Slides at:
dakotamurray.me/2020-northeastern/

Scientific success in context

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

SCHOOL OF
INFORMATICS, COMPUTING,
AND ENGINEERING

The first law of success:

“Performance drives success but when performance can’t be measured, networks drive success”

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

The first law of success:

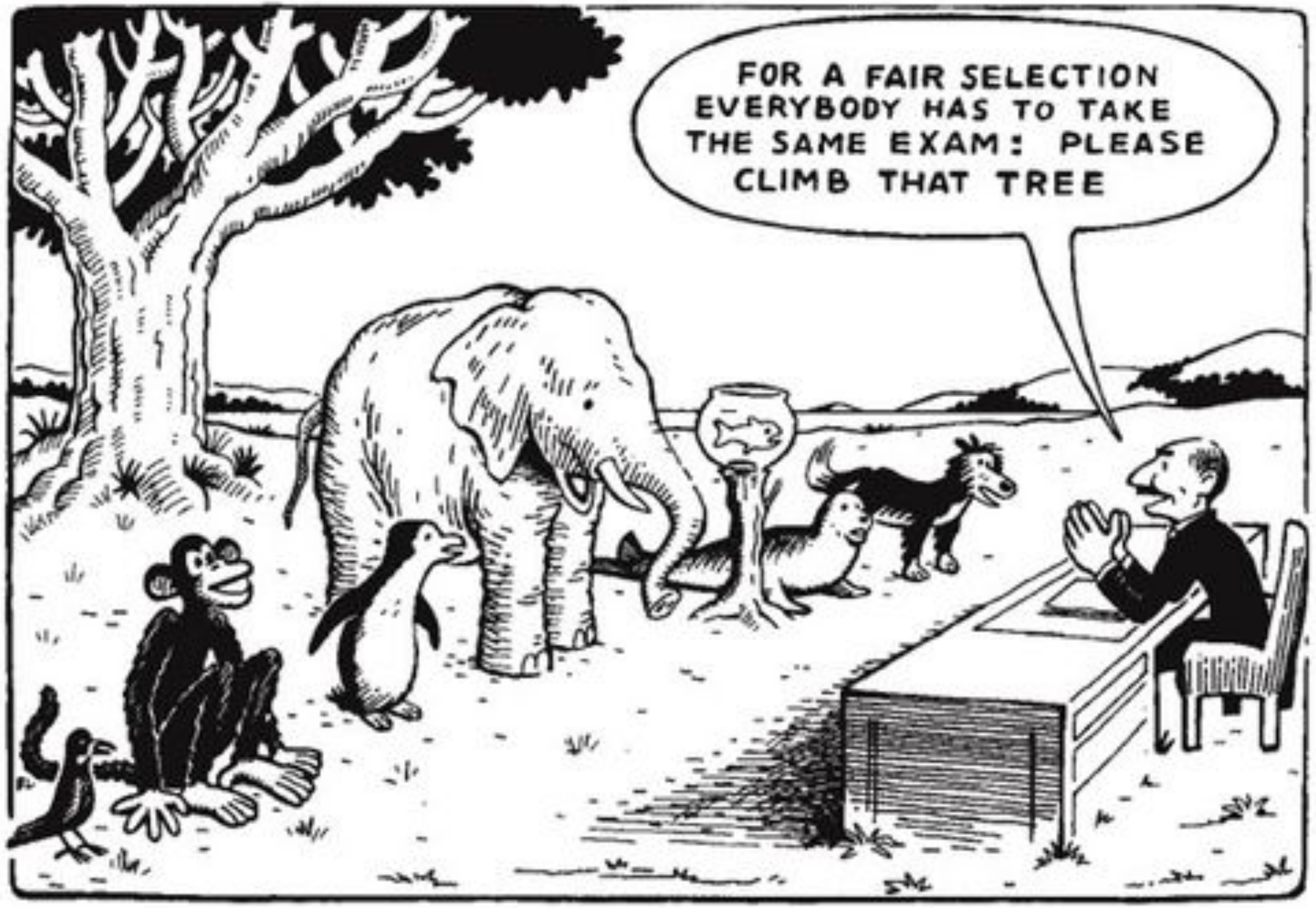
"Performance drives success but when performance can't be measured, networks drive success"

Success isn't about you, its about us

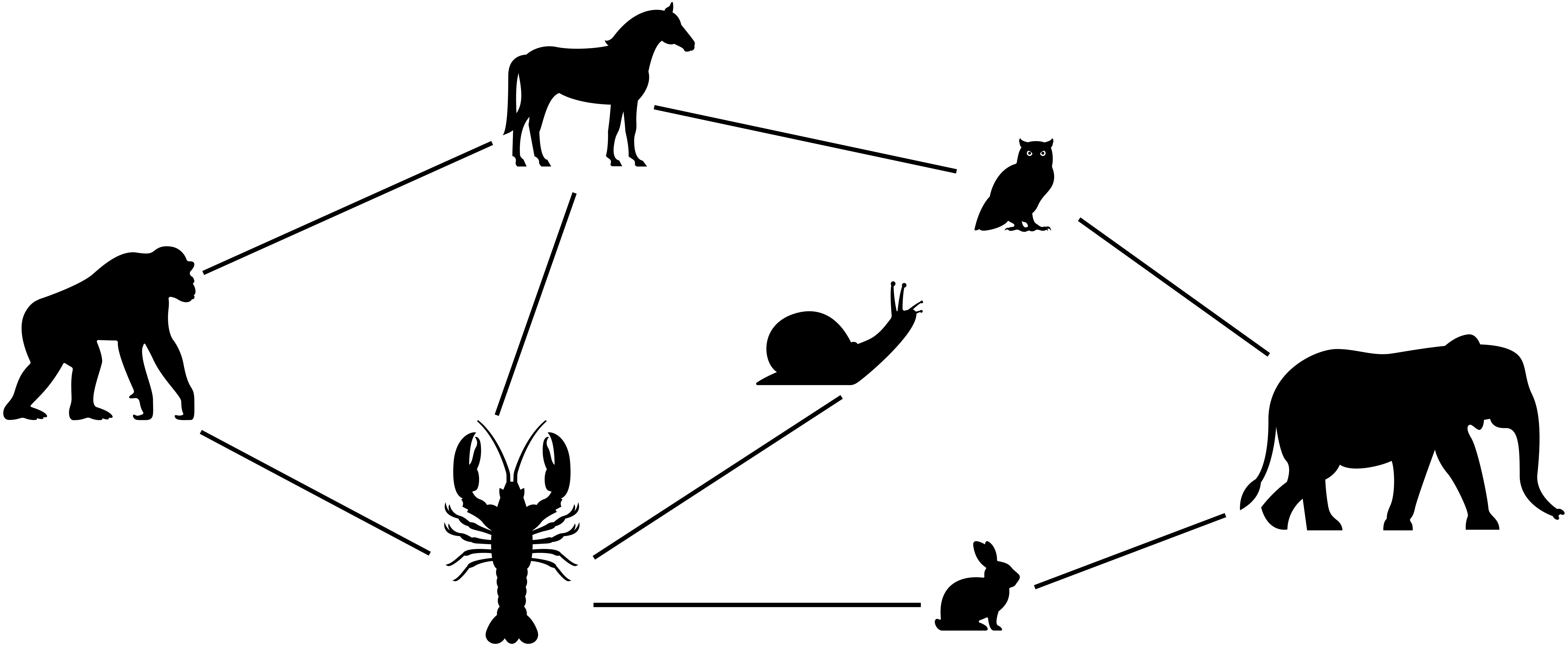
Barabási, A.-L. (2018). *The Formula: The Universal Laws of Success*. Little, Brown and Company.

What does it mean to perform well in science?

Evaluation without context is flawed

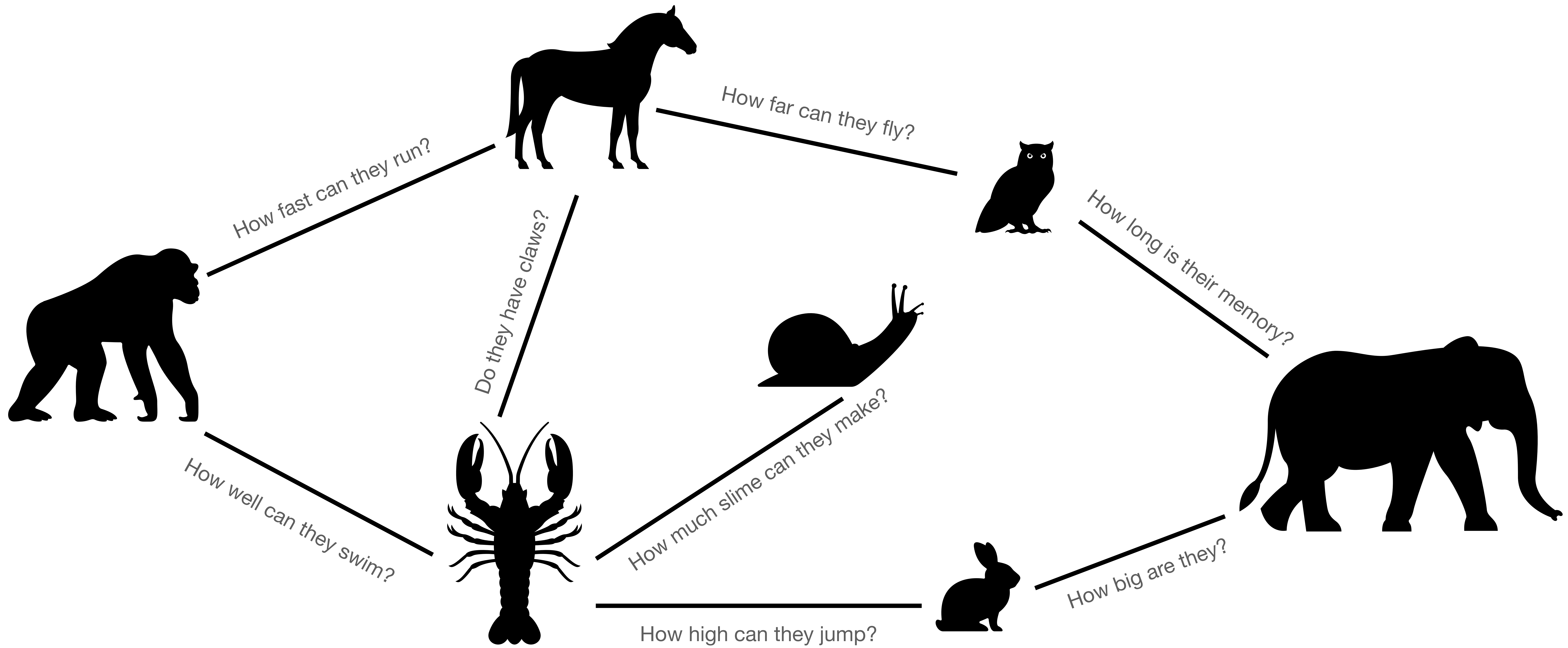


Networks matter!



