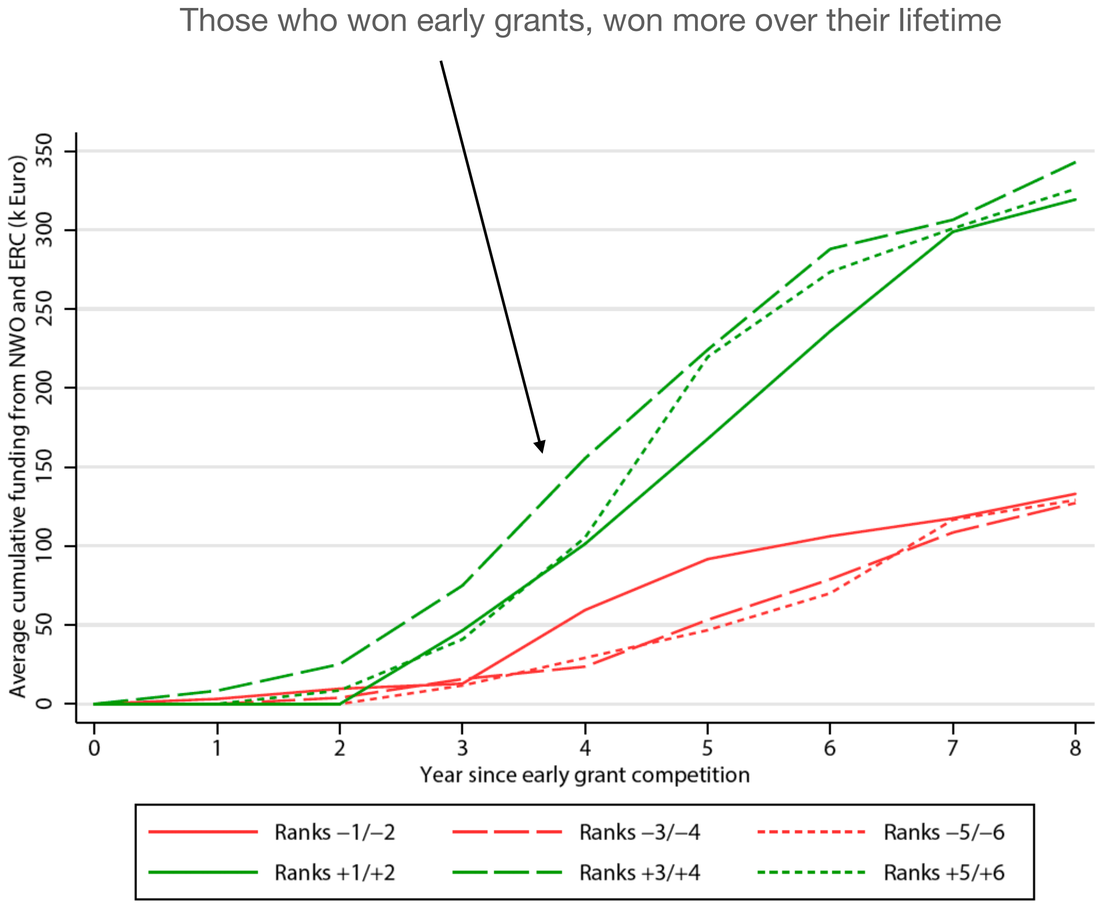
Hinges on the identities, cultures, and relationships of both the evaluators, and the ones being evaluated

Context is often ignored in evaluation

Matthew effects

Small biases, big disparities

- Early winners will continue to win
- Inequality in peer review, teaching ratings, citation impact, and mobility can compound into large disparities
- Drive who stays and who leaves



Barabási, A.-L. (2018). *The Formula: The Universal Laws of Success*. Little, Brown and Company.

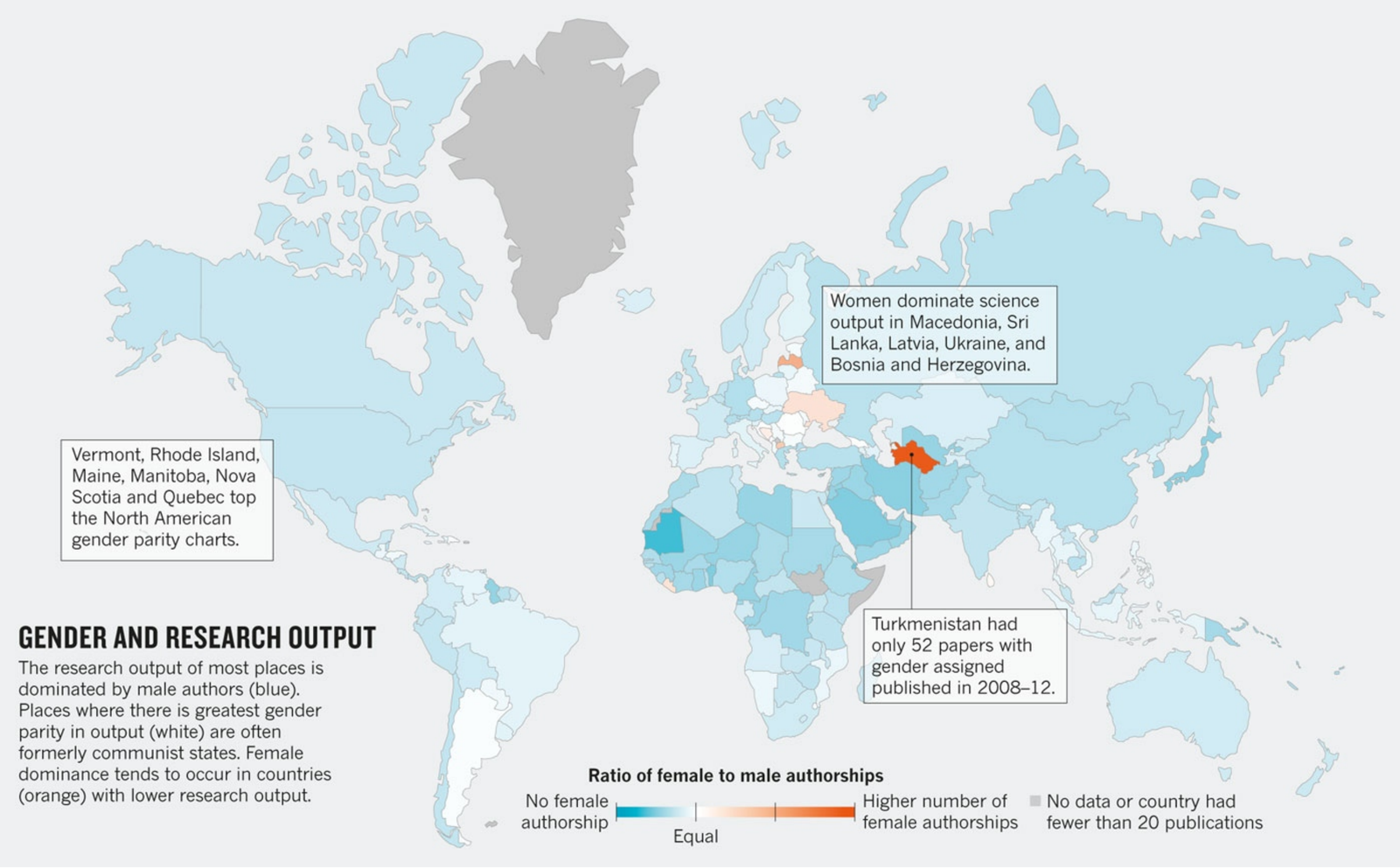
Merton, R. K. (1968). The Matthew Effect in Science. *Science*, *159*(3810), 56–63.

Bol, T., Vaan, M. de, & Rijt, A. van de. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 201719557.

There are known disparities in science

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, *504*(7479), 211.

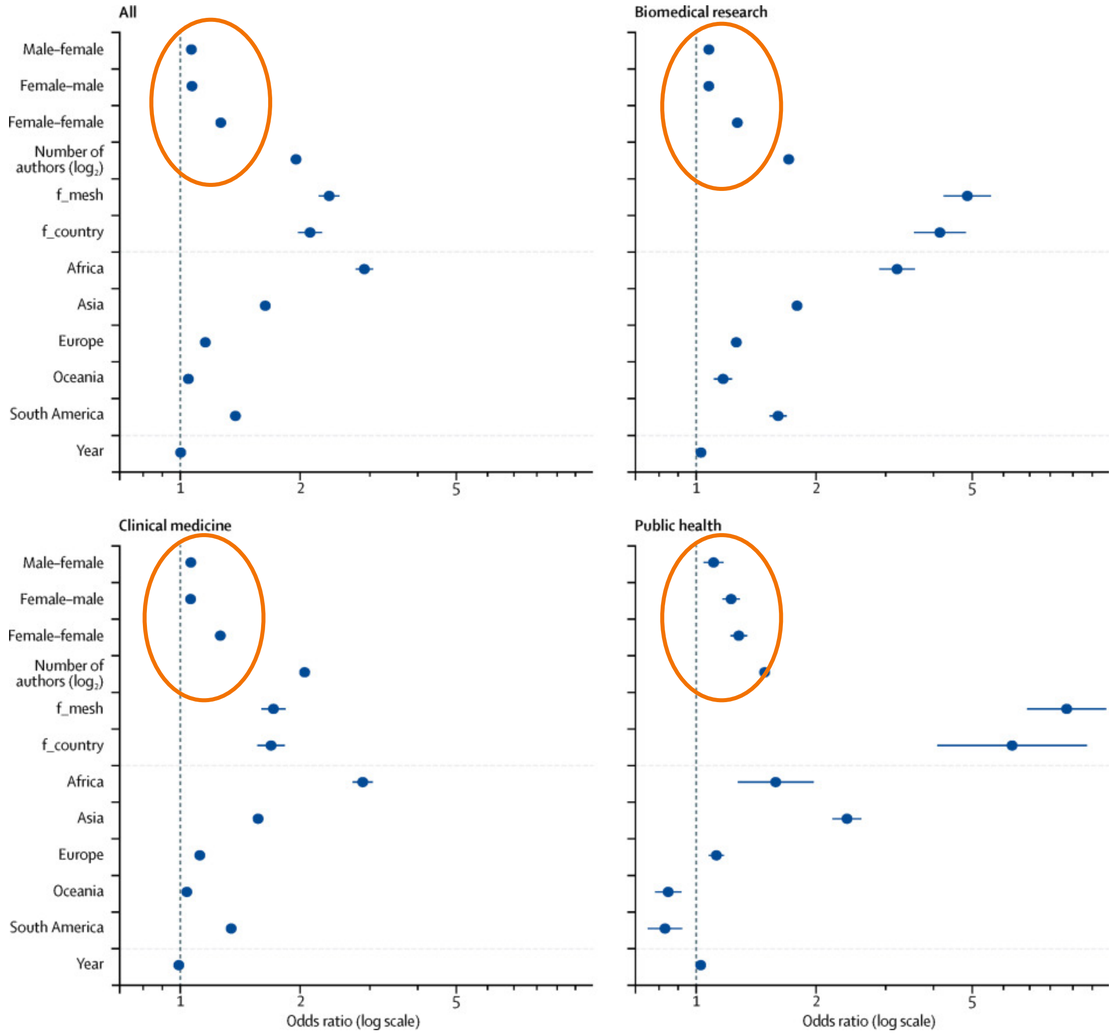
Women make up only around 30% of science worldwide



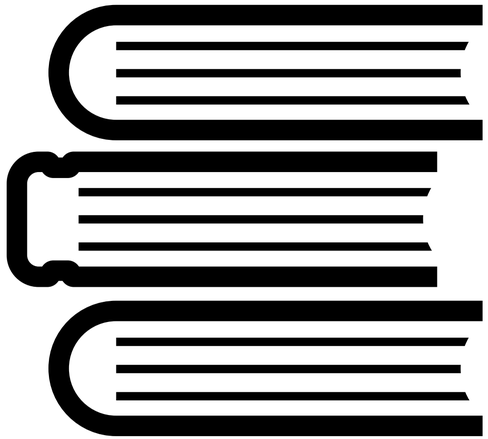
These disparities shape what we know about the world

Sugimoto, C. R., Ahn, Y.-Y., Smith, E., Macaluso, B., & Larivière, V. (2019). Factors affecting sex-related reporting in medical research: A cross-disciplinary bibliometric analysis. *The Lancet*, *393*(10171), 550–559.

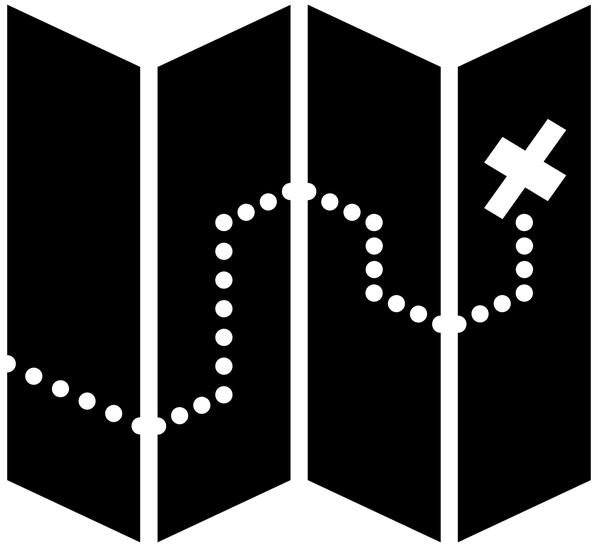
Across fields of medicine, papers with a female author are more likely to report results based on sex



Diverse science is more effective!



Knowledge



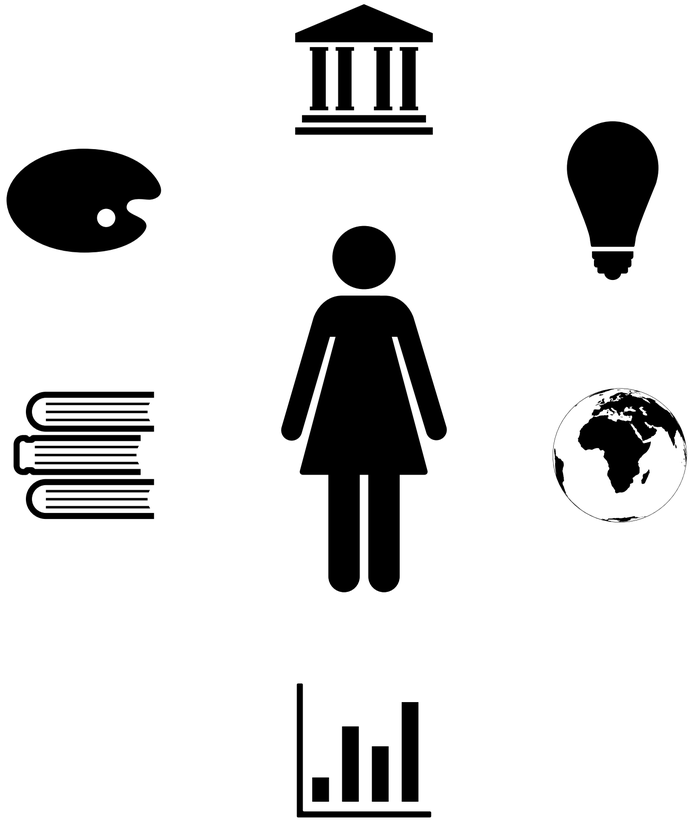
Perspectives

Cognitive toolboxes

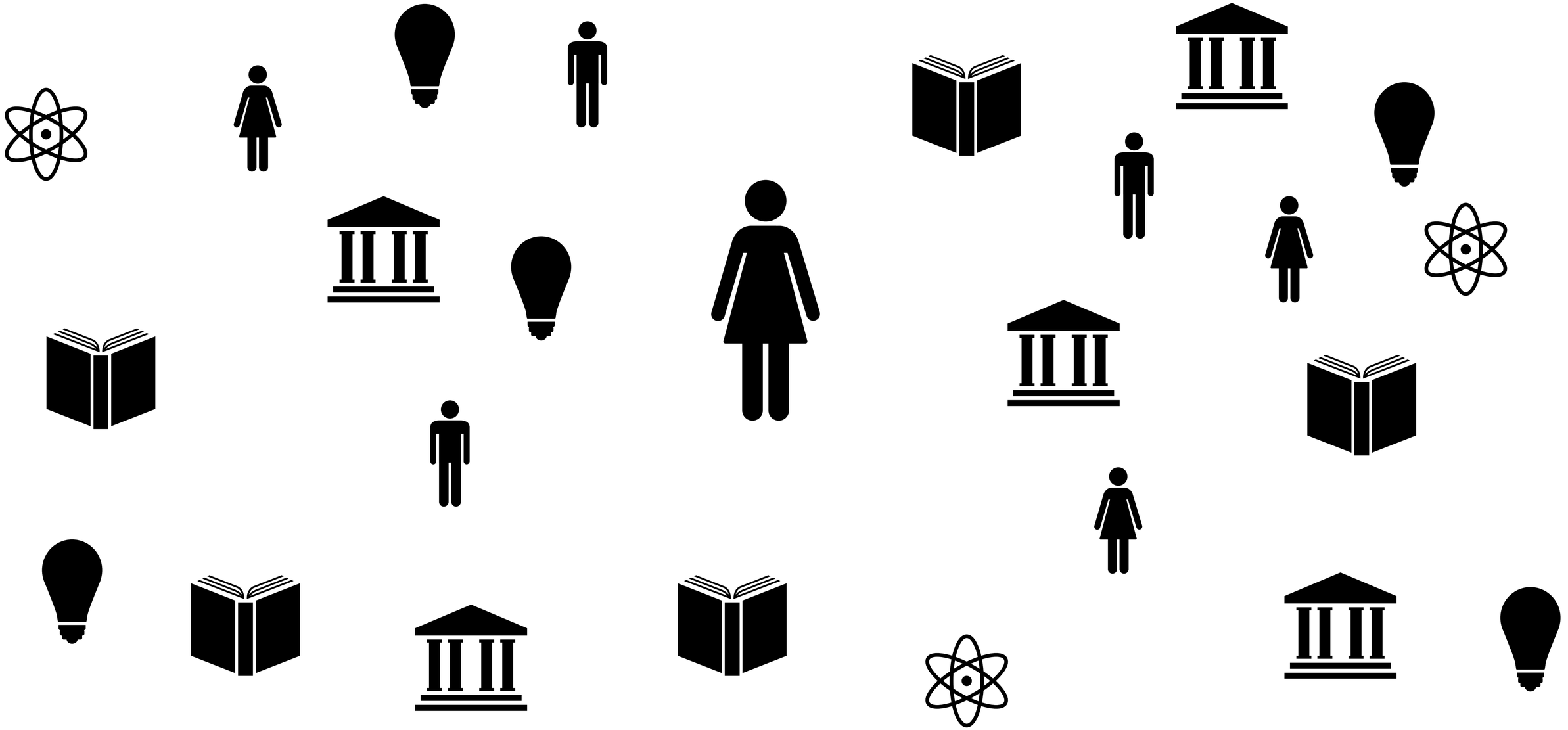
Page, S. E. (2008). *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies.* Princeton University Press.