Networks matter!

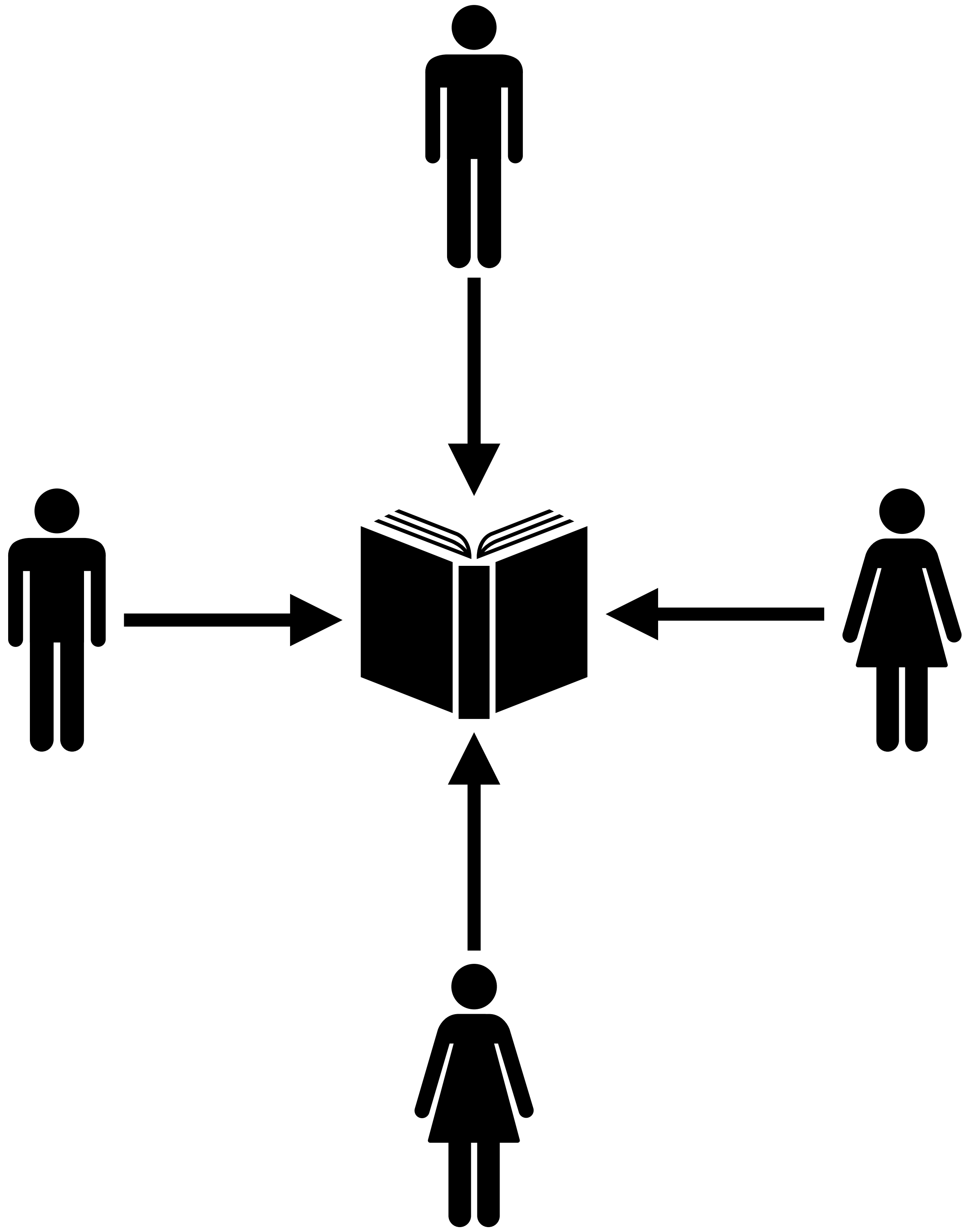
But its even more complicated



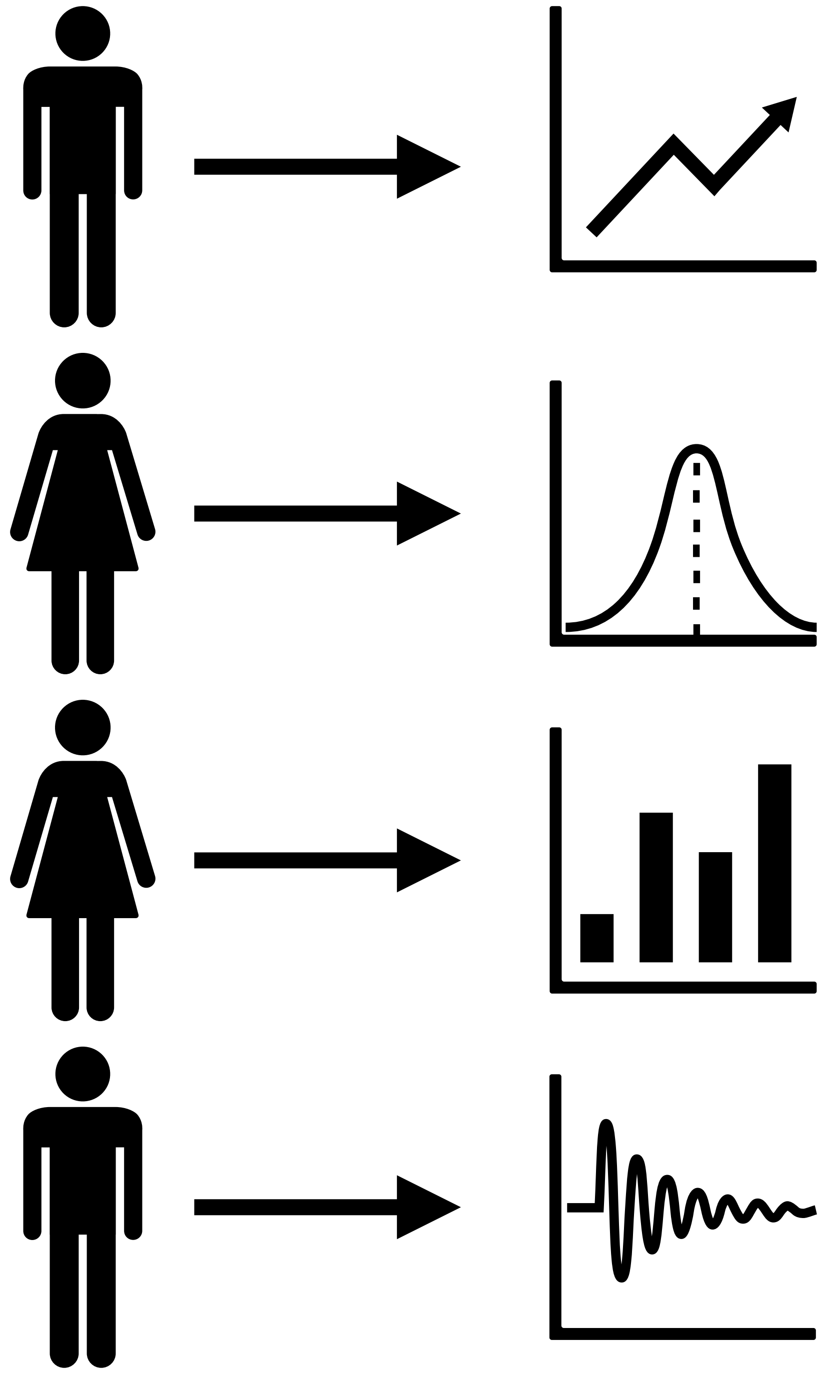
Evaluation in science

Hiring, promotion, publication, funding, and reputation

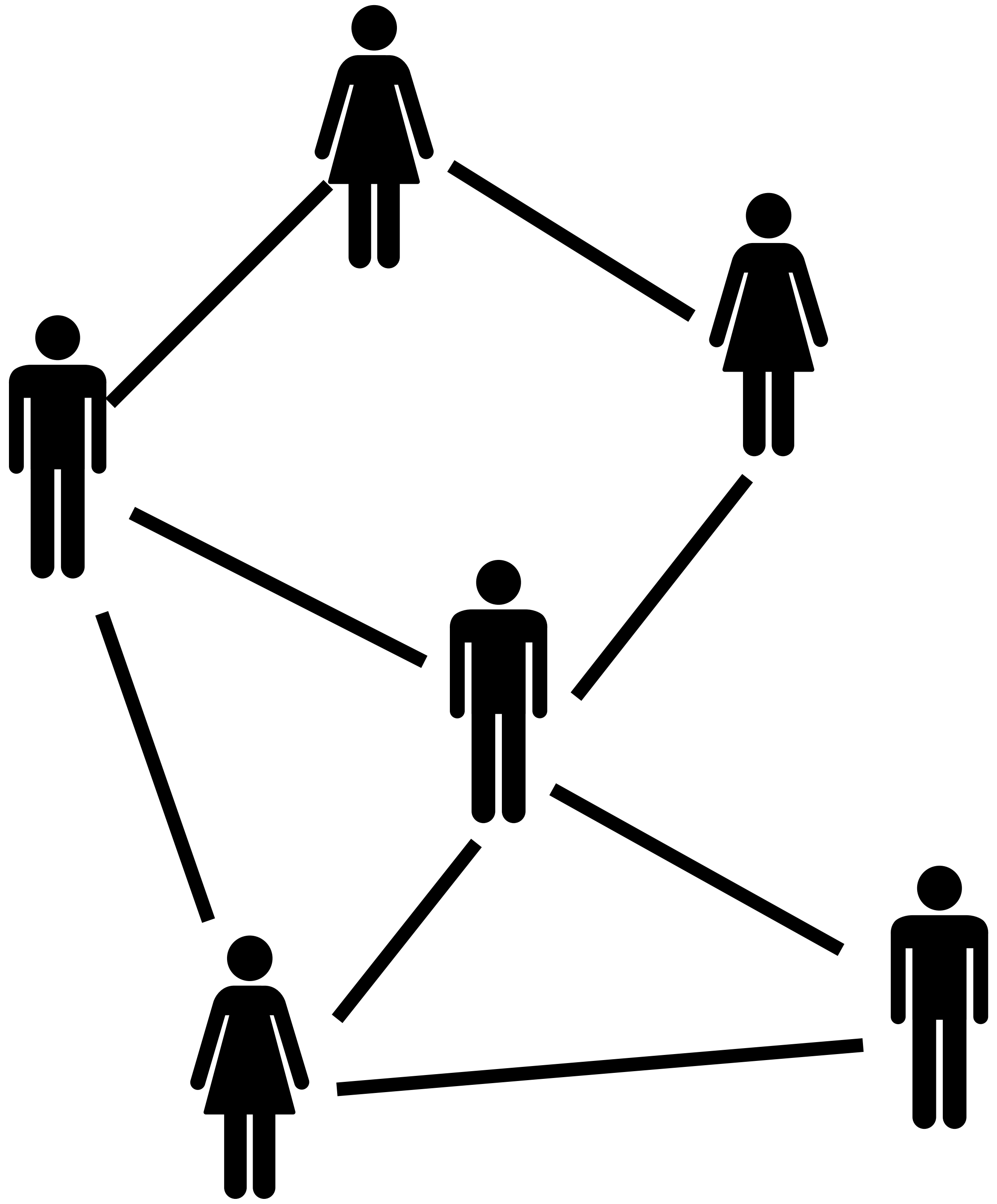
Peer review



Performance metrics



Peer-to-peer judgements
(networks)