Researchers are not isolated

They have diverse characteristics that shape their careers



They are *embedded* in a wider context



Thank you!

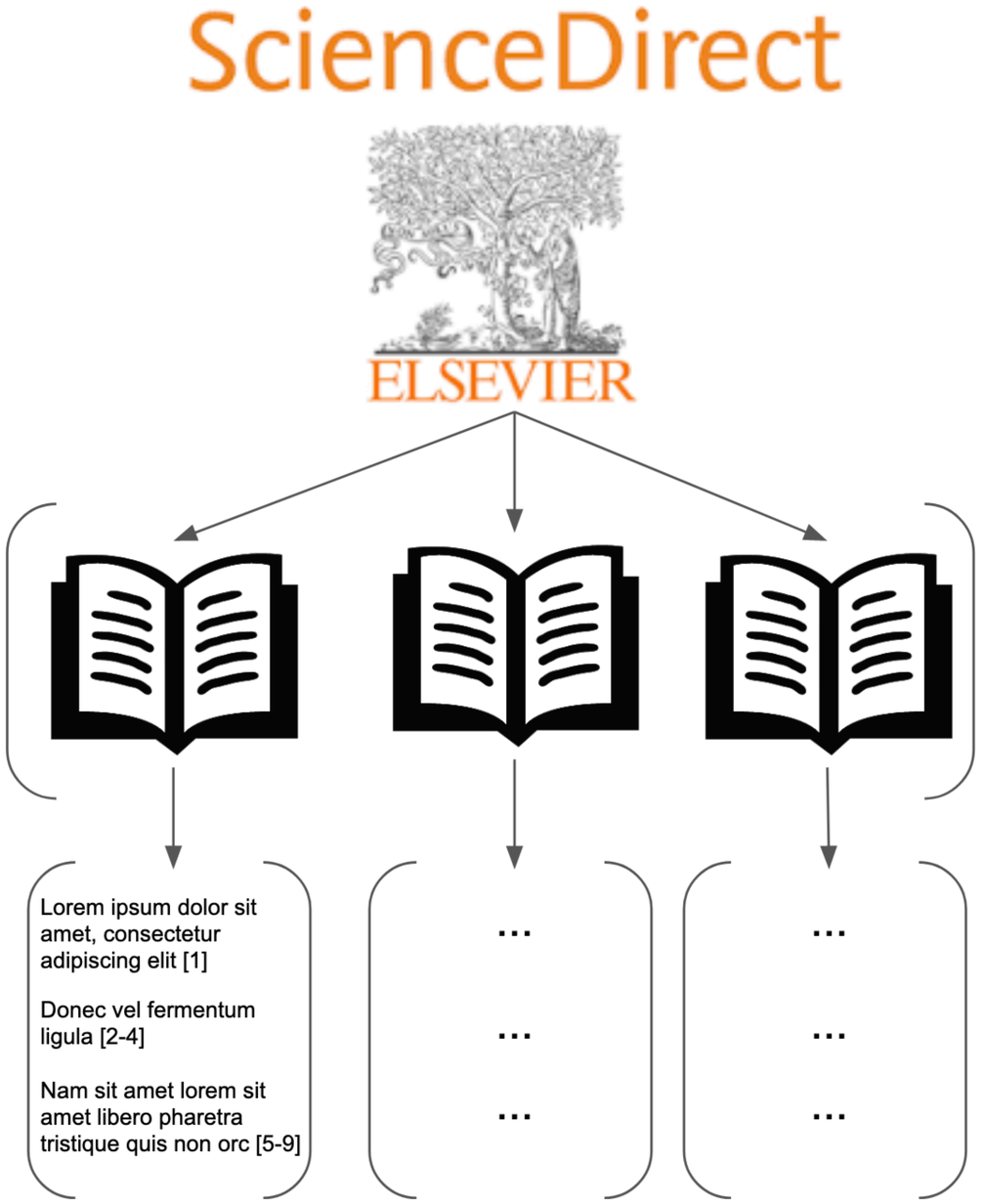
Dakota Murray
@dakotasmurray
dakota.s.murray@gmail.com

Slides at: dakotamurray.me/talk/2021-northwestern/

Appendix

Data

- Disagreement between texts
- Extract citation sentences
- Over 3 million full-text English-language article
- Identify disagreement citations

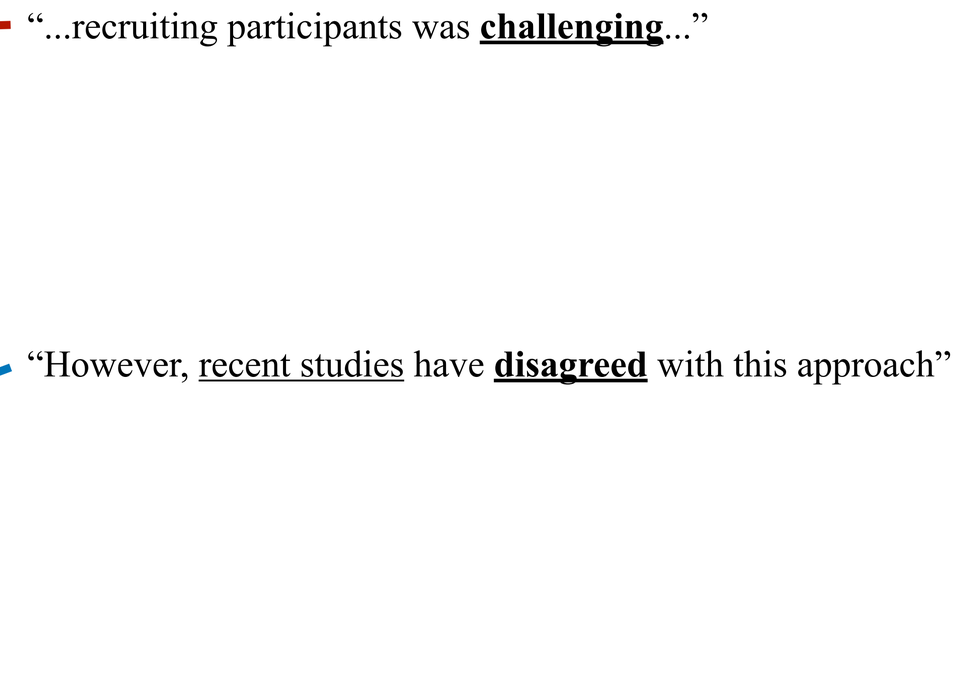


Signal & filter terms

	*standalone*	+studies	+ideas	+methods	+results
Challenge*					
Conflict*					
Contradict*					
Contrary					
Contrast*					
Contravers*					
Debat*					
Differ*					
Disagree*					
Disprov*					
No consensus					
Questionable*					
Refut*					

Signal & filter terms

	*_standalone_*	+studies	+ideas	+methods	+results
Challenge*					
Conflict*					
Contradict*					
Contrary					
Contrast*					
Contravers*					
Debat*					
Differ*					
Disagree*					
Disprov*					
No consensus					
Questionable*					
Refut*					



Signal & filter terms

	*_standalone_*	+studies	+ideas	+methods	+results
Challenge*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Conflict*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Contradict*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Contrary	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Contrast*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Contravers*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Debat*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Differ*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Disagree*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Disprov*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
No consensus	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Questionable*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*
Refut*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*	*50 citation sentences*

Which combinations are most valid?

Sampled 50 citation sentences for every combination

Two coders independently labeled them as Valid disagreement or Invalid

Take the most valid as our indicator of disagreement

Signal & filter terms

	standalone	+studies	+ideas	+methods	+results
Challenge*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Conflict*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Contradict*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Contrary	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Contrast*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Contravers*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Debat*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Differ*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Disagree*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Disprov*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
No consensus	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Questionable*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences
Refut*	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences	50 citation sentences

Which combinations are most valid?

Sampled 50 citation sentences for every combination

Two coders independently labeled them as Valid disagreement or Invalid

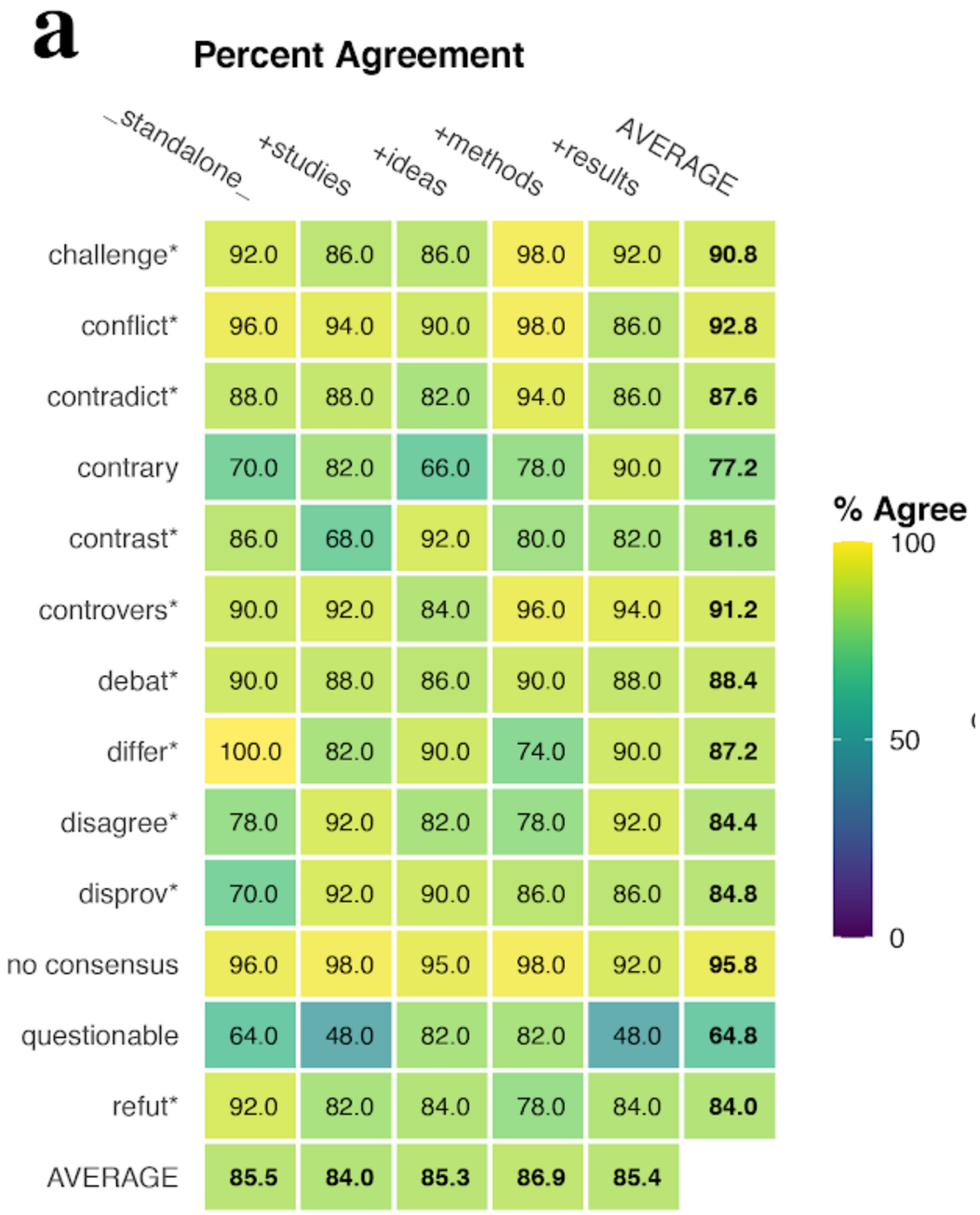
Take the most valid as our indicator of disagreement

23 queries representing ~450,000 citation sentences

Non-exhaustive, but precise!

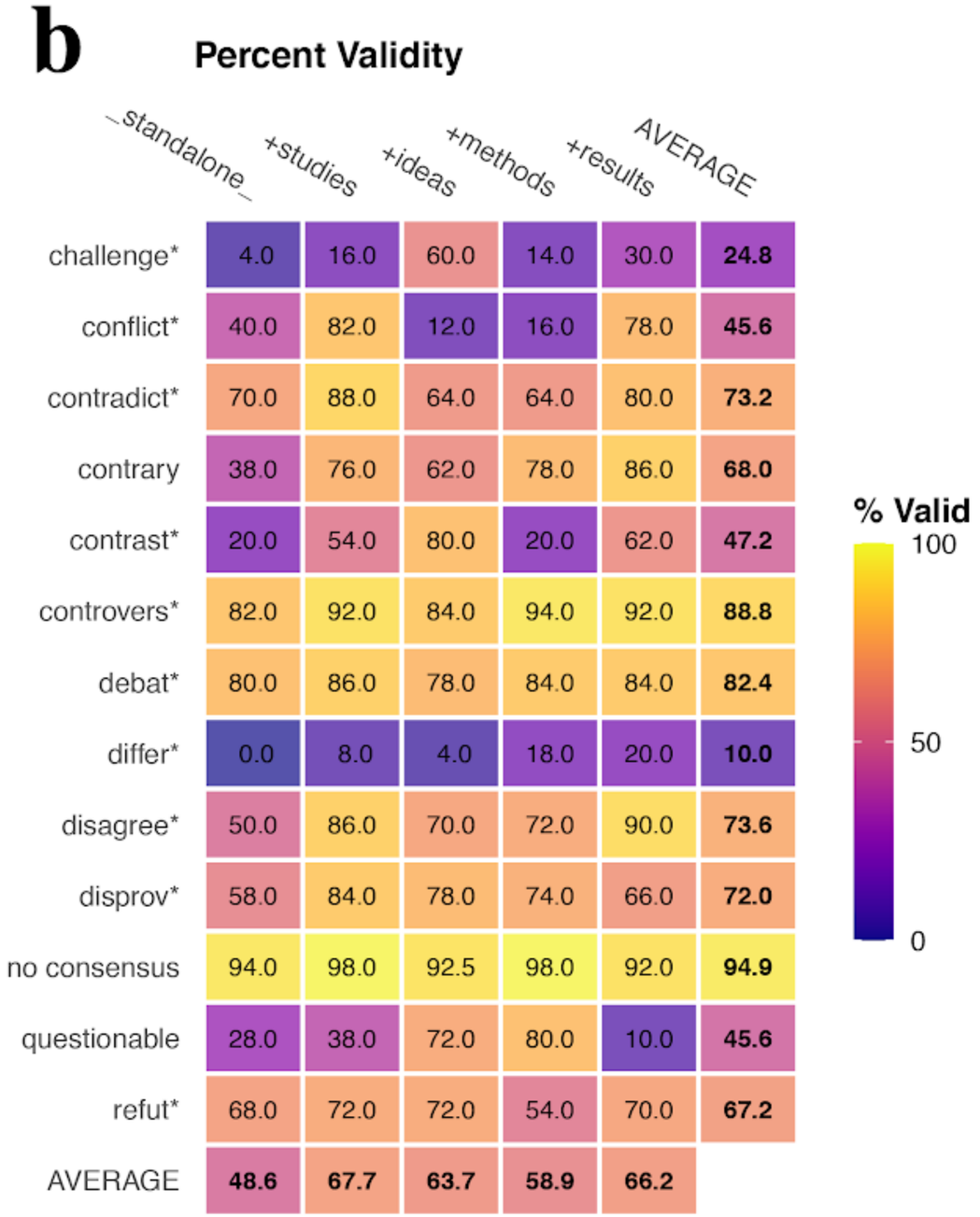
Validity

Coder 1: ✓ Valid **Coder 2: ✓ Valid**	Coder 1: ✓ Valid Coder 2: X Invalid
Coder 1: X Invalid Coder 2: ✓ Valid	**Coder 1: X Invalid** **Coder 2: X Invalid**