Ideally, evaluation should capture true merit

Success is about us

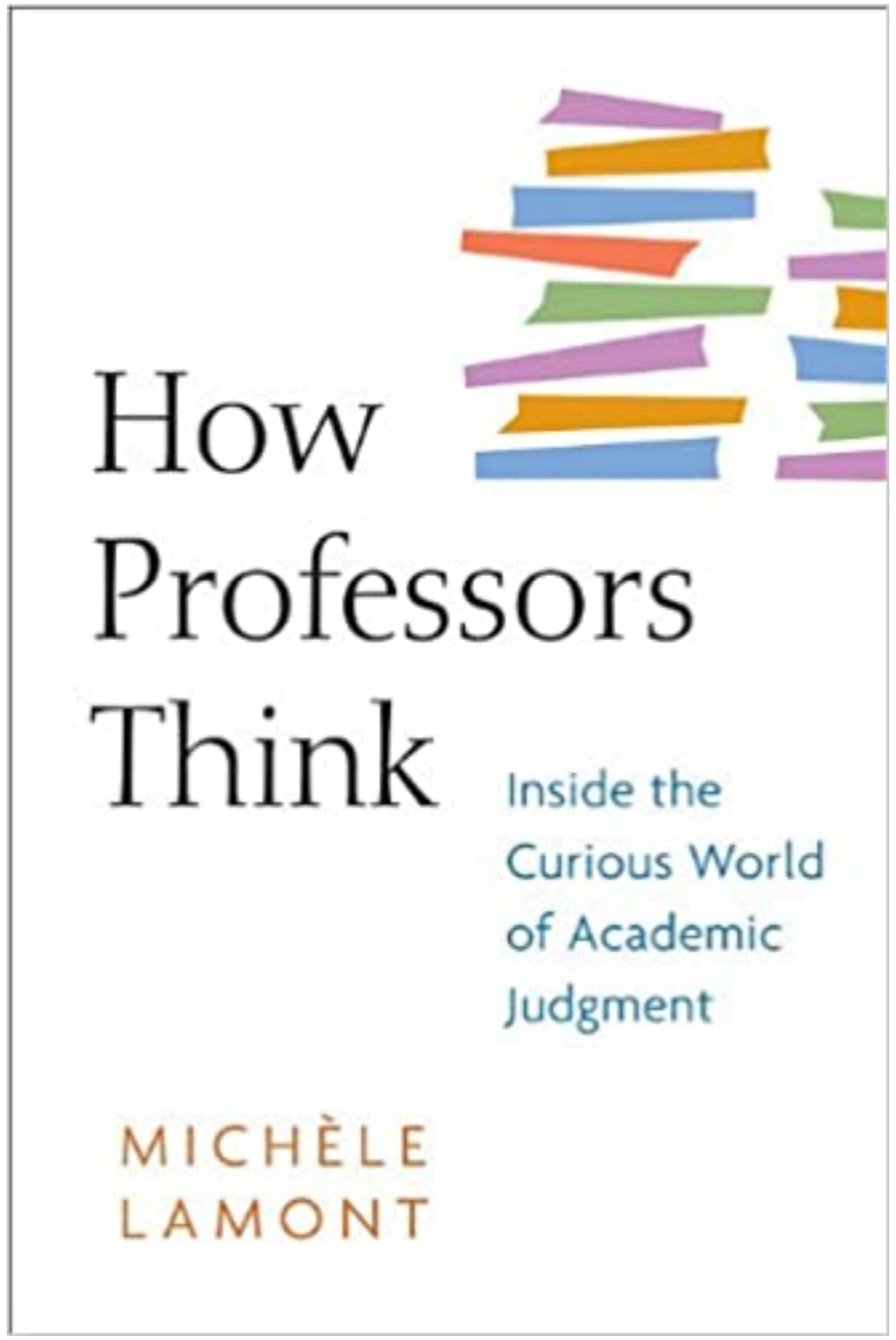
But we are embedded in a social, political, and cultural context

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)*

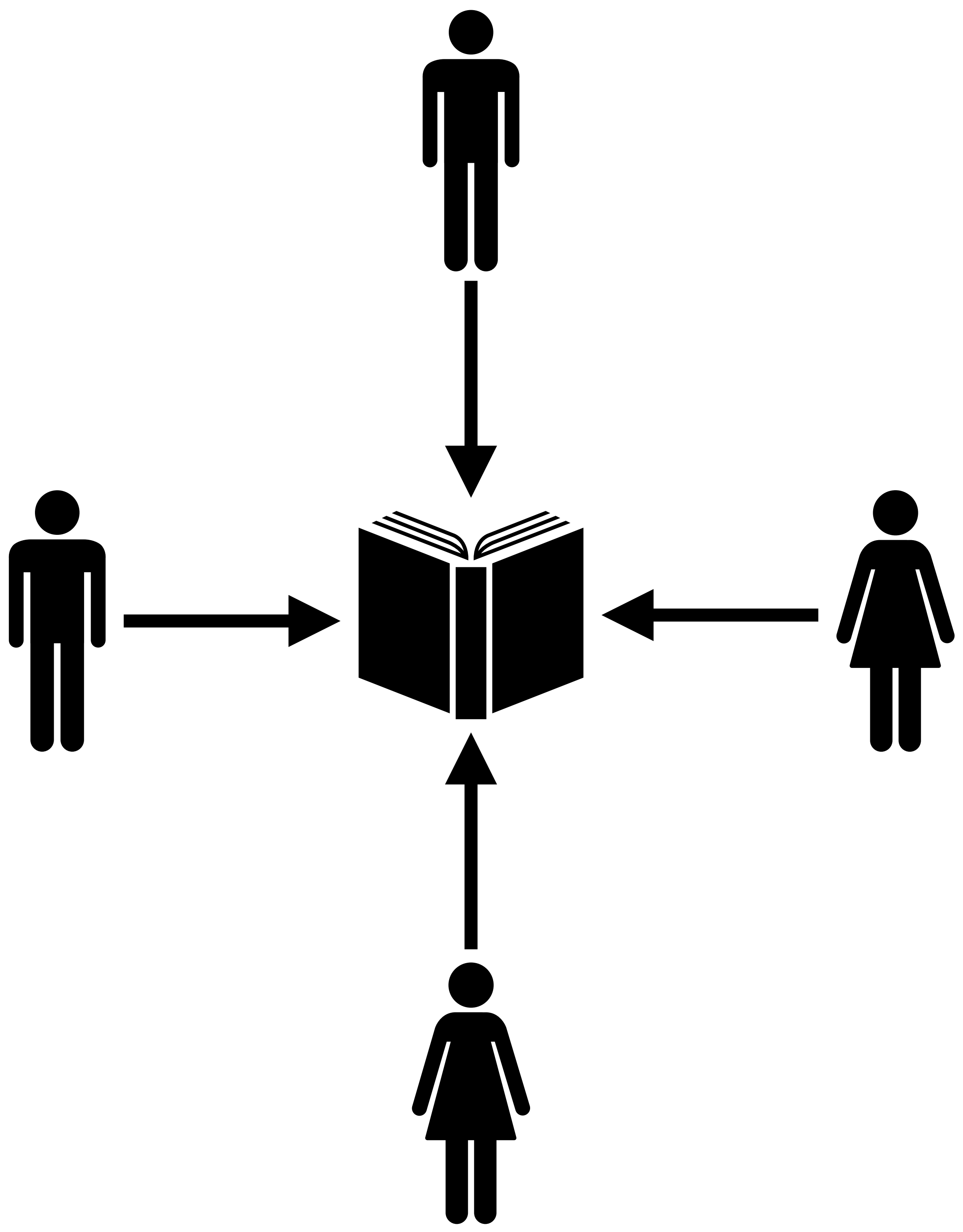


Lamont, M. (2009). *How Professors Think: Inside the Curious World of Academic Judgment*. Harvard University Press.

My research direction:

Investigate the contextual factors that drive scientific evaluation and success

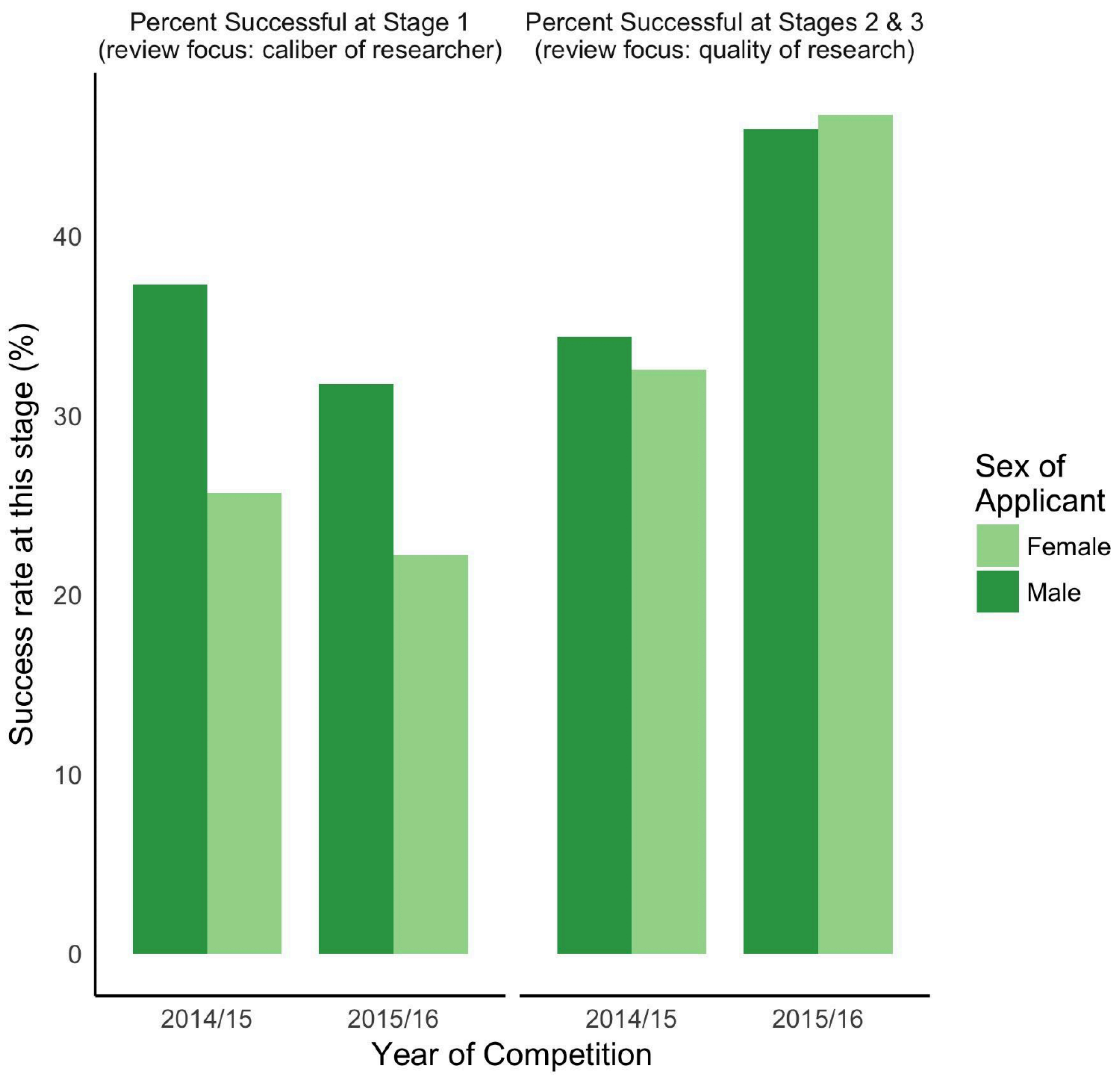
Peer review



Gender bias

Grant peer review at CIHR

- 2-stage review process

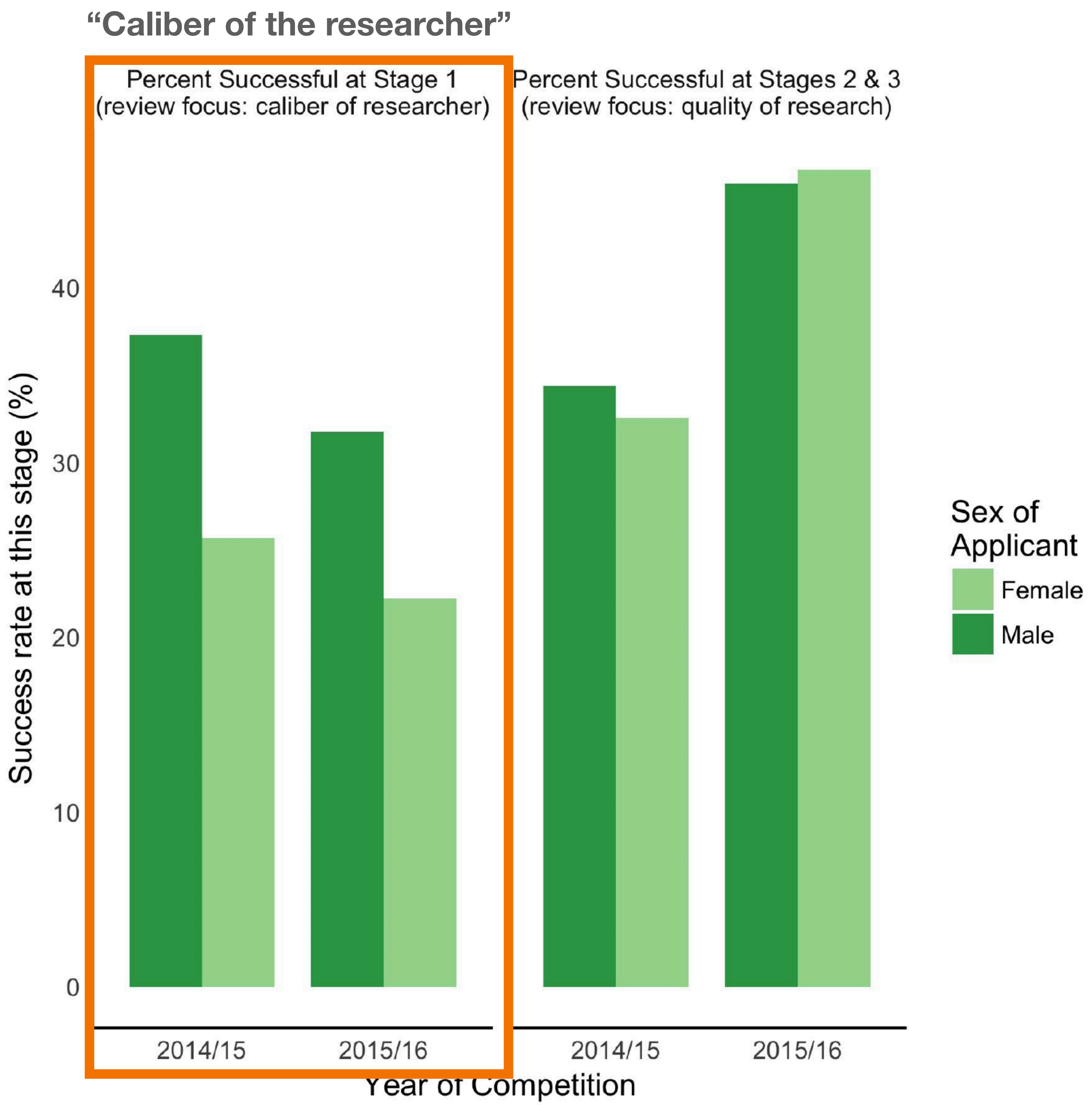


Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2017). Female grant applicants are equally successful when peer reviewers assess the science, but not when they assess the scientist. *BioRxiv*, 232868.

Gender bias

Grant peer review at CIHR

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

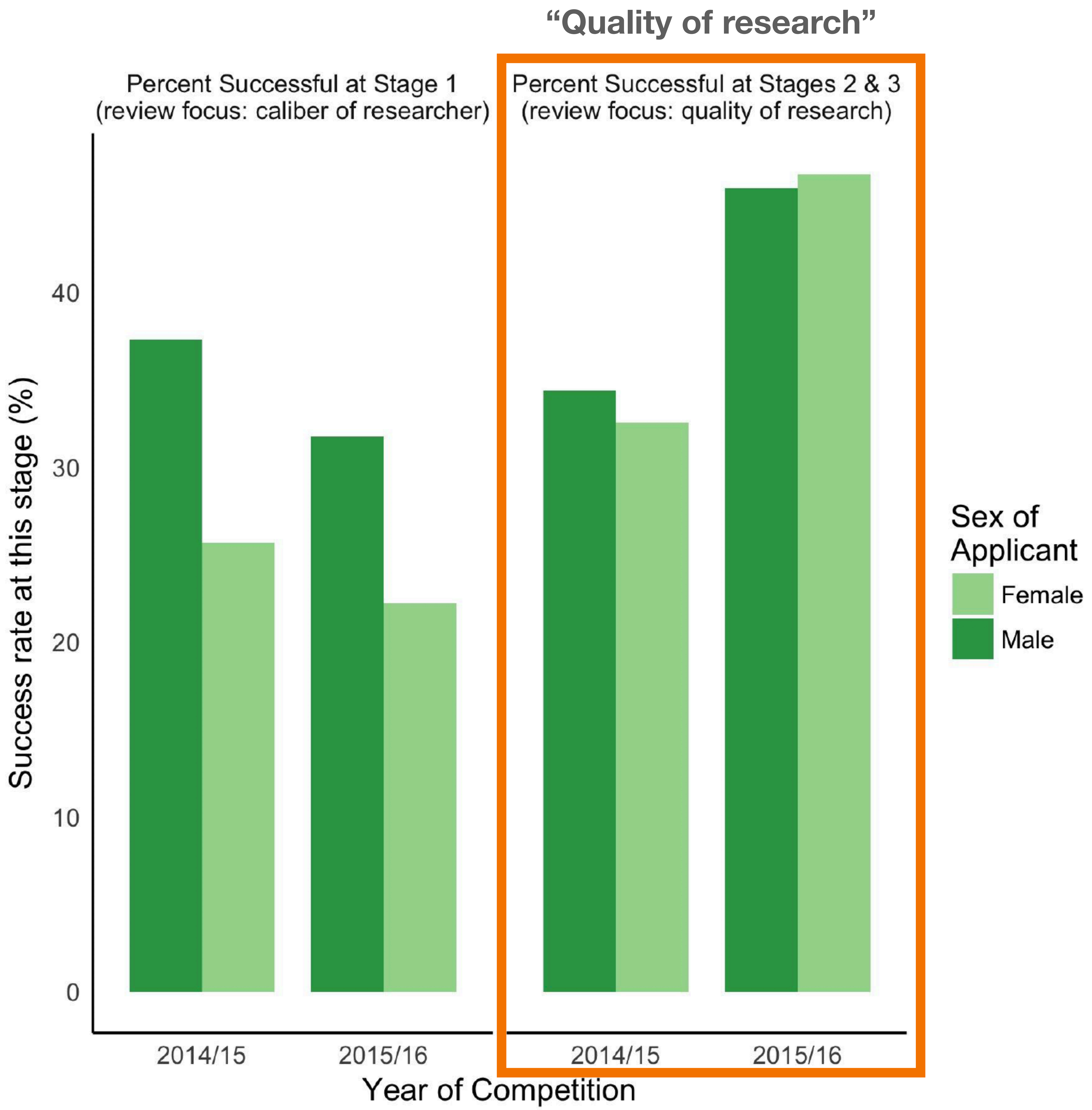


Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2017). Female grant applicants are equally successful when peer reviewers assess the science, but not when they assess the scientist. *BioRxiv*, 232868.

Gender bias

Grant peer review at CIHR

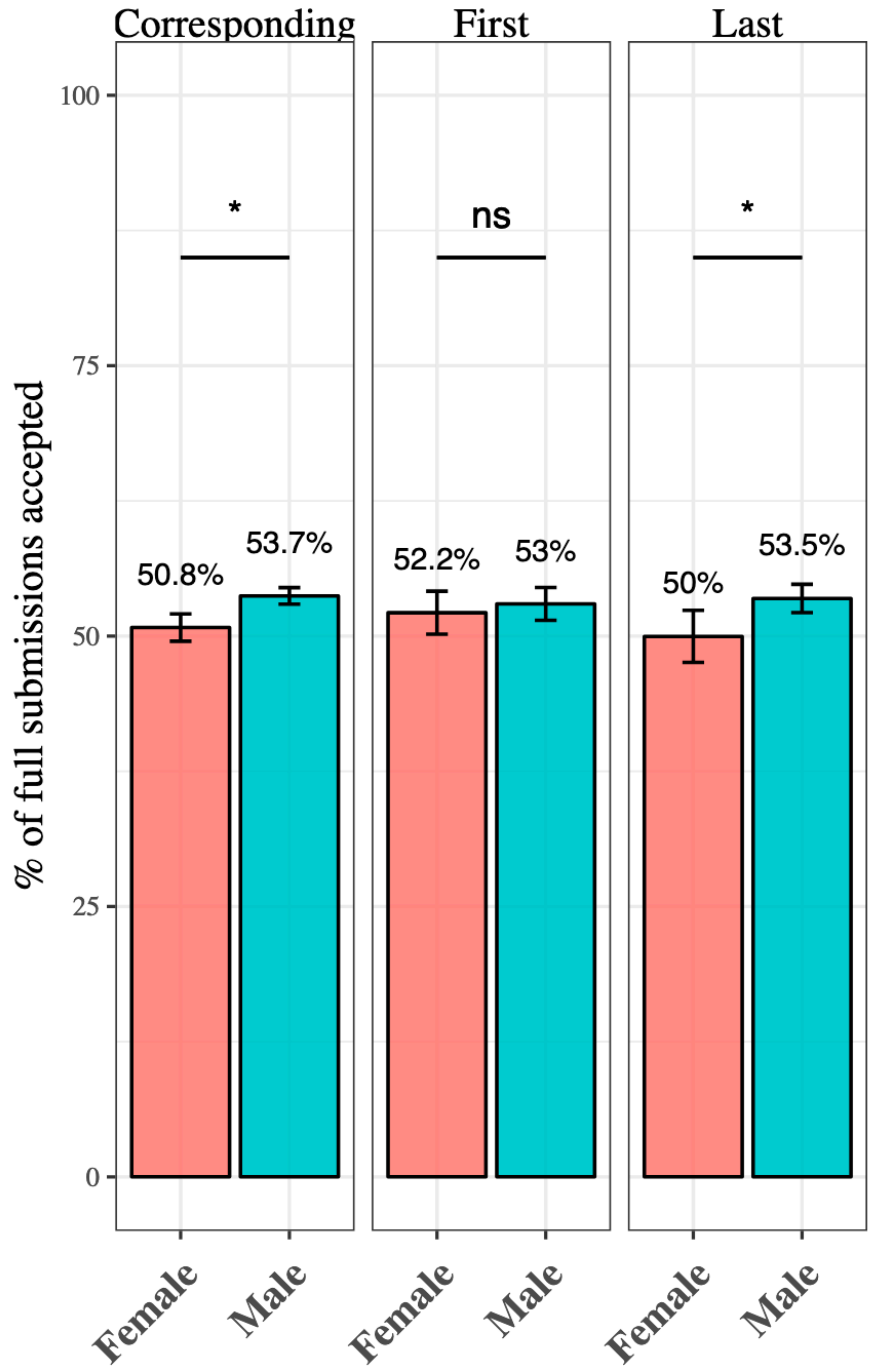
- 2-stage review process
- Women rated lower in “Caliber of researcher”
- Gender equity when rating on “quality of researcher”



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.

Journal peer review at eLife

Women's papers accepted less

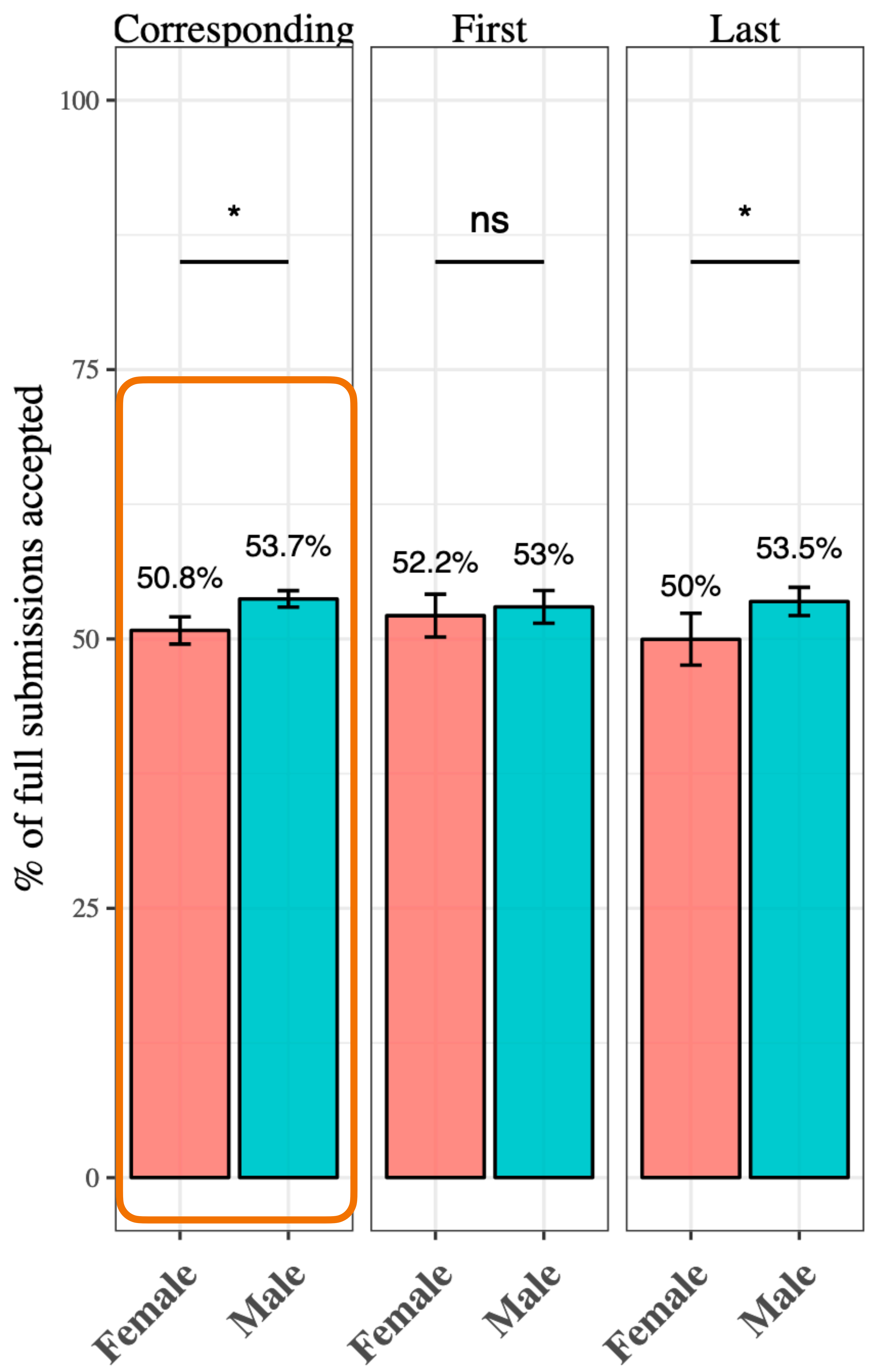


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

Women’s papers accepted less

- 2.9 % point gender disparity for corresponding authorship

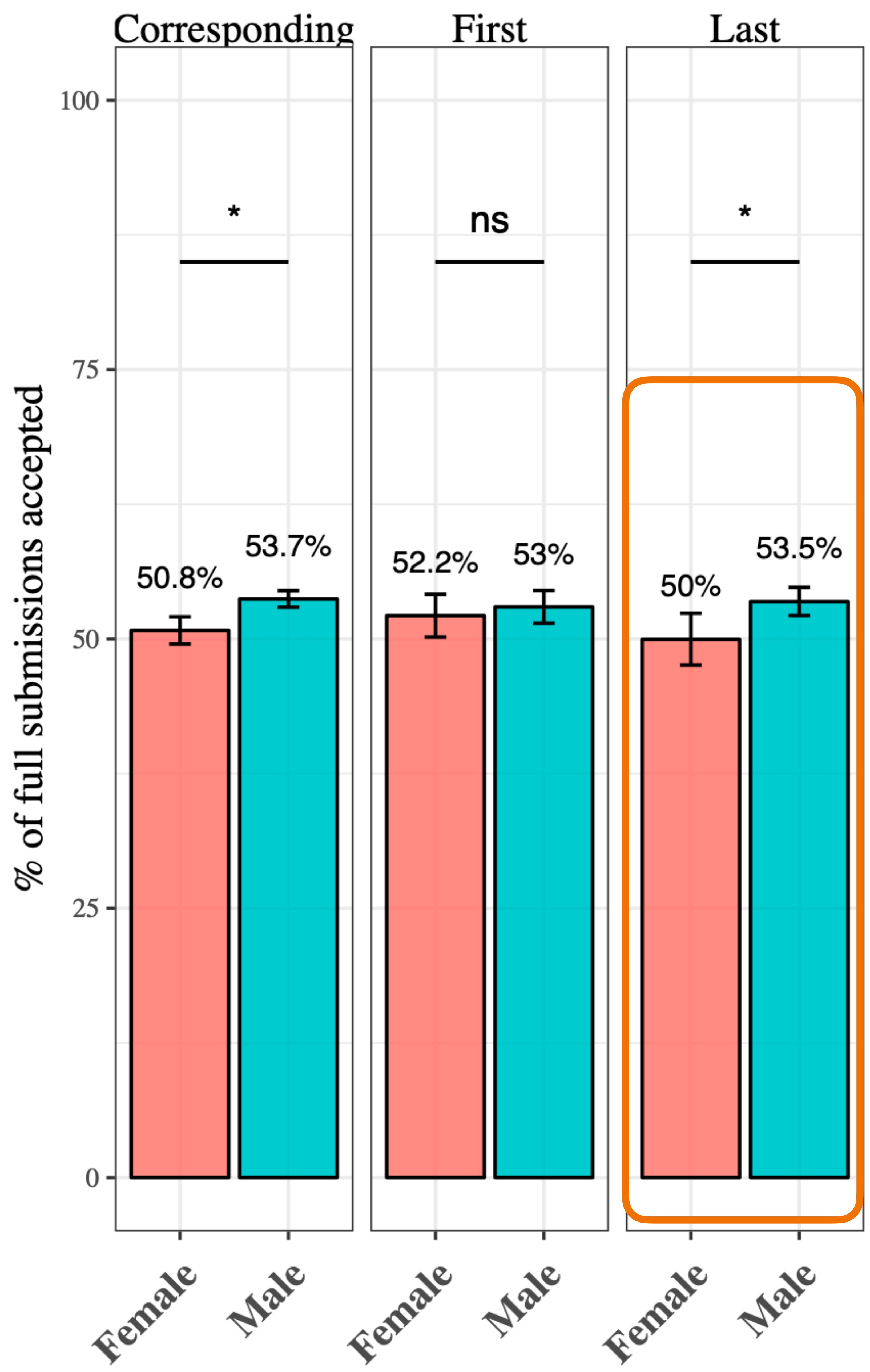


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

Women’s papers accepted less

- 2.9 % point gender disparity for corresponding authorship
- 3.9% among last authors