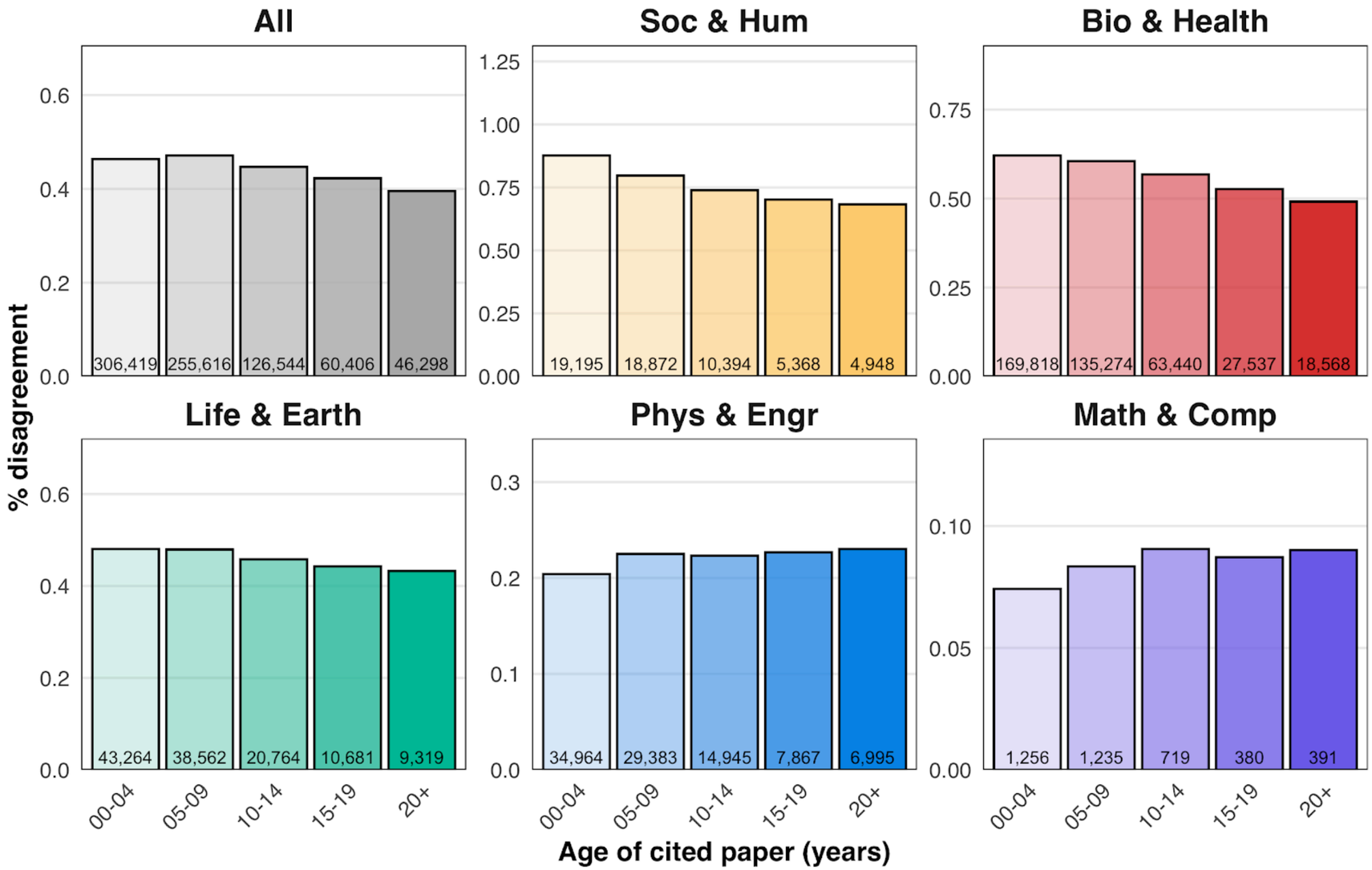


Validity

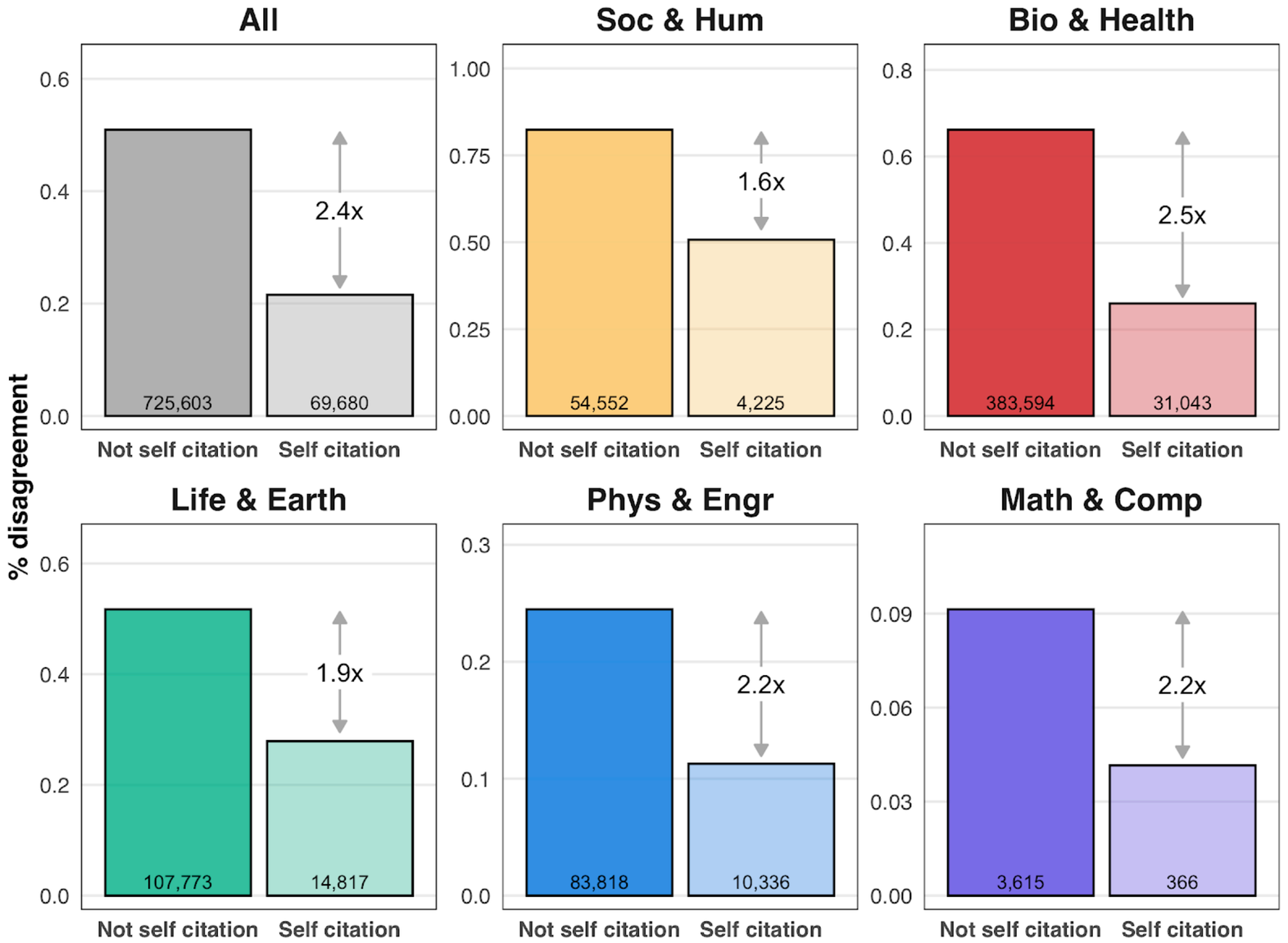
Coder 1: ✓ Valid **Coder 2: ✓ Valid**	Coder 1: ✓ Valid Coder 2: X Invalid
Coder 1: X Invalid Coder 2: ✓ Valid	Coder 1: X Invalid Coder 2: X Invalid



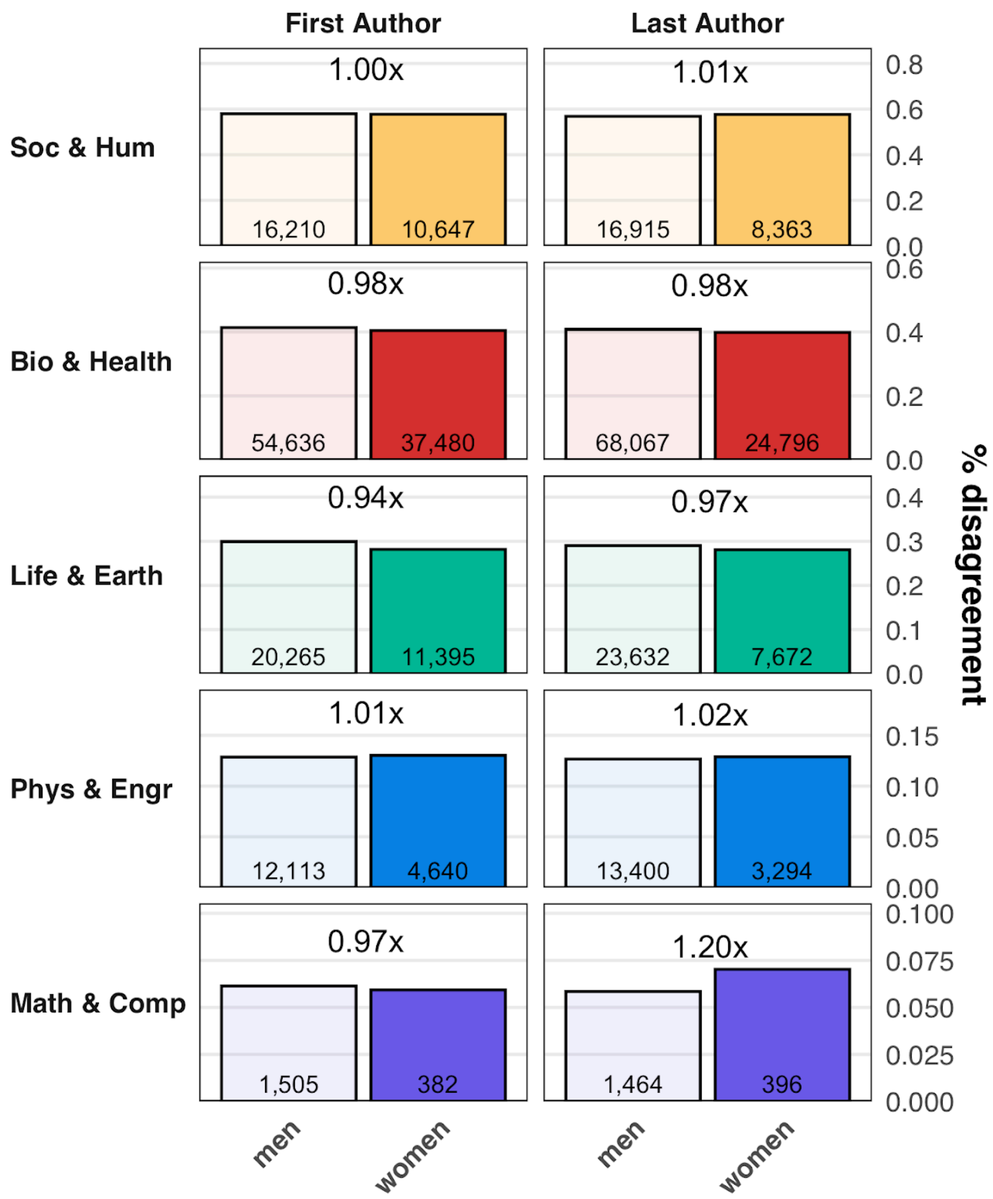
Less disagreement for older papers



Disagreement and self-citation



Disagreement and gender



Appendix — Mobility

Mobility is central to science

Institutionalized in evaluation

Article 19: Requirement for International Visits
When applying for promotion to full professor or equivalent rank, the applicants who were born after January 1, 1970, must complete at least a 6-month international visit.

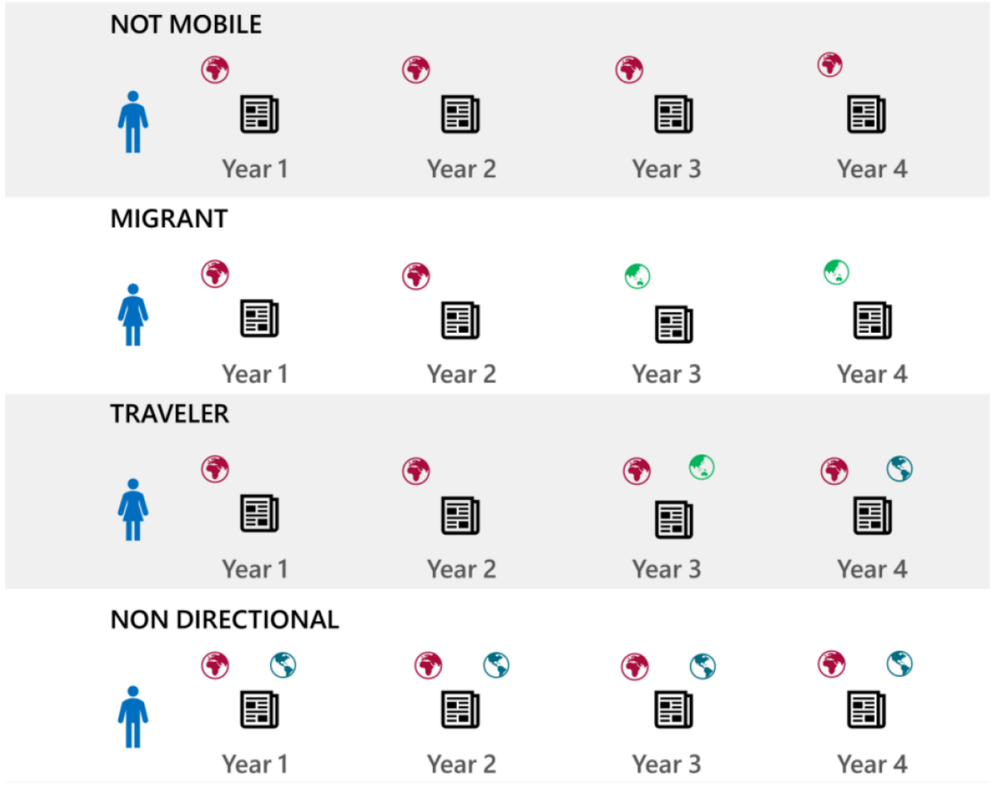
- Hangzhou Dianzi University

Thanks to Fei She

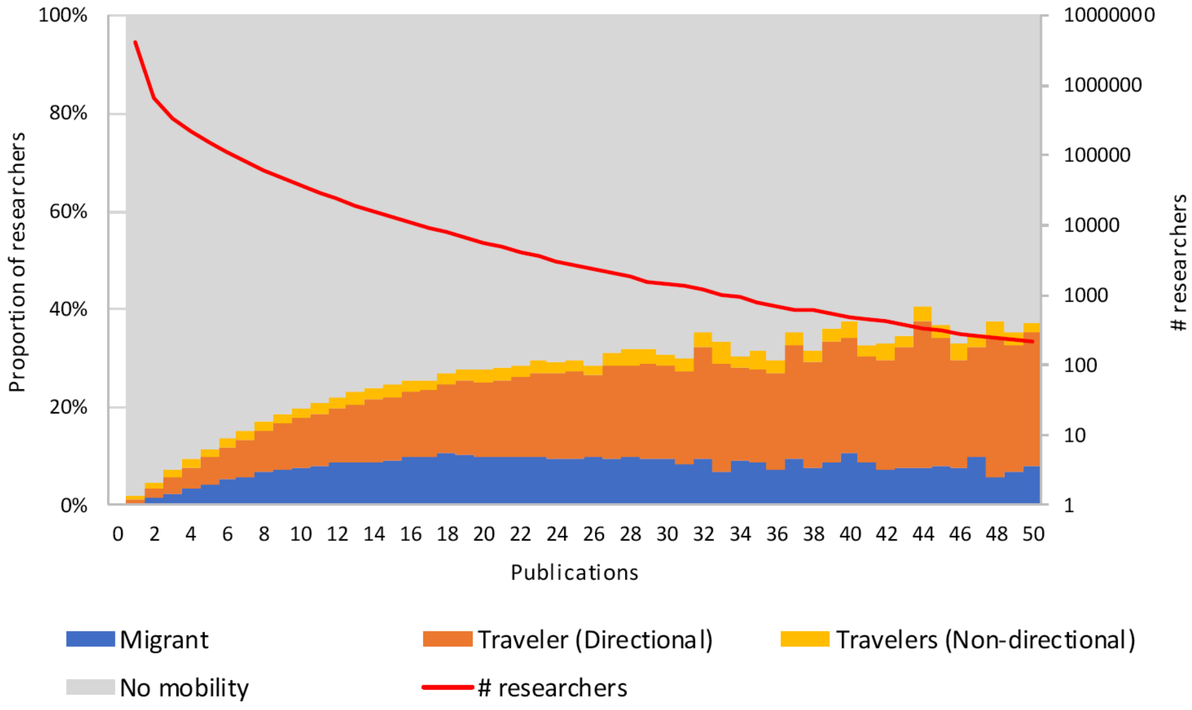
Mobility drives economies, cultural exchange, epidemics

Hanson, R., Mouton, C. A., Grissom, A. R., & Godges, J. P. (2020). *COVID-19 Air Traffic Visualization: Decisionmakers Should Base Travel Restrictions on Infection Rates Per Capita and Air Traffic Levels*.