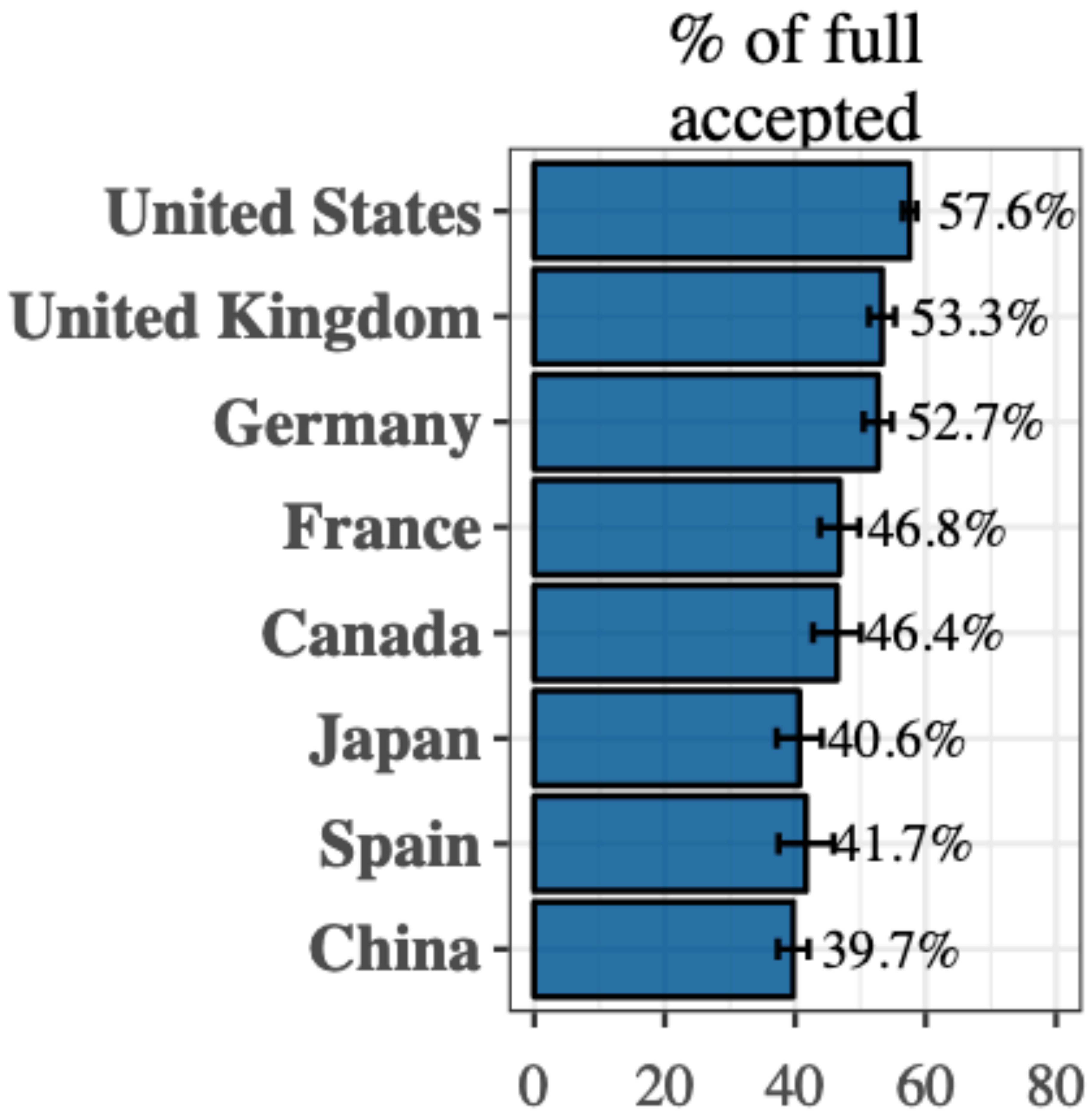


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

Journal peer review at eLife

U.S. papers accepted more

- Country of corresponding author

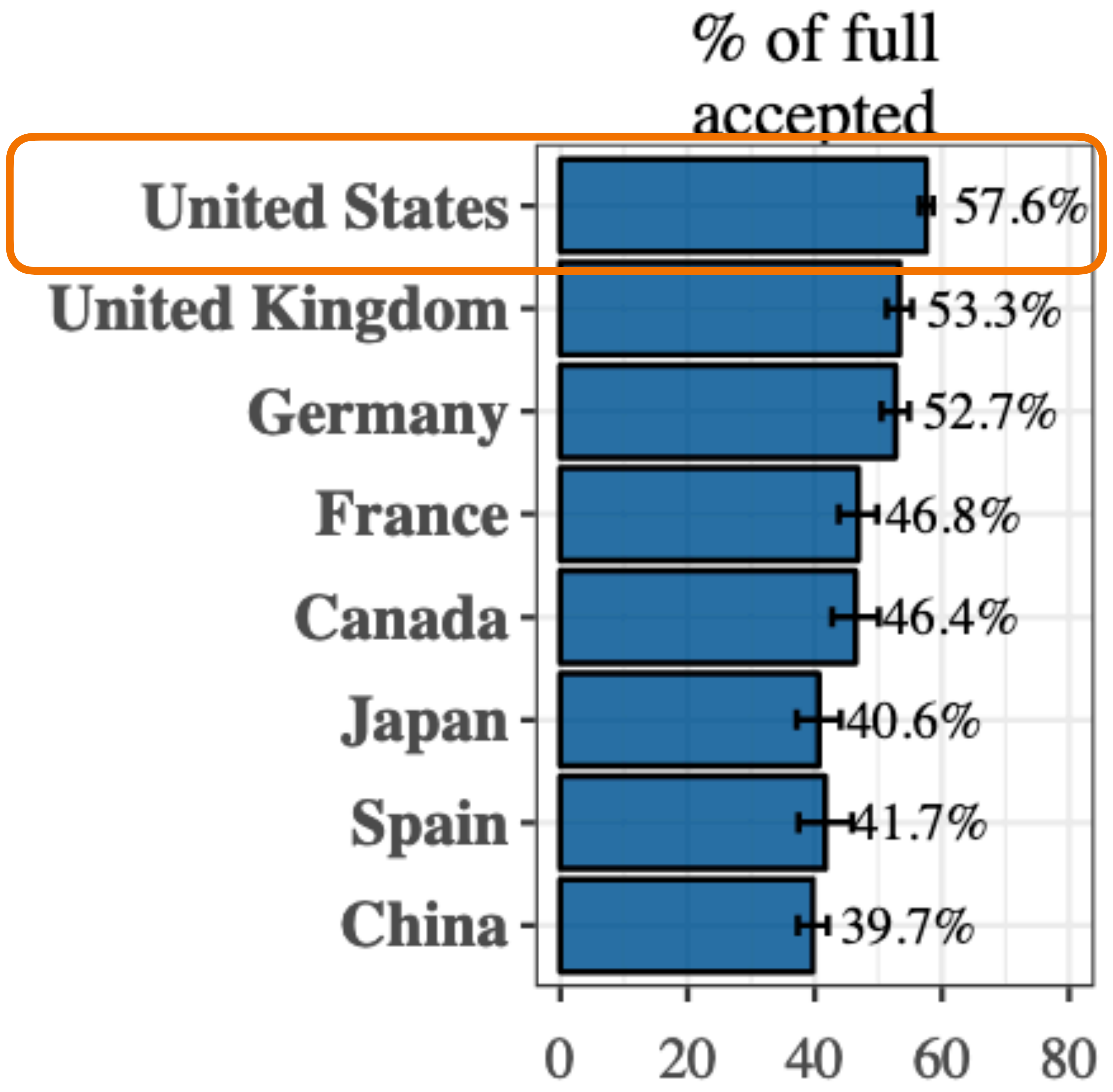


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

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

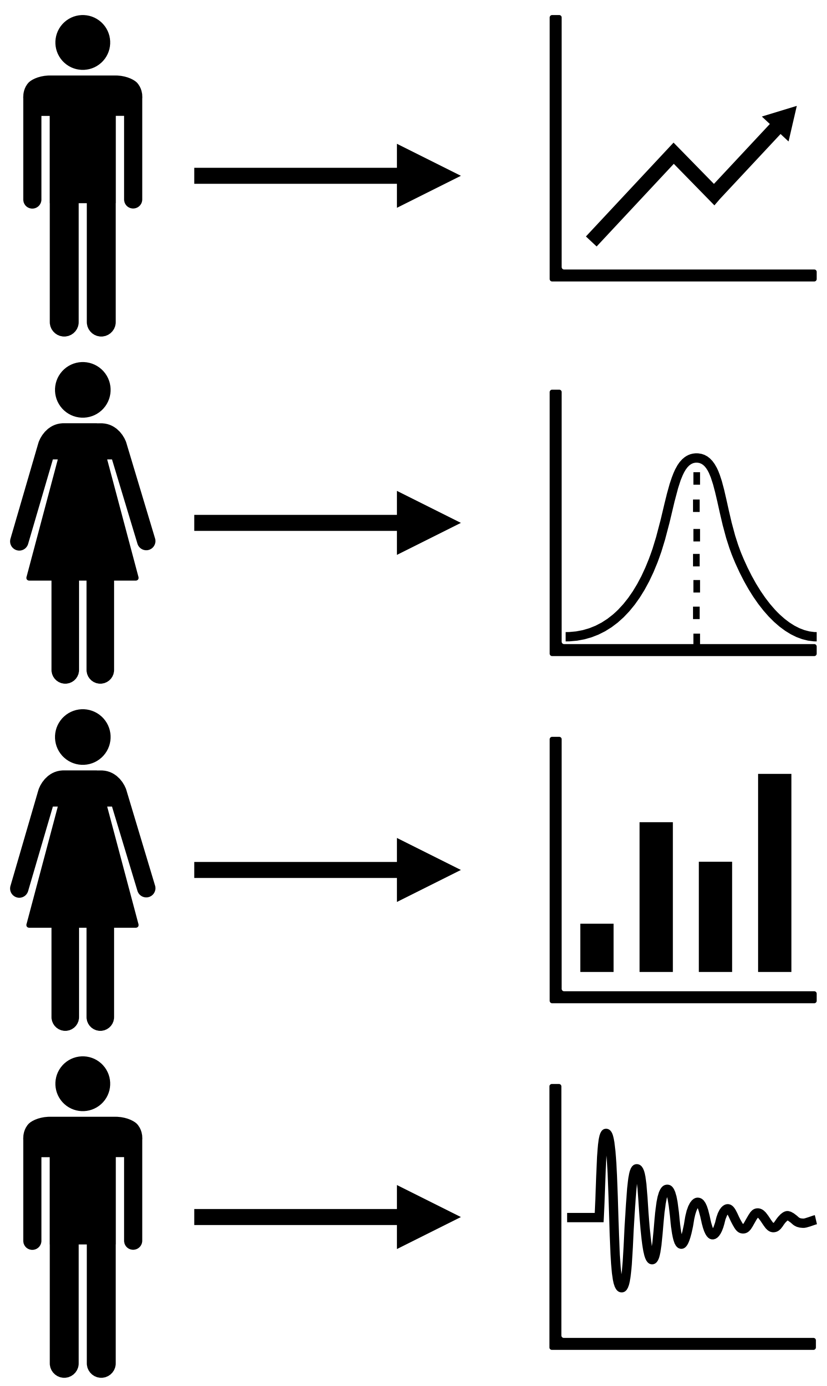


Murray, D., Siler, K., Larivière, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Author-Reviewer Homophily in Peer Review. *BioRxiv*, 400515.