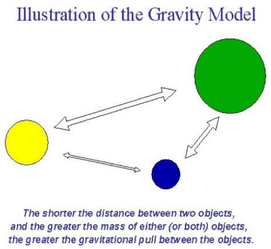
Mobility is complicated



Robinson-Garcia, N., Sugimoto, C. R., **Murray, D.**, Yegros-Yegros, A., Larivière, V., & Costas, R. (2018). Scientific mobility indicators in practice: International mobility profiles at the country level. *El Profesional de La Información*, *27*(3), 511.



Robinson-Garcia, N., Sugimoto, C. R., **Murray, D.**, Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, *13*(1), 50–63.



Gravity Model

- Popular model of mobility
- ***Flows*** between places (co-affiliations) a function of their size and distance $T_{ij} = C\frac{m_i m_j}{f(r_{ij})}$
- We use two kinds of distance measure.
  1. Geographical distance
  2. Embedding distance $d_{ij} = 1 - \frac{v_i \cdot v_j}{|v_i||v_j|}$

A basic principle: a good *representation* allows *prediction*.

$$P(w_t|w_c) \gg P(w_{\text{random}}|w_c)$$

Target Context

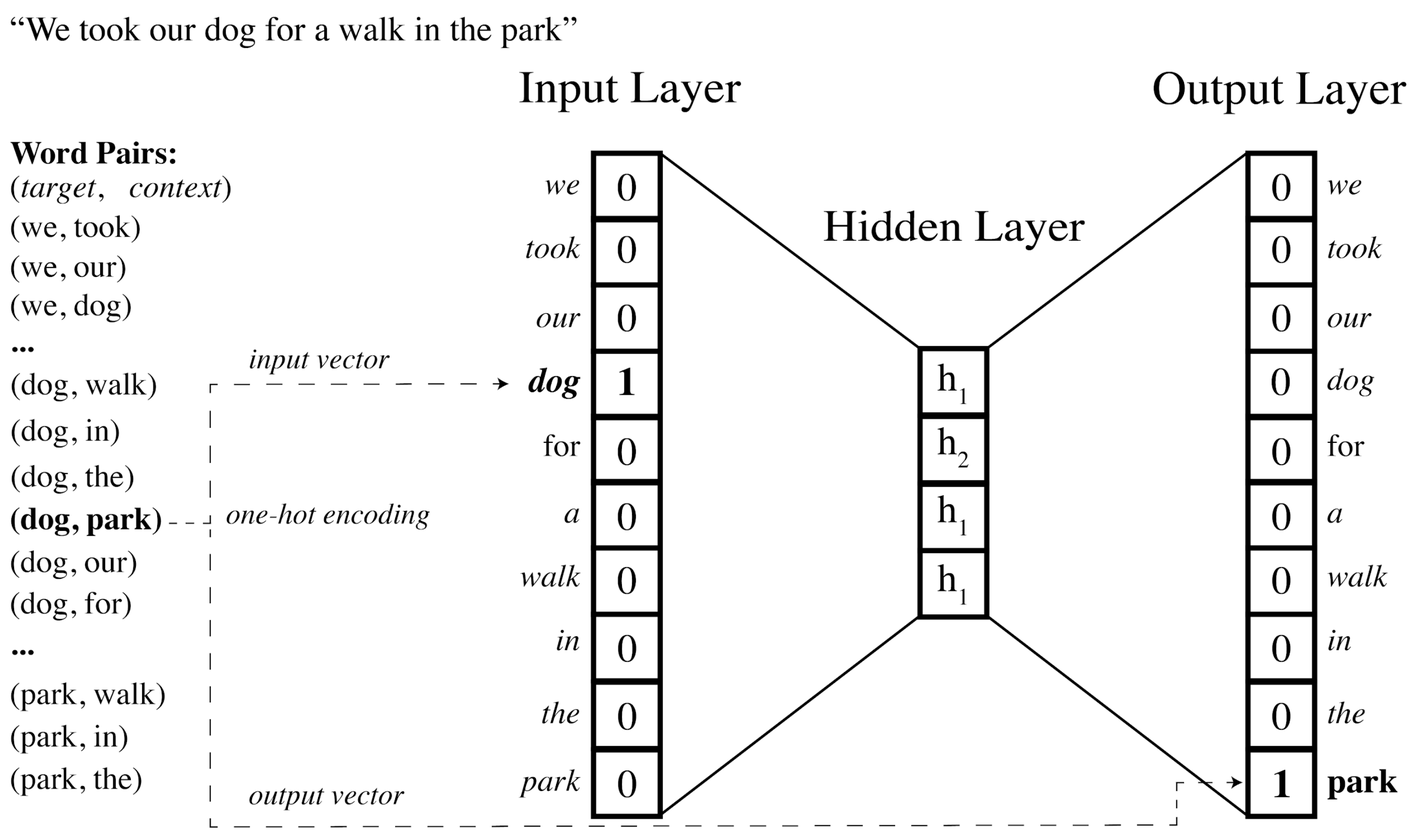
Let's assume that we can calculate the conditional probability with a function (e.g. dot product) of word **vectors** and **learn those vectors with a neural network**.

Learning word embeddings

Train a neural network to predict context words given a target

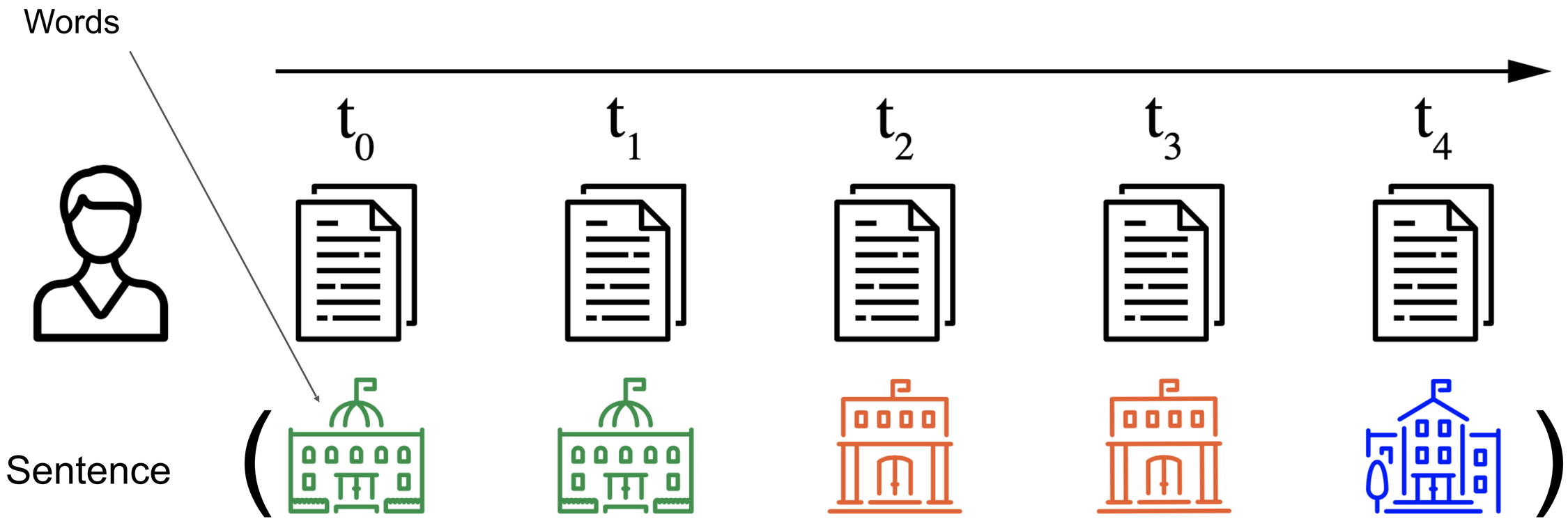
The hidden layer maps targets to concepts!

Words with similar contexts will have a similar “mapping” vector in the hidden layer



They don't have to be real "words" or "sentences"

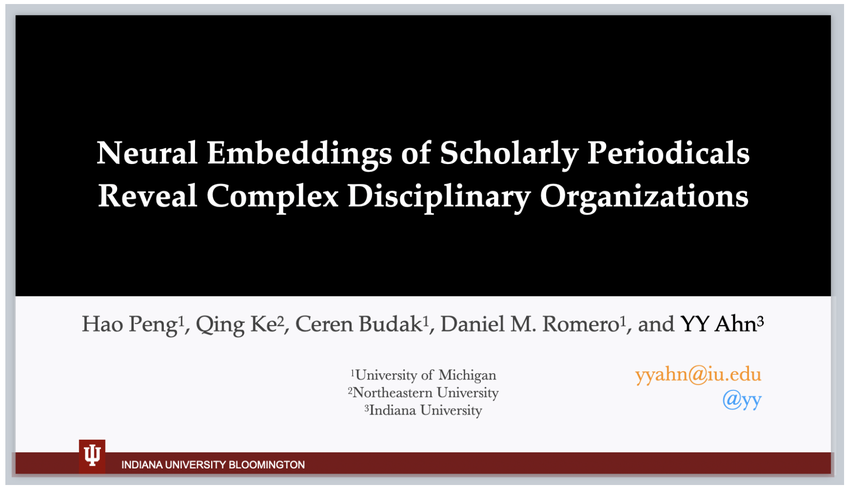
Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use **trajectories of scientists** as **sentences** and **organizations** as **words**.



They don't have to be real "words" or "sentences"

Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use the **trajectories of scientists** as **sentences** and **institutions** as **words**.

(

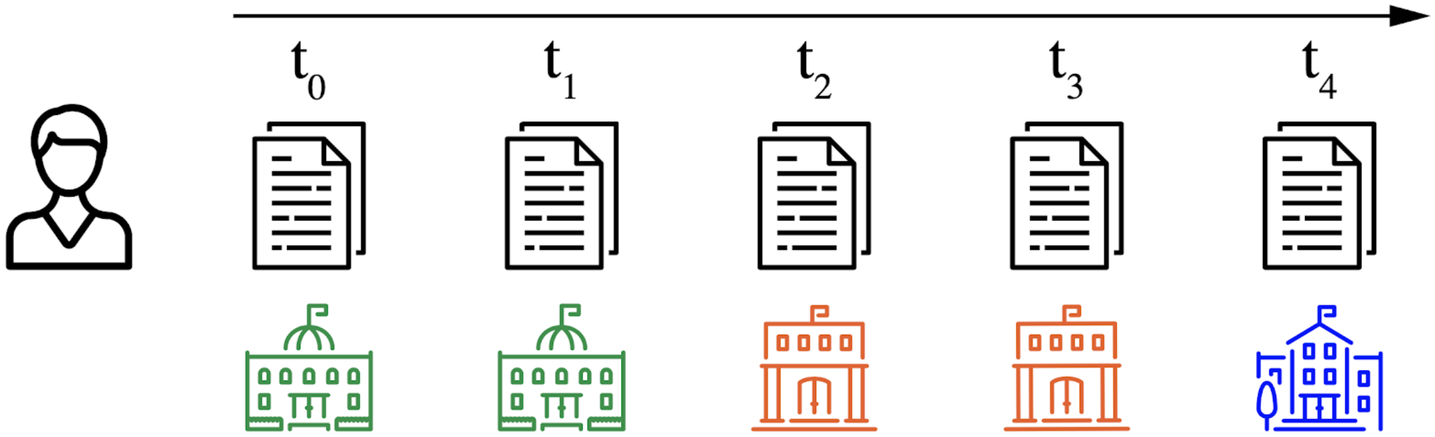


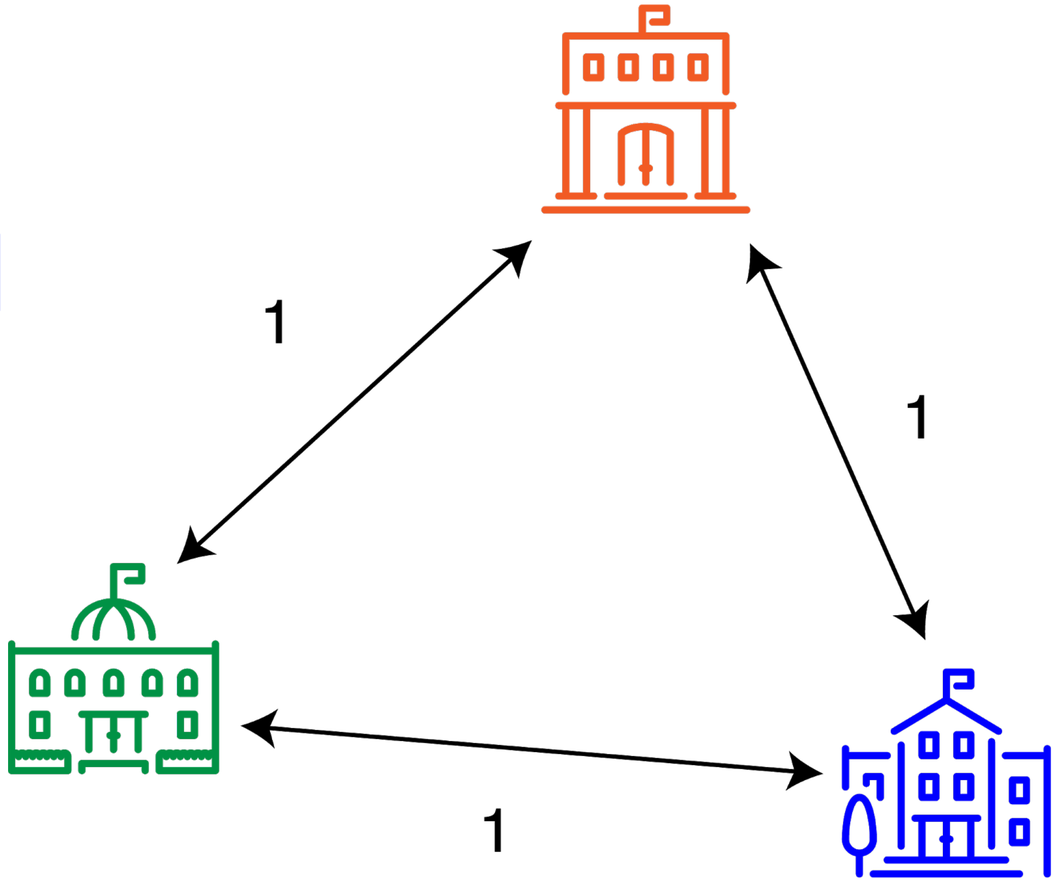
IC2S2 "embedding" session

)

Each a trajectory of organizations

Derive “flux” from scientists career trajectories





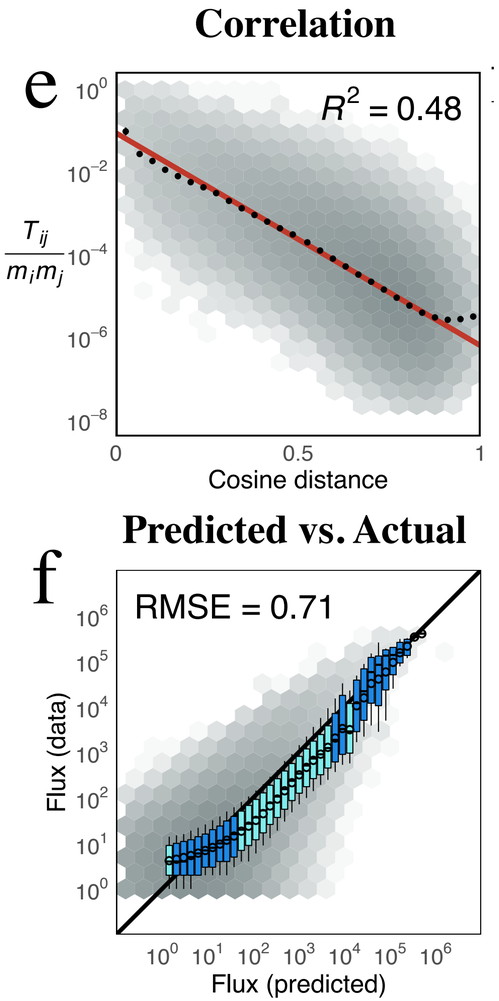
Aside: why does it work?

Relationship between word2vec and the Gravity Law

- Distance in the embedding space maps well to the gravity model
- Both correlation and prediction
- Word2vec finds a representation to predict adjacent words, but the gravity law emerges
- Deep connection between them?
- Preliminary work, but ideas welcome!

$$T_{ij} = C m_i m_j f(r_{ij})$$

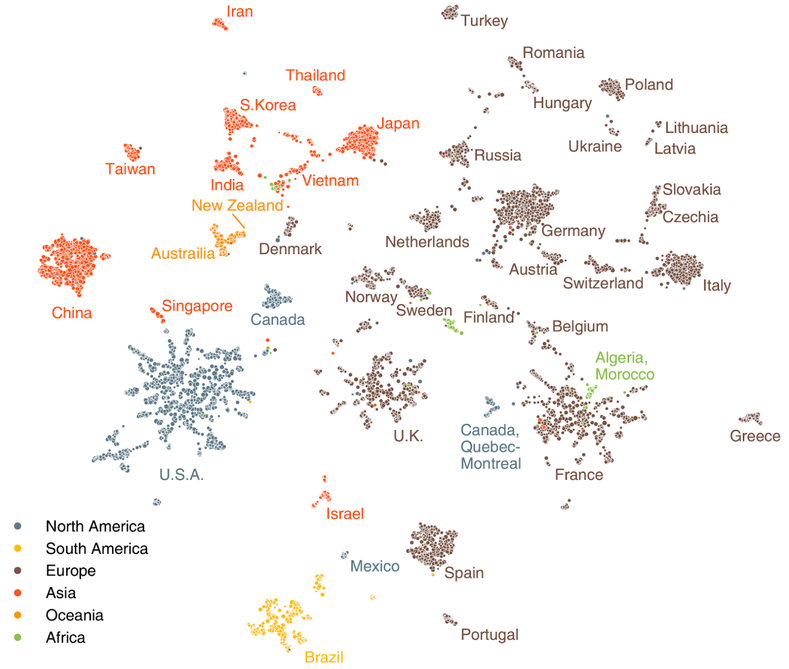
$$\frac{T_{ij}}{m_i m_j} = f(r_{ij})$$

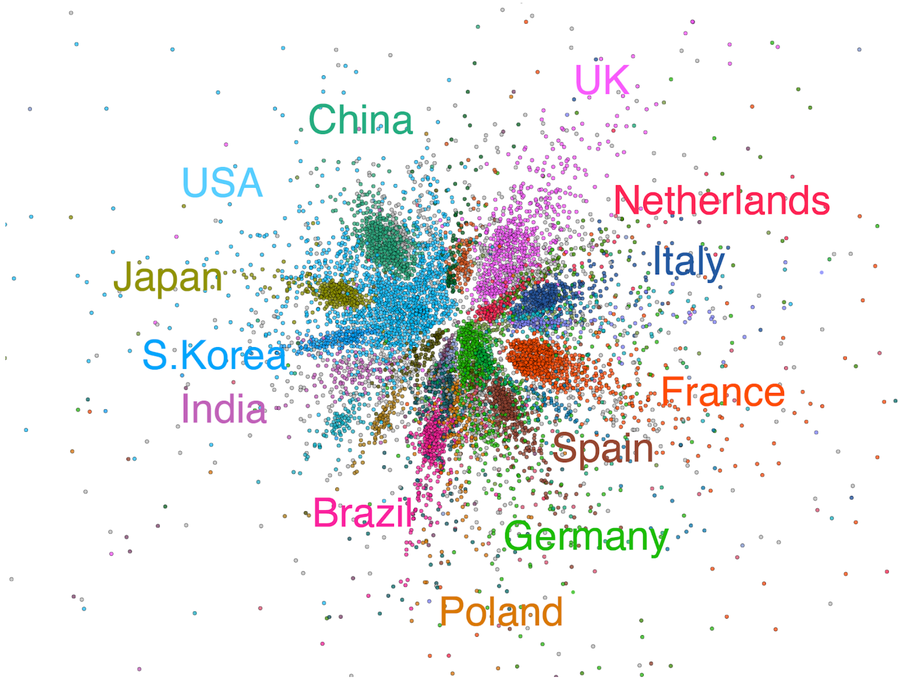


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

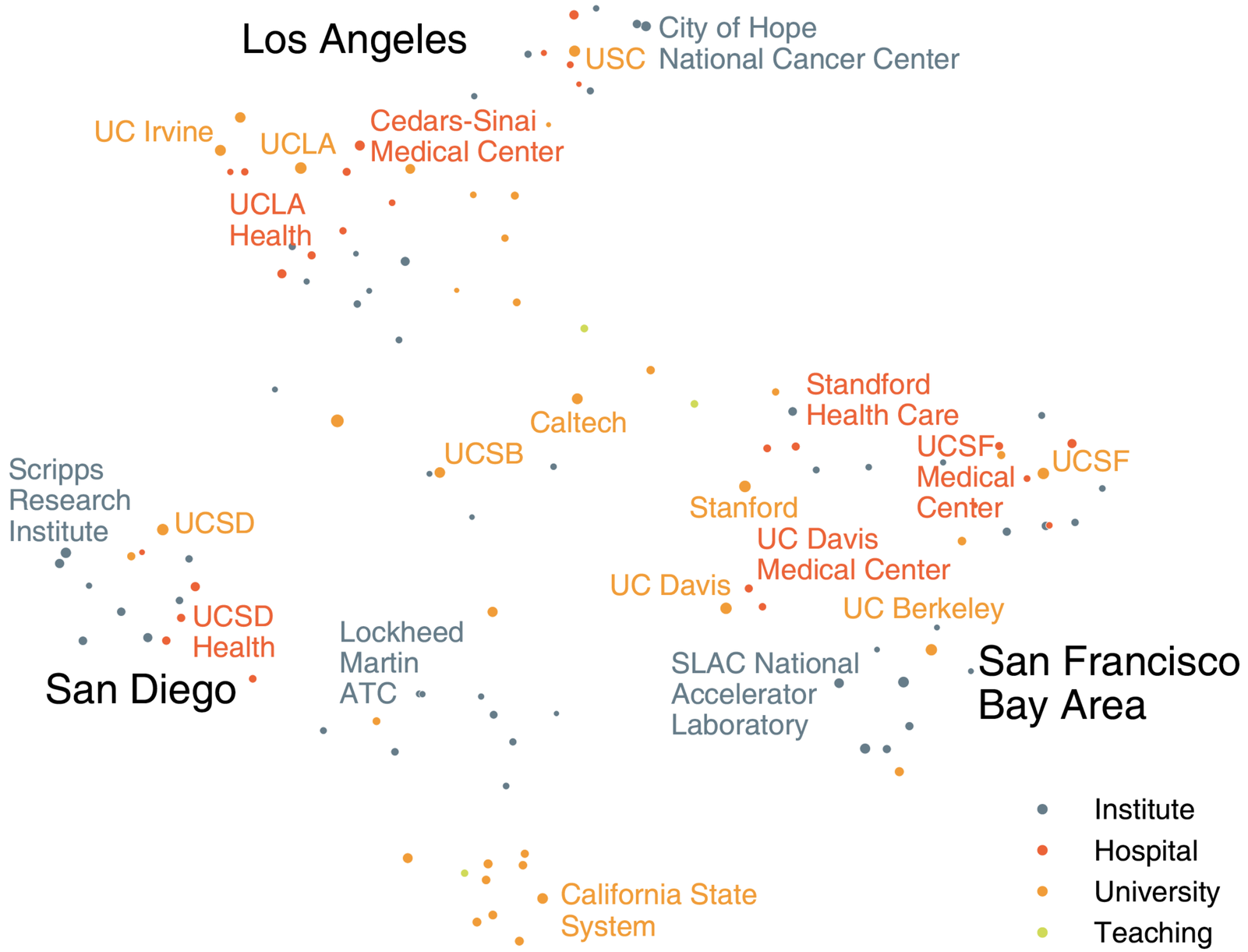
Why not networks?

- Structure obfuscated in visualizations
- Poor performance (reported in the paper)
- Many techniques, tuning required
- Missing edges
- Embeddings provide access to many interesting techniques

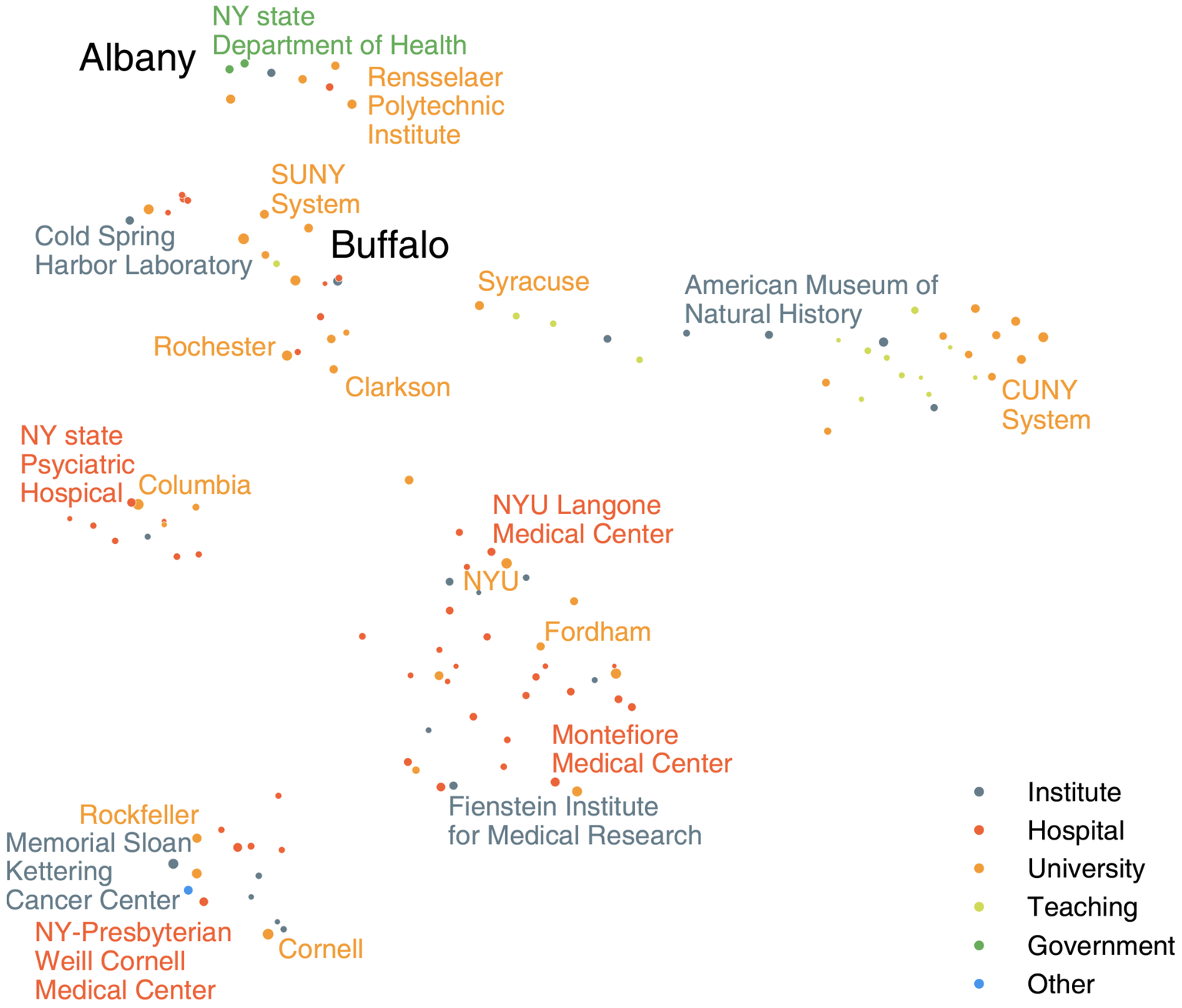




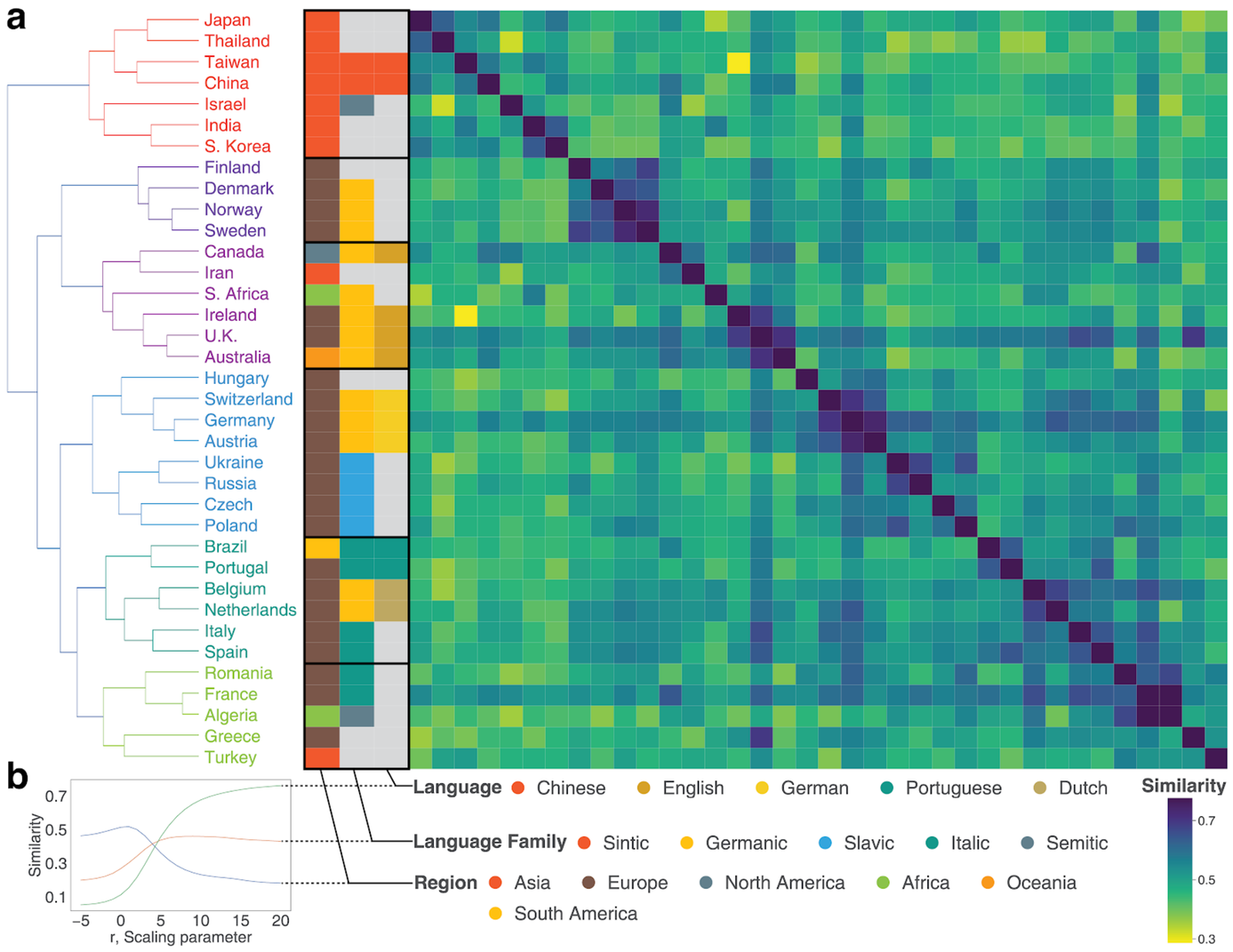
California Structure



New York Structure



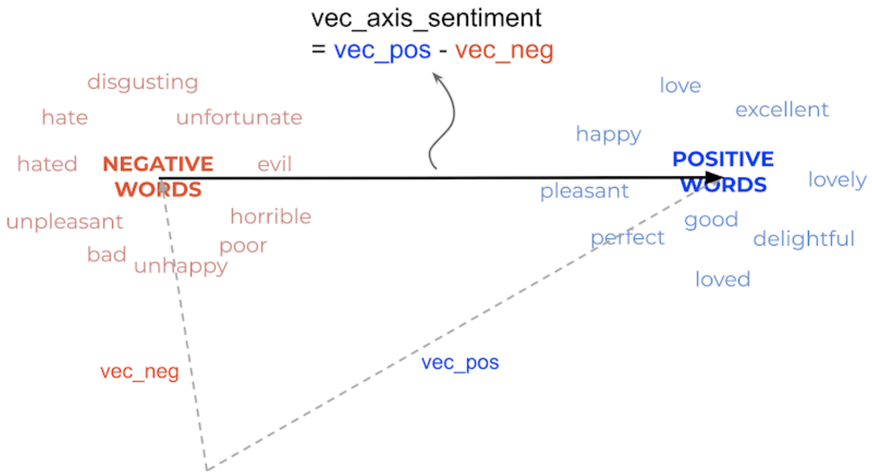
Geography, then language, structure the vector space



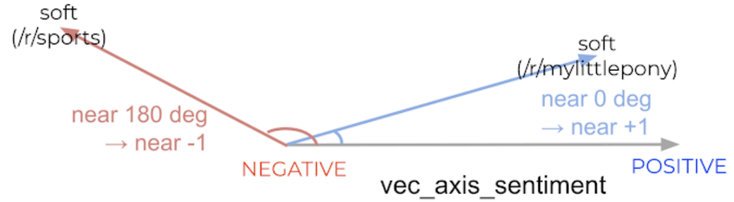
But also prestige…

Leverage the semantic properties of the embedding

Step 1: Define a semantic axis

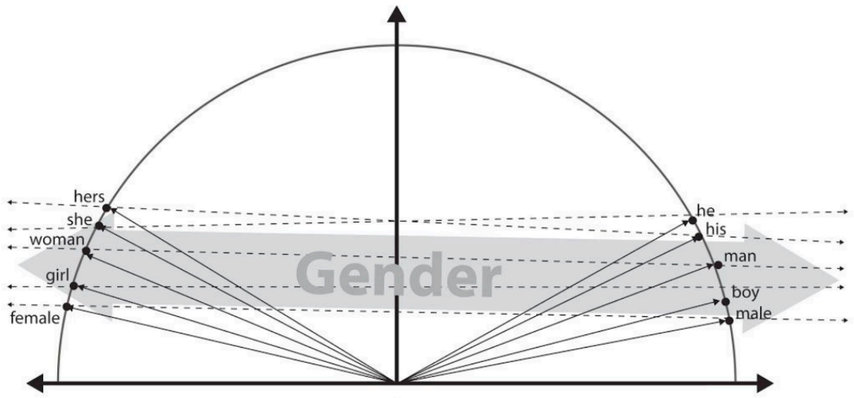


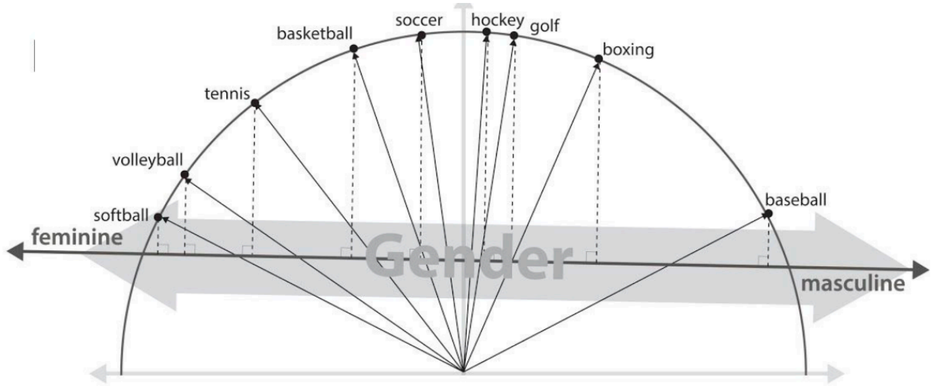
Step 2. Locate words on the semantic axis