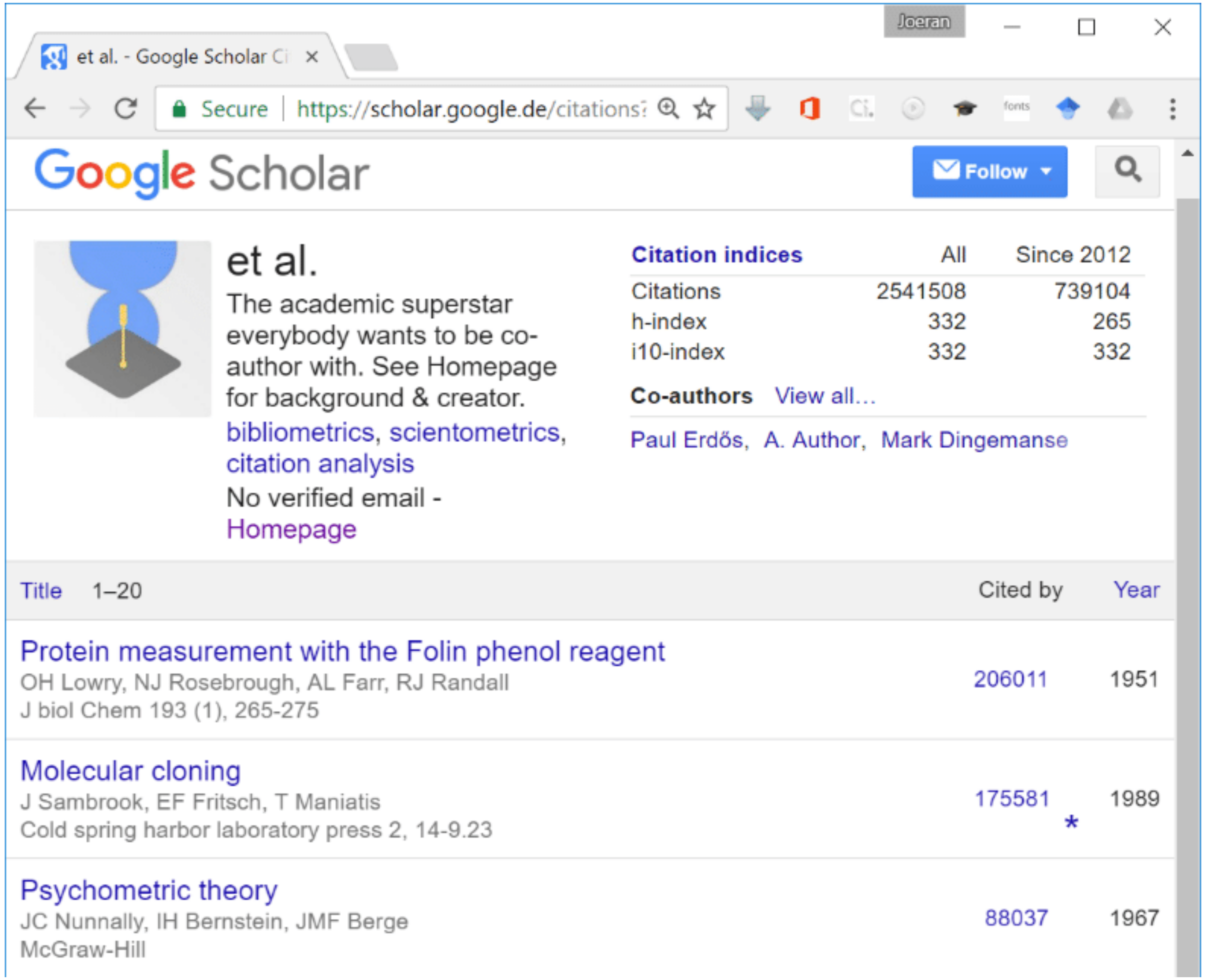
Peer review is subject to contextual factors

What about more “objective”, quantitative, metrics?