Compute a cosine similarity between a word vector and axis vector

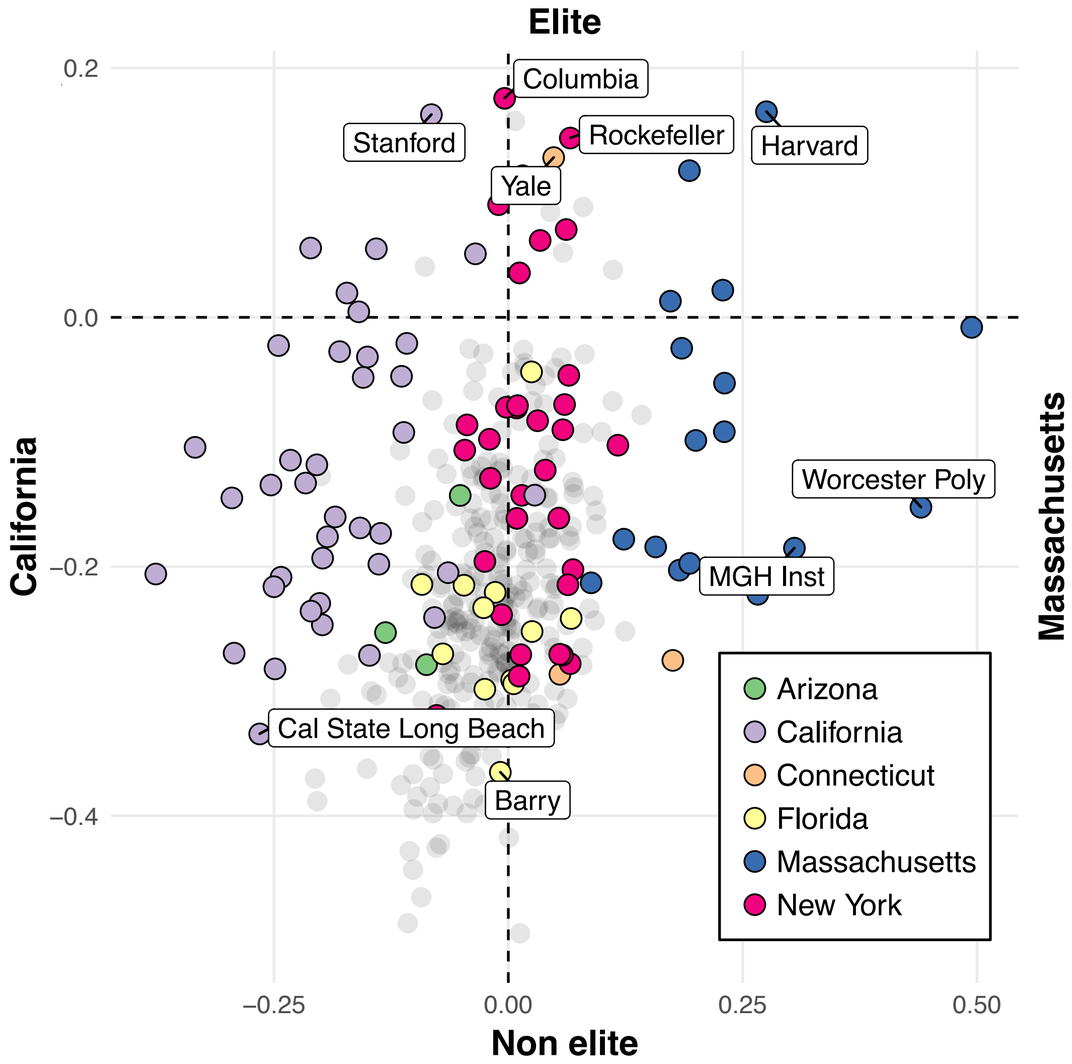
An, J., Kwak, H., & Ahn, Y.-Y. (2018). SemAxis: A Lightweight Framework to Characterize Domain-Specific Word Semantics Beyond Sentiment. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2450–2461.



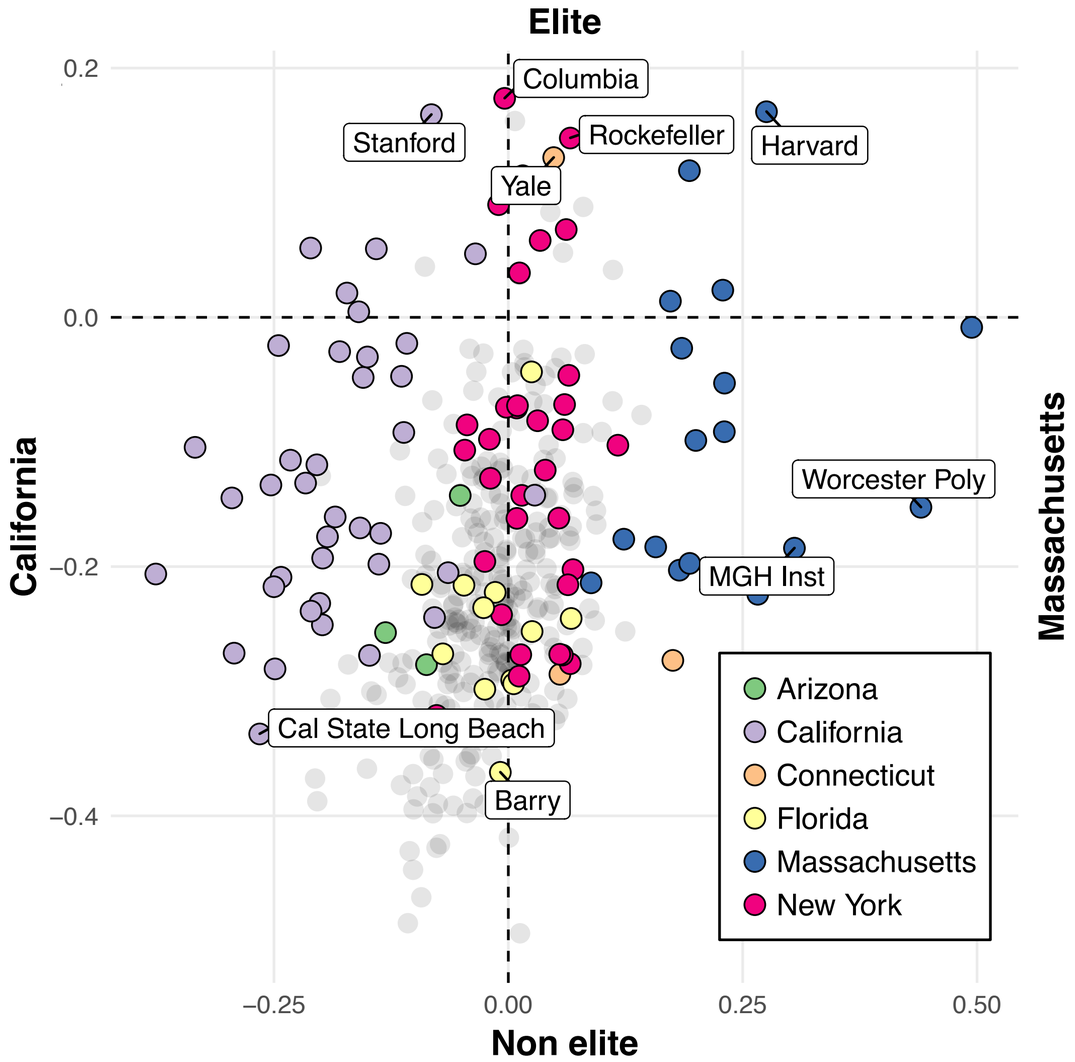


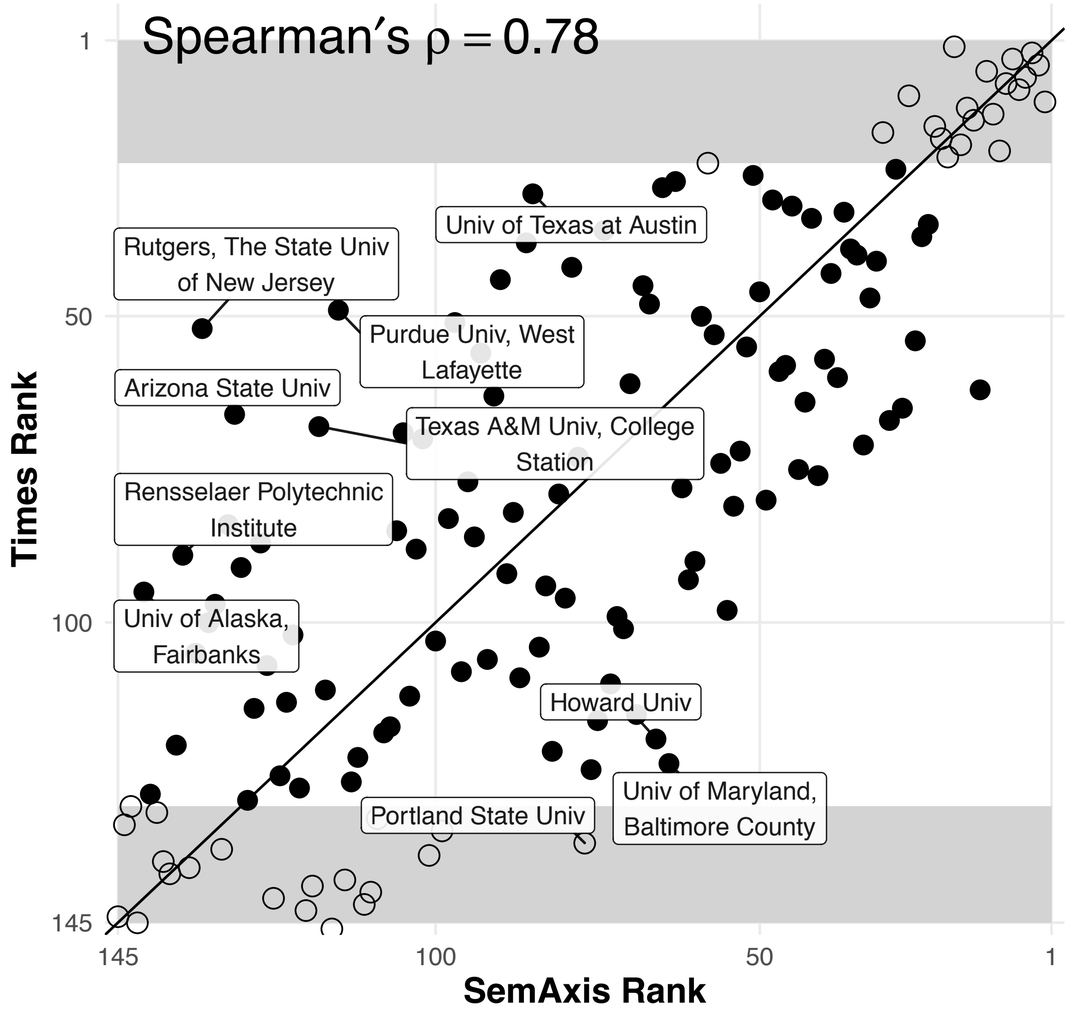
Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*, *84*(5), 905–949.

SemAxis using Geography and Prestige



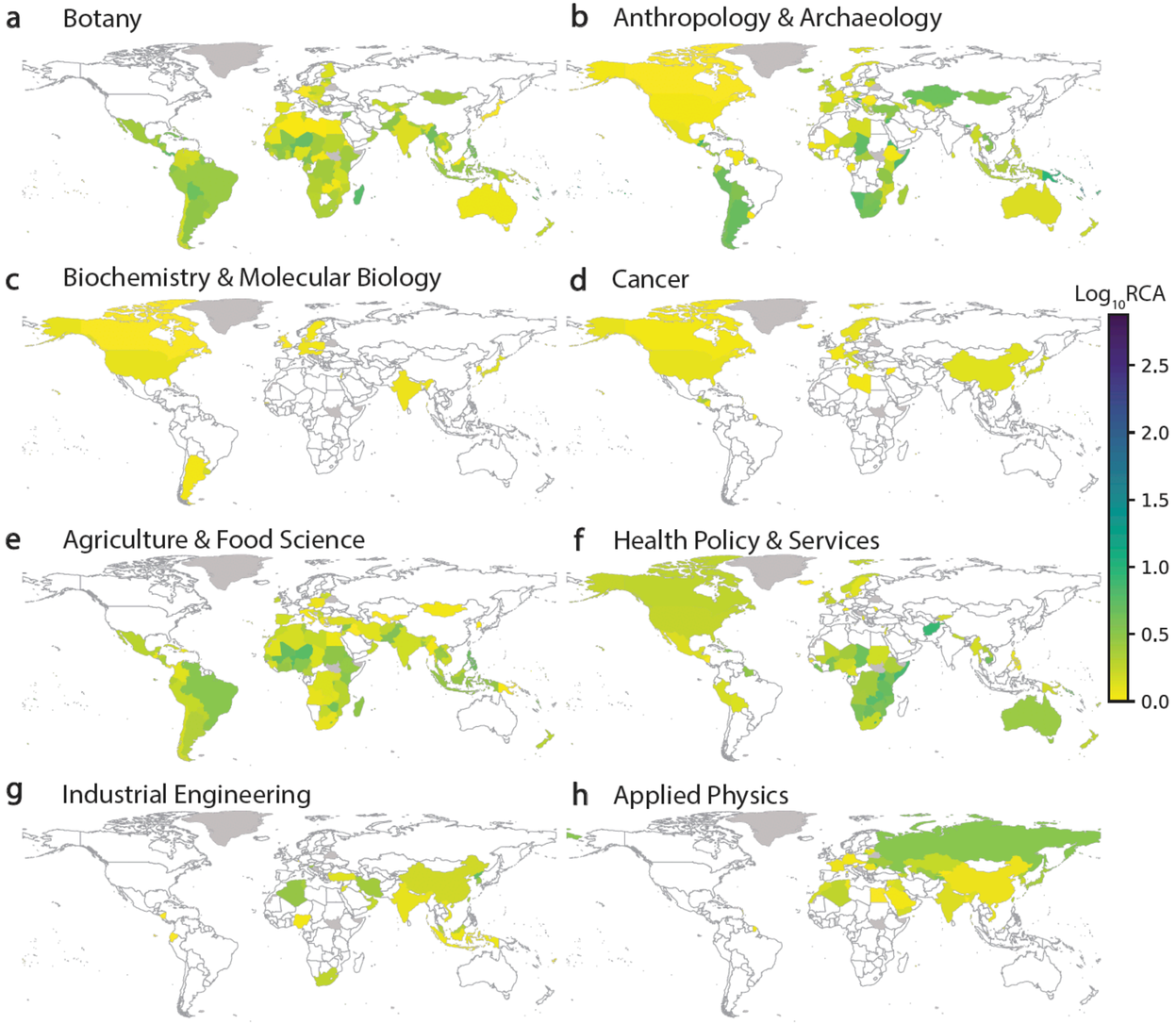
SemAxis reconstructs university prestige





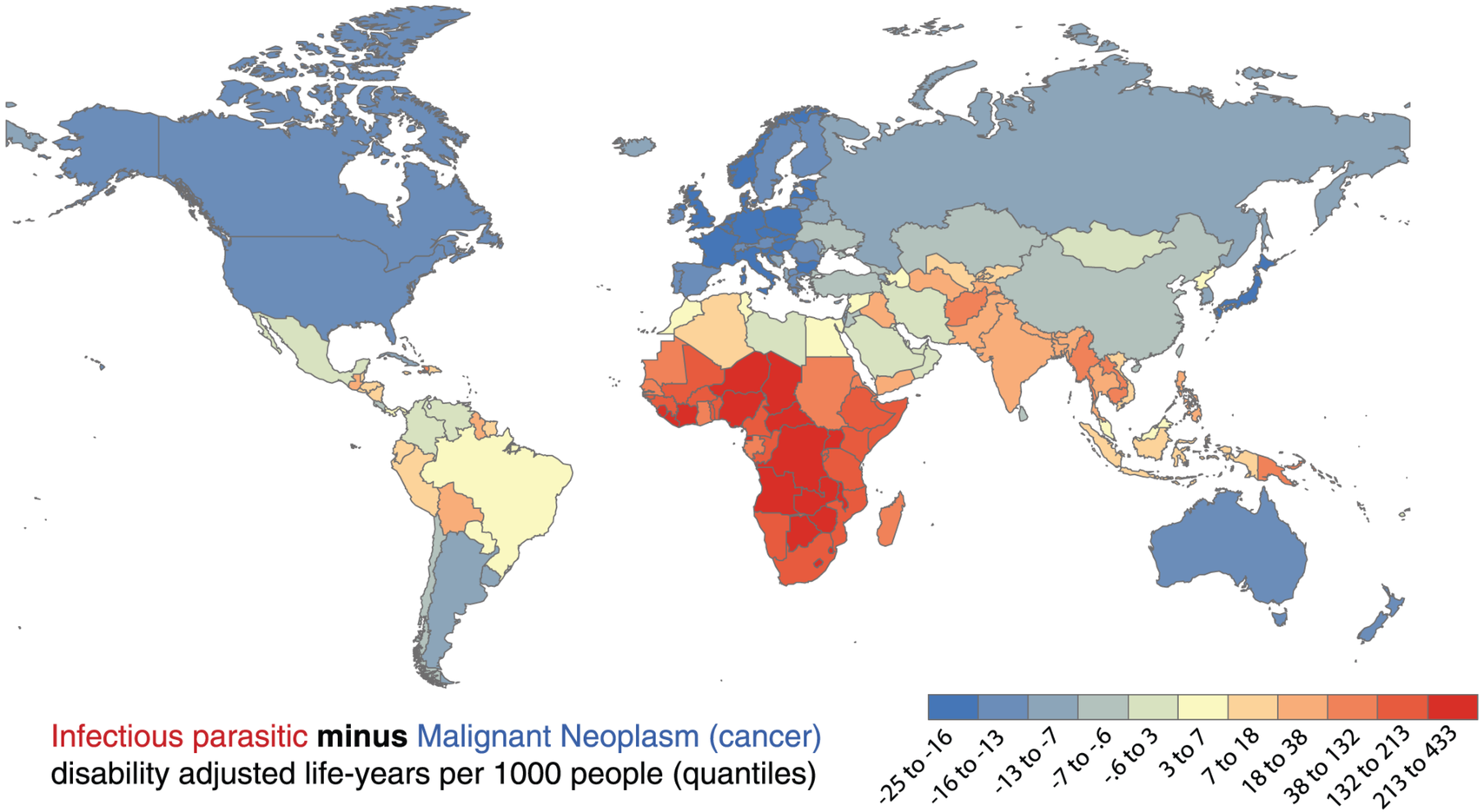
Countries do different kinds of research

Miao, L., **Murray, D.**, Jung, W., Larivière, V., Sugimoto, C. R., Ahn, Y., The scientific development of nations. (In Preparation)



These countries research locally-relevant topics

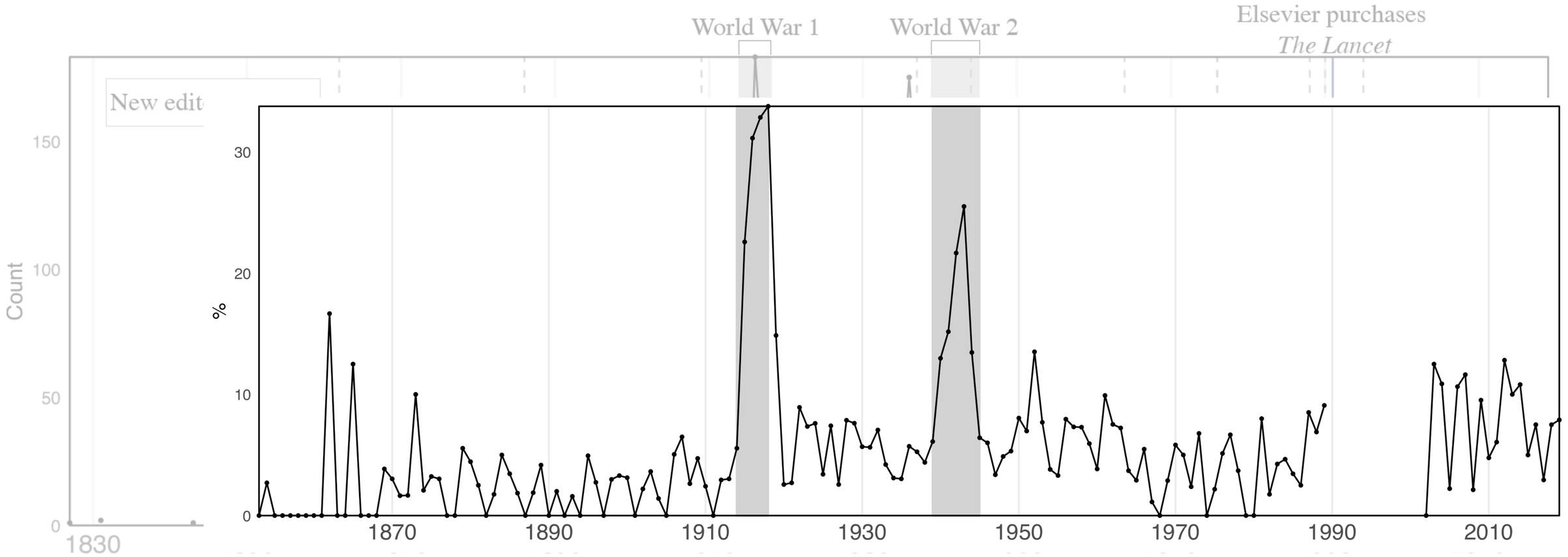
Evans, J. A., Shim, J.-M., & Ioannidis, J. P. A. (2014). Attention to Local Health Burden and the Global Disparity of Health Research. *PLOS ONE*, *9*(4), e90147.



Appendix—obituary

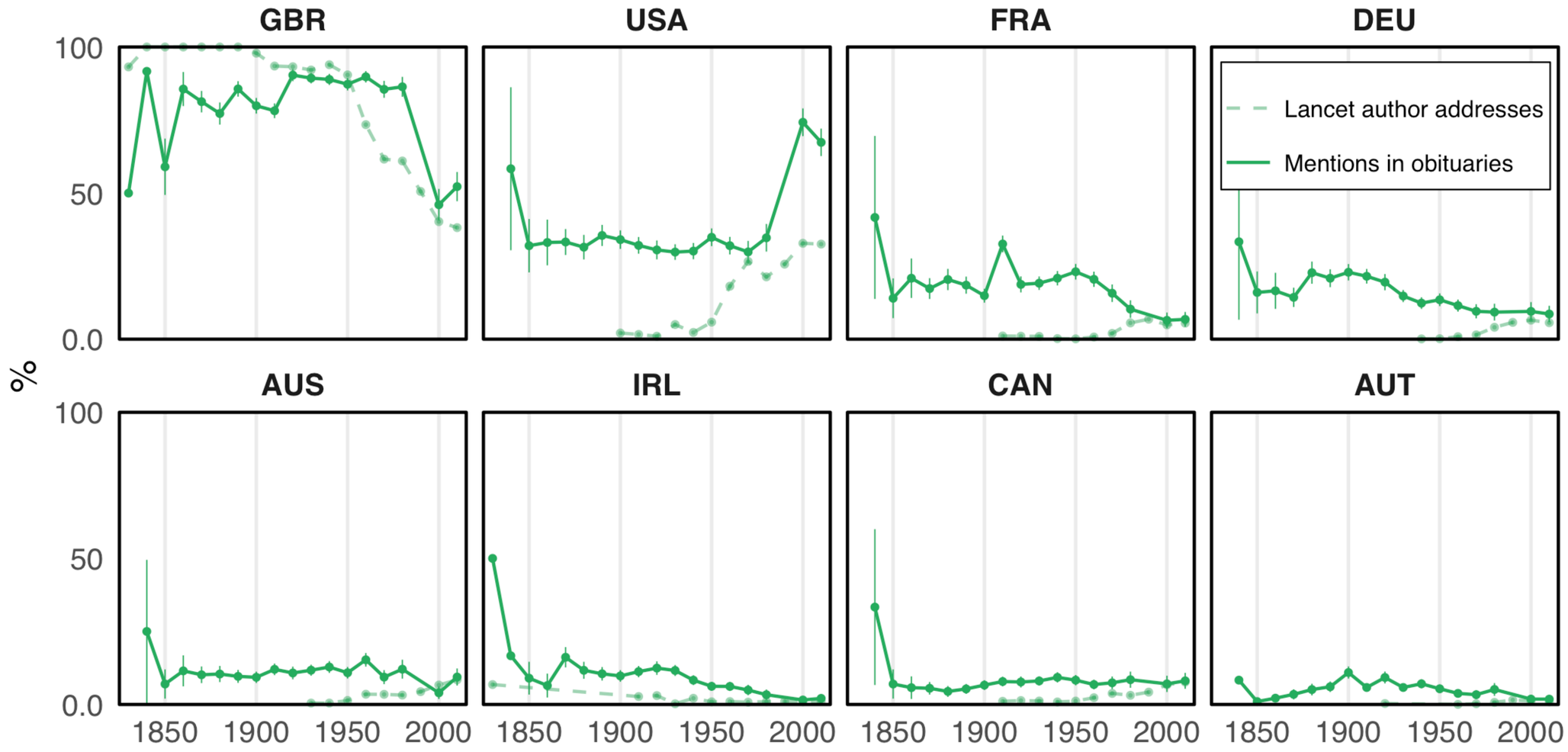
Chart the history of the *The Lancet*

The word “killed” appearing in the obituary



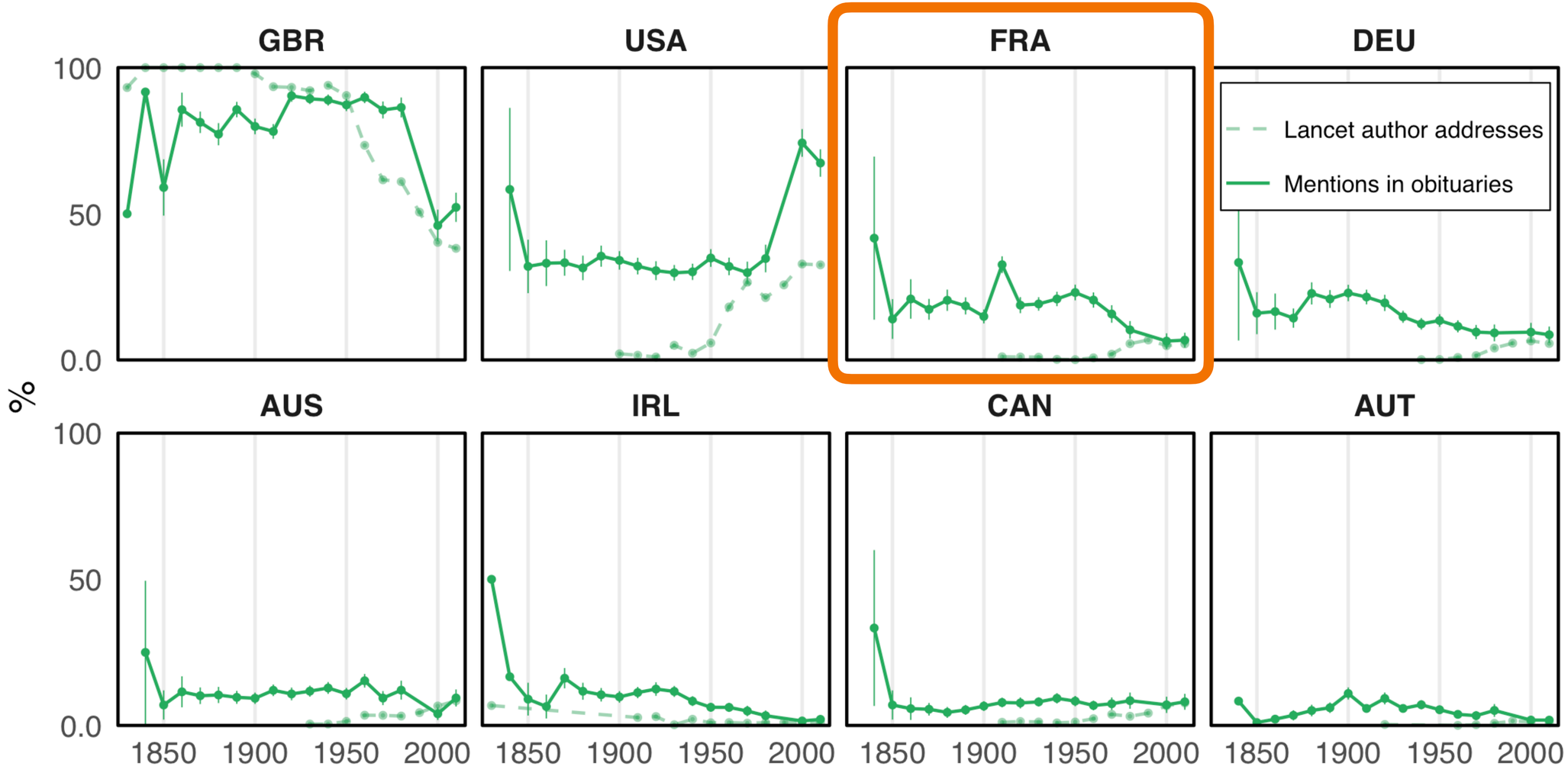
British journal, British authors

Placenames extracted from obituary text, affiliations from Lancet papers



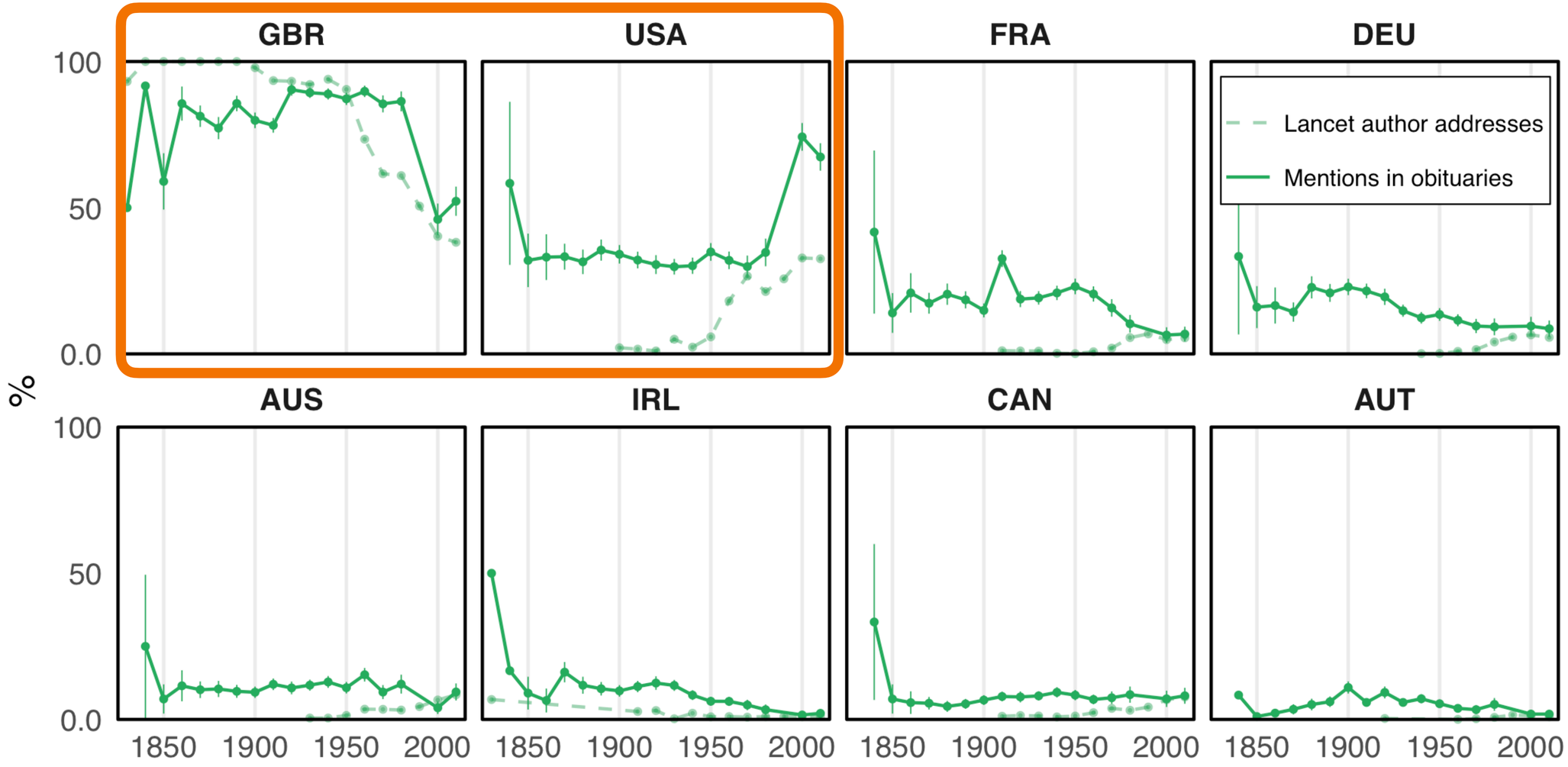
Uptick of French place names in WW1

In the 1910s



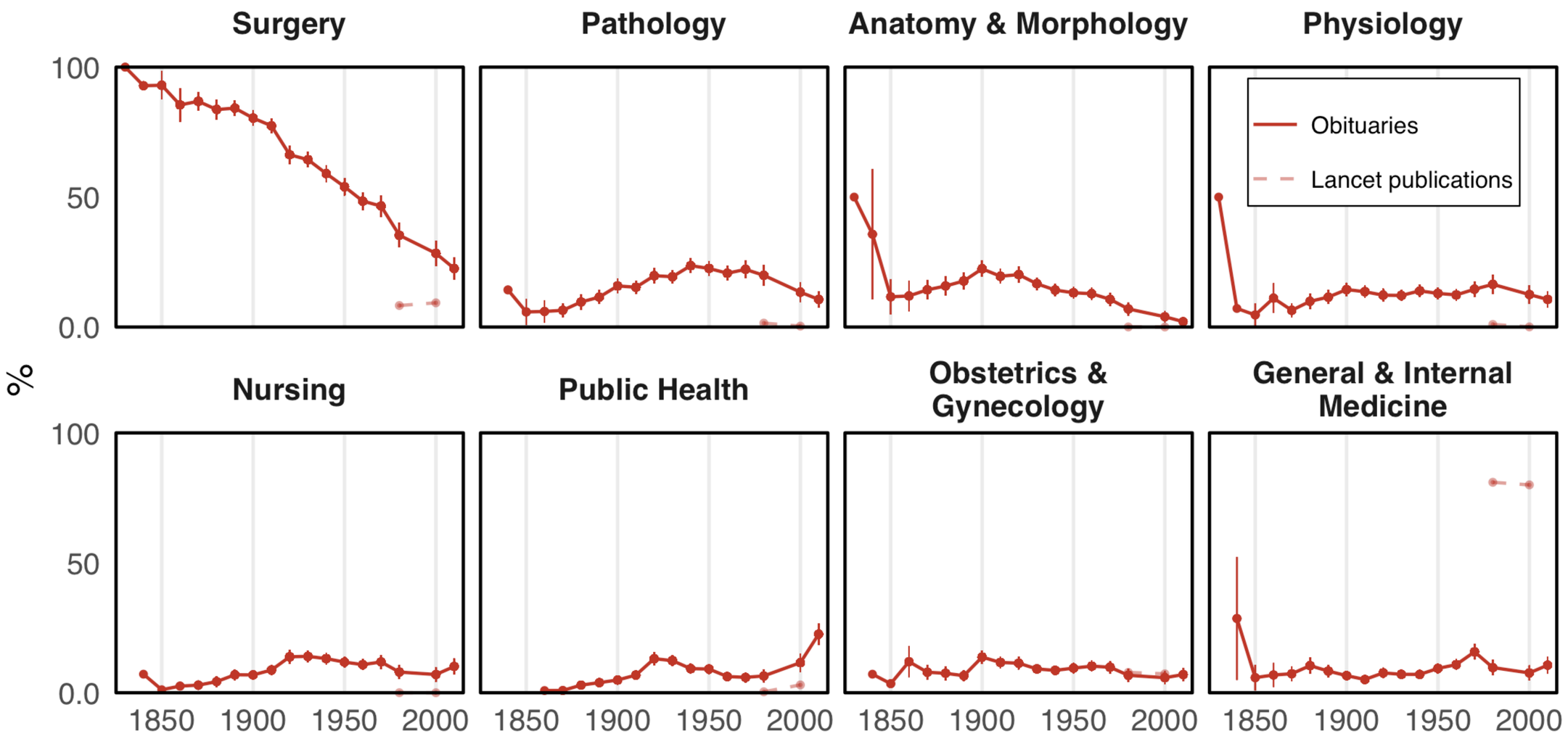
An international shift (well…moslty U.S.) after Elsevier's buy

Between 1980s, and 2000s



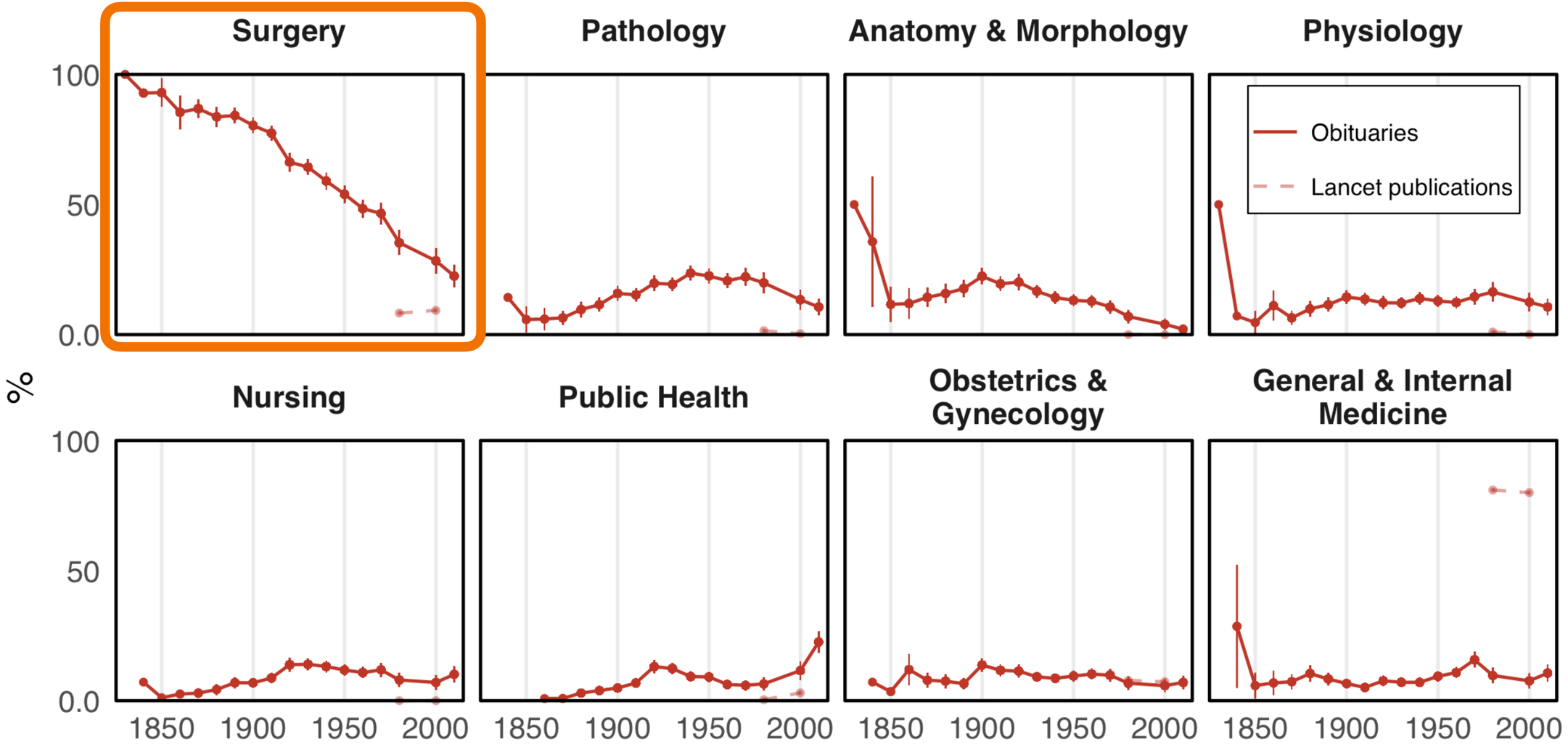
The evolving specialization of the profession

Extracting terms like “Surgeon” and “Pathologist” from text



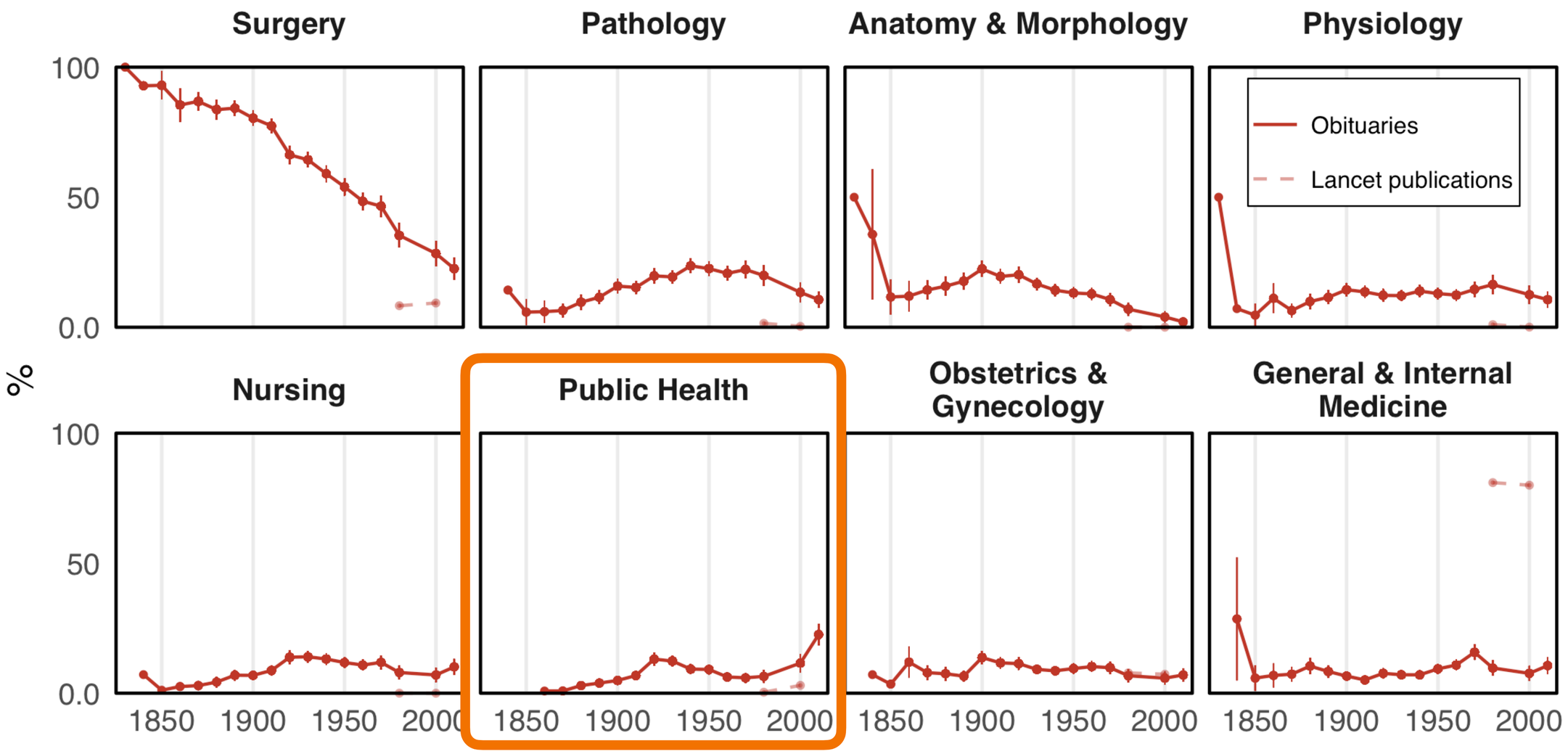
Early on, everyone was a surgeon

This changed as the field became more specialized



More Public Health obituaries in 1940s, 2000s

Why? ideas?



Appendix — misc

Improving evaluation in science

Diversity

Get more people involved

- Differing perspectives mitigate latent biases
- Provide unique important information during evaluation
- It also makes science stronger
- Has improved, but still needs work

Underrepresented minorities as a percent of various U.S. faculty appointment types

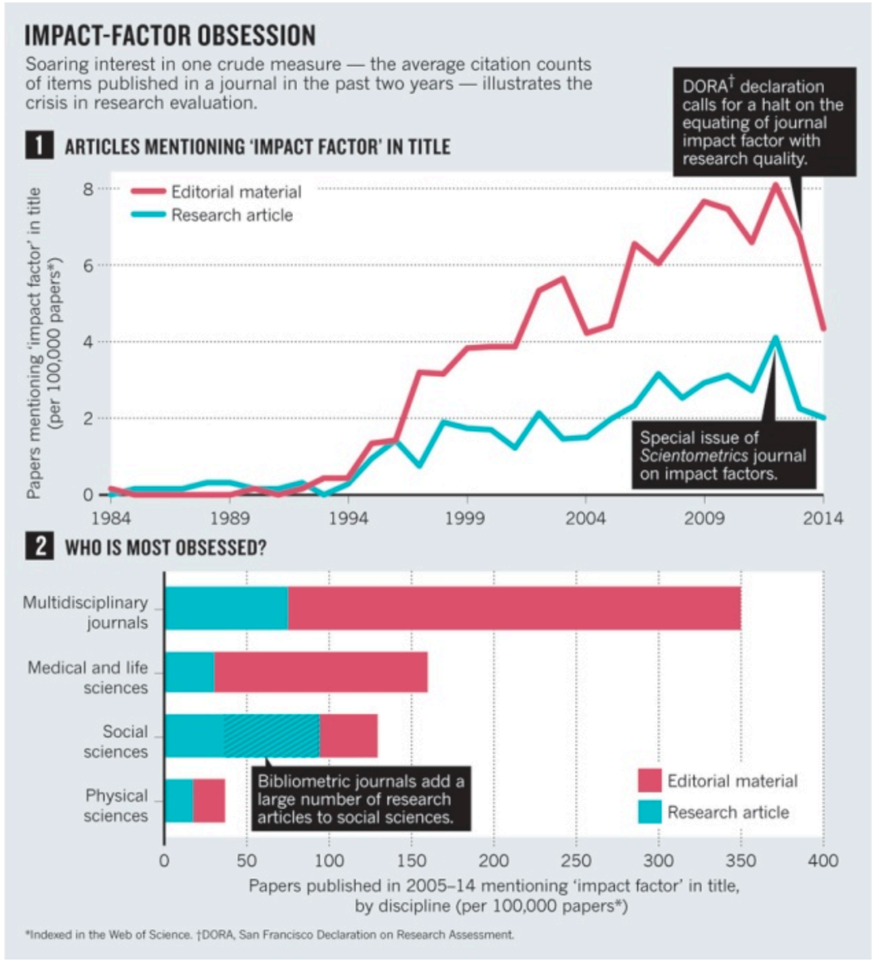
	1993	2003	2013
All faculty	8.6%	10.5%	12.7%
Full-time faculty	8.2%	9.9%	11.1%
Tenured	7.1%	9.2%	10.2%
Tenure Track	10.0%	11.3%	11.7%
Nontenure Track	9.1%	10.0%	12.0%
Part-time faculty	9.2%	11.2%	14.2%

Finkelstein, M., Conley, V. M., & Schuster, J. H. (2016). *Taking the Measure of Faculty Diversity*. TIAA Institure.

Responsible evaluation

Stop being so obsessed with numbers

- Researchers, journals, editors, and administrators overuse metrics like the JIF
- Metrics are associated with objectivity (often wrongly)
- Promote awareness of issues
- Adhere to best principles
  - Leiden Manifesto

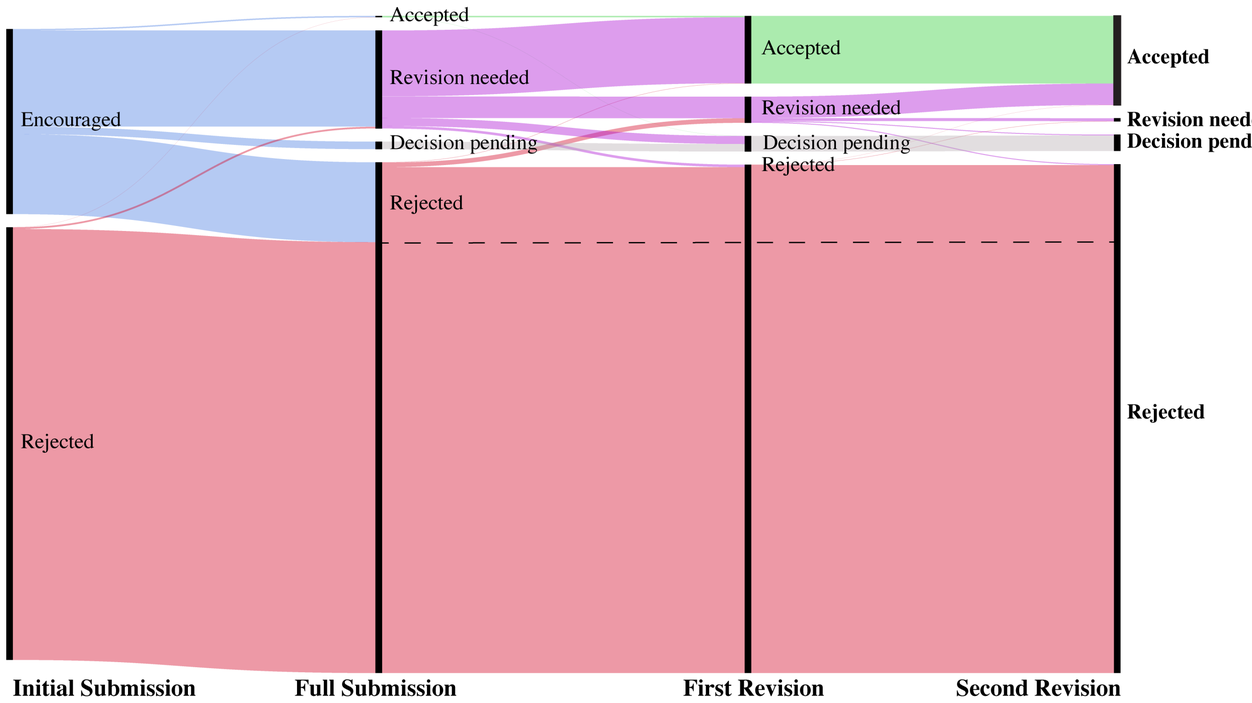


Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News*, *520*(7548), 429.

Openness

Making the process open to scrutiny

- A key tenet of science is openness
- Promotes reproducibility, transparency, trust
- Data on scholarly activity should also be open
- Allows researchers to interrogate the scientific process, identify problems, and propose solutions



Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.