Performance metrics



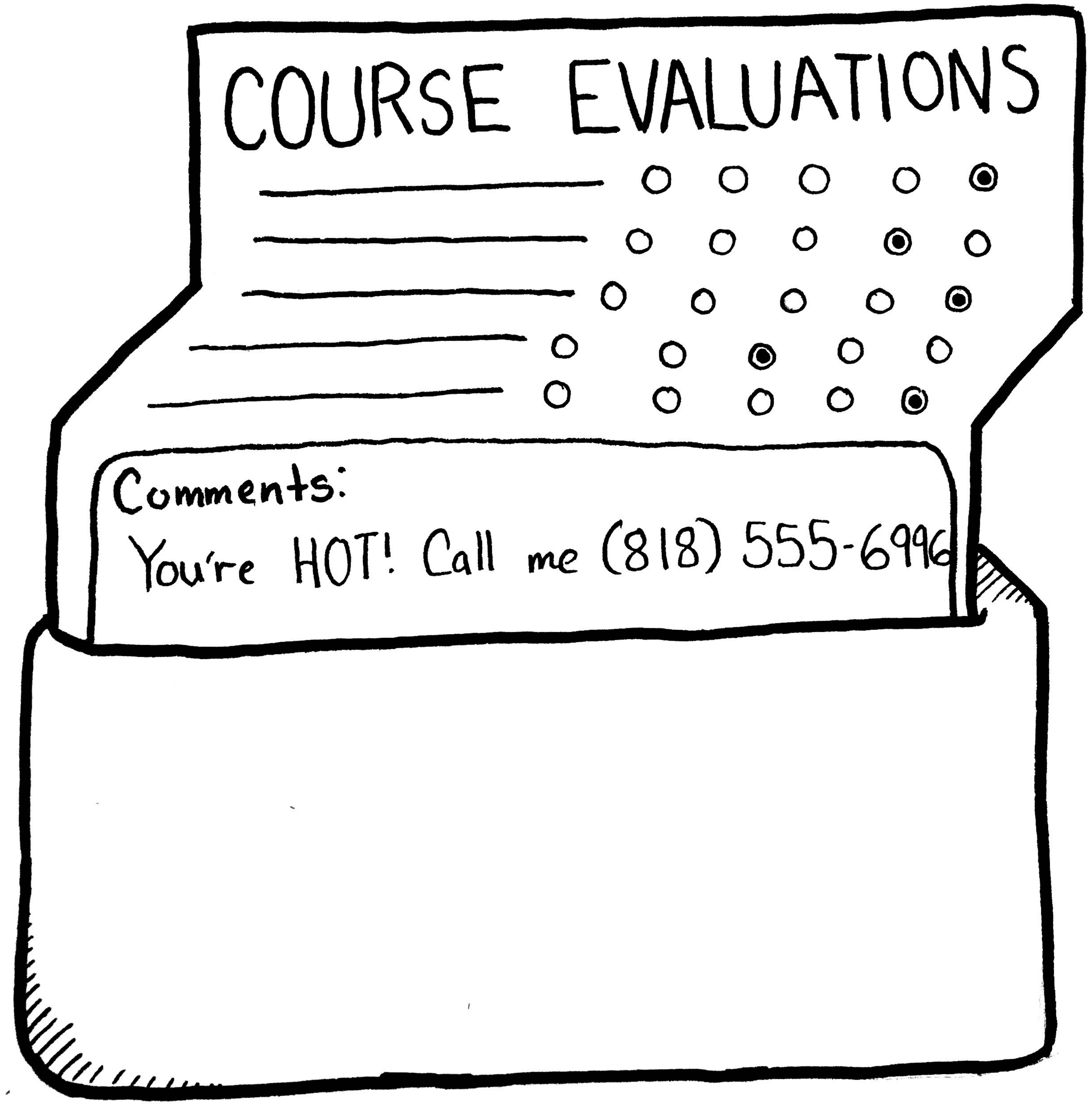
Metrics—far from ideal



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

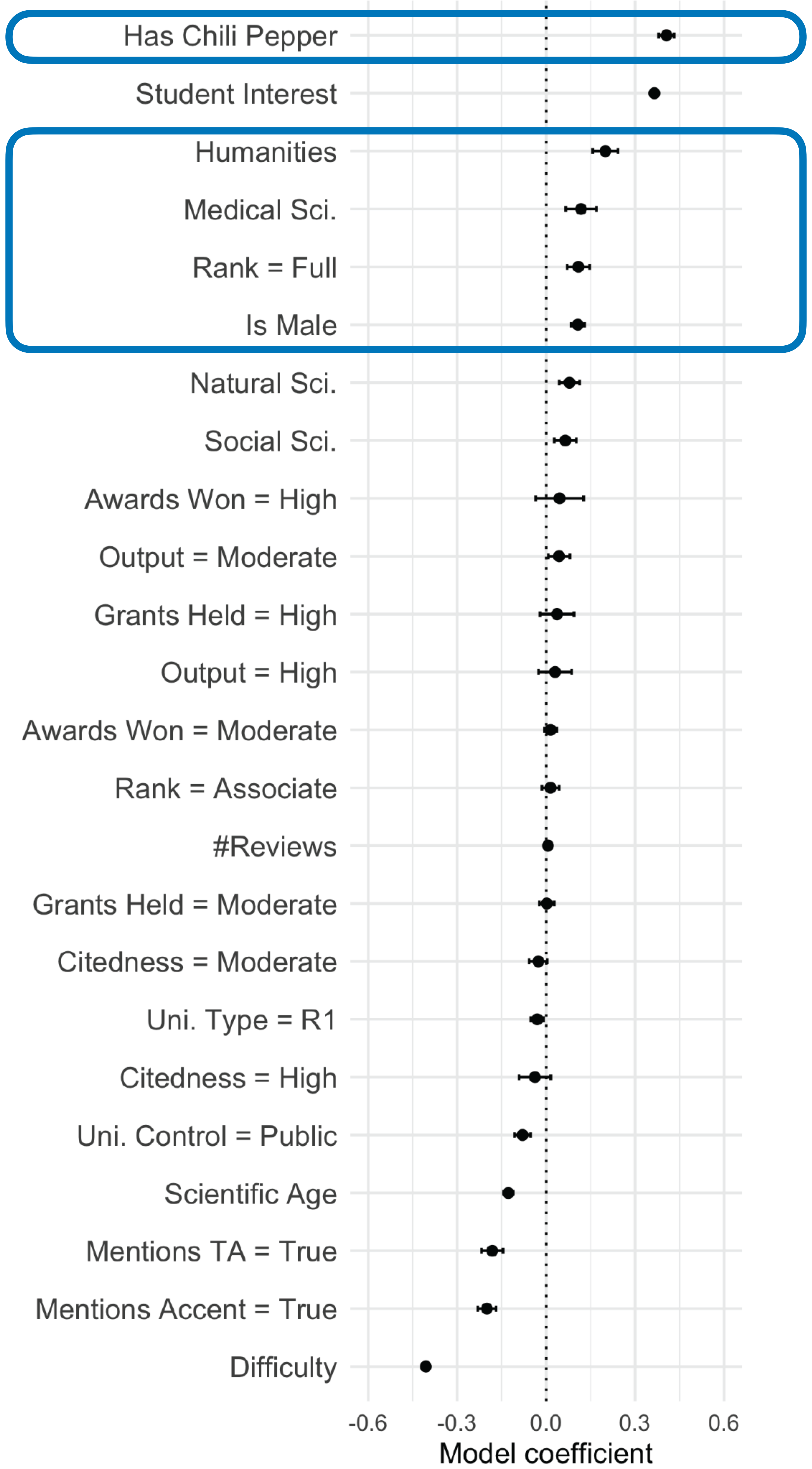
- Attractive profs rated higher

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

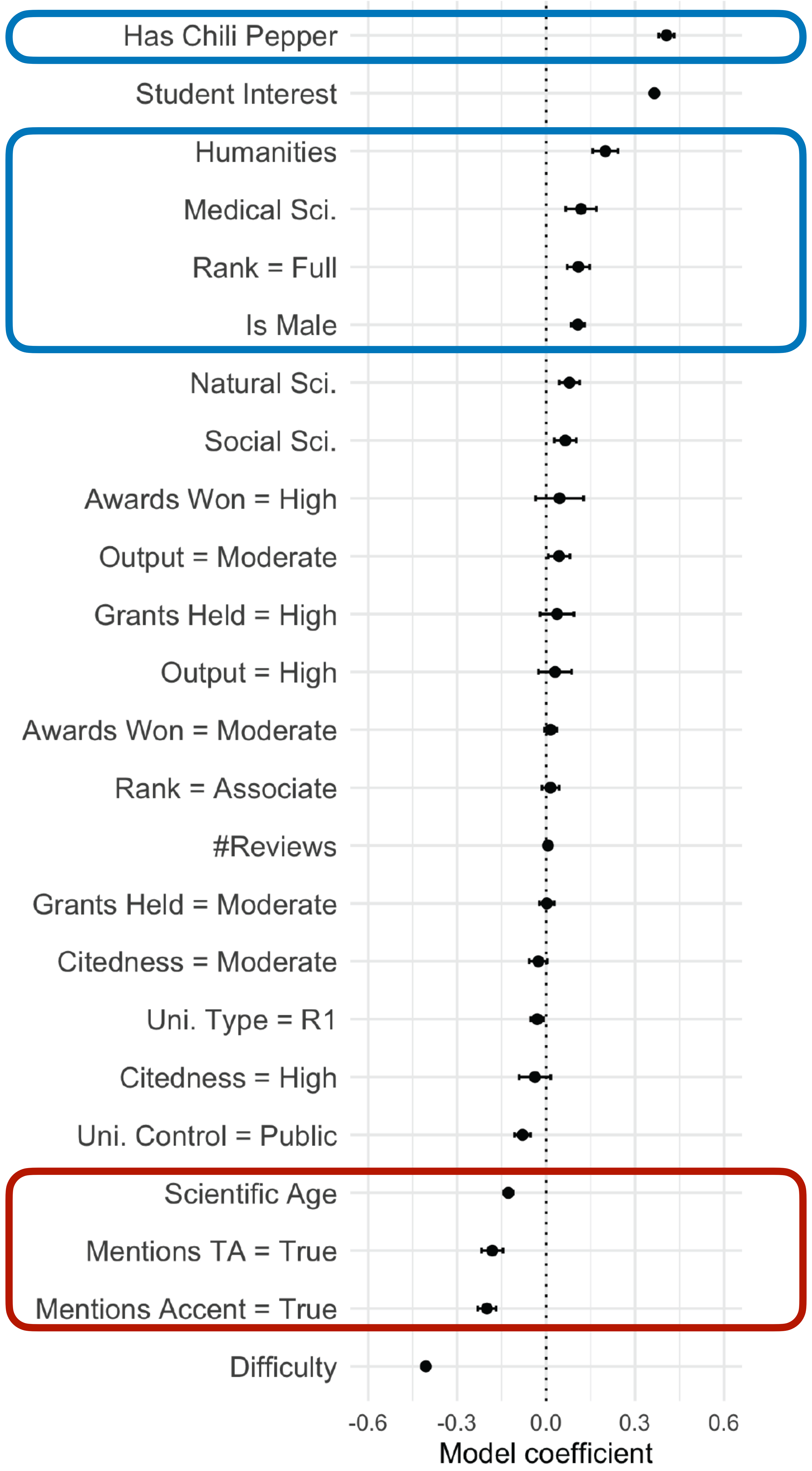


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

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

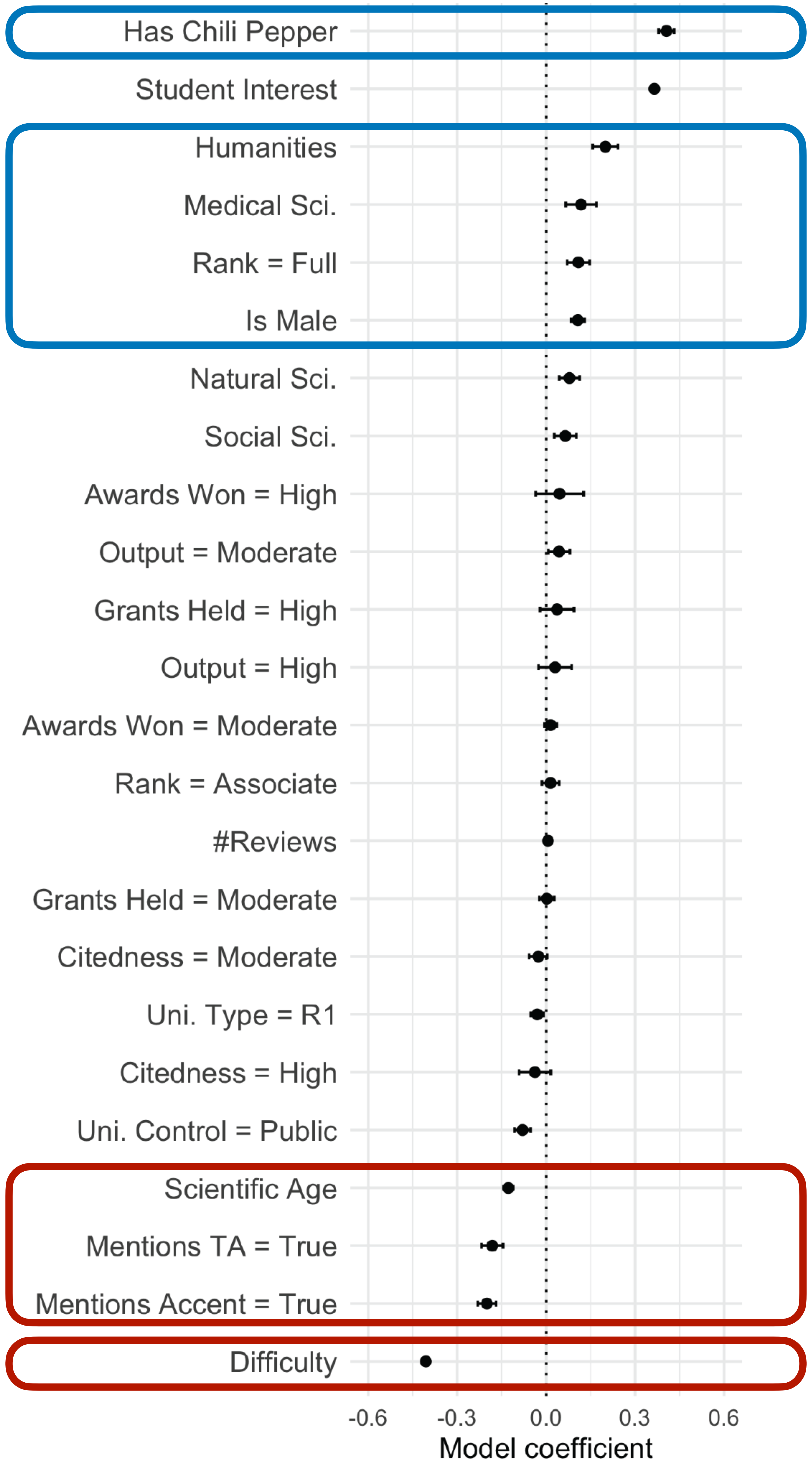


Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

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



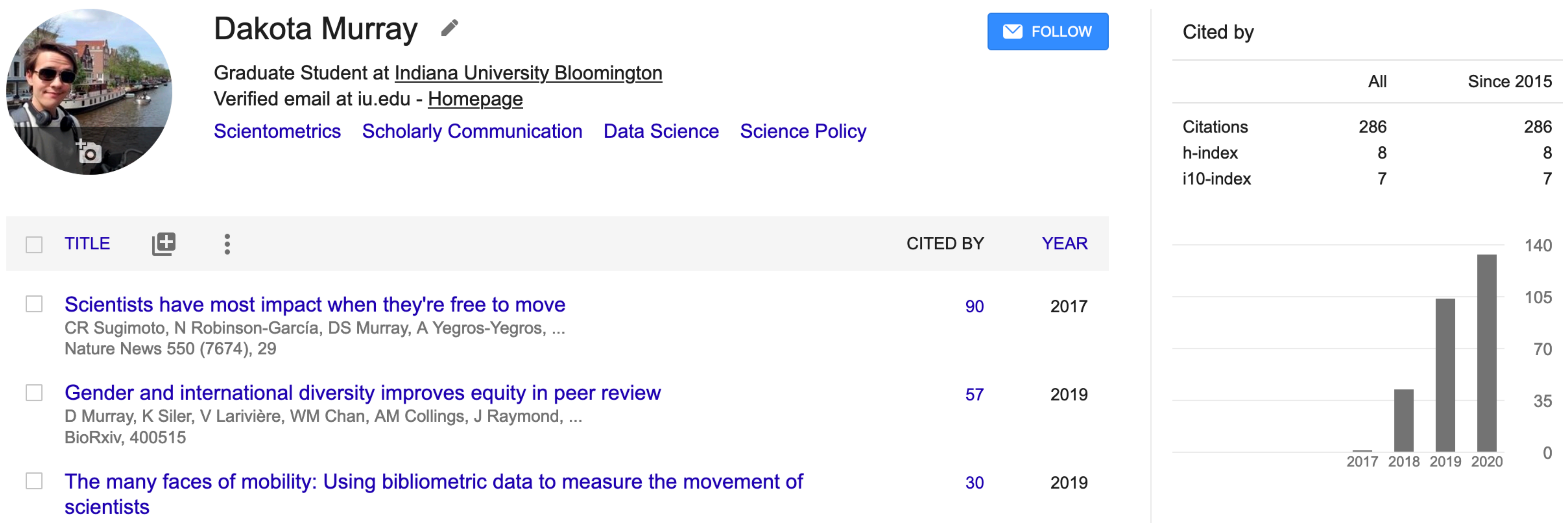
Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2020). Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLOS ONE*, *15*(6), e0233515.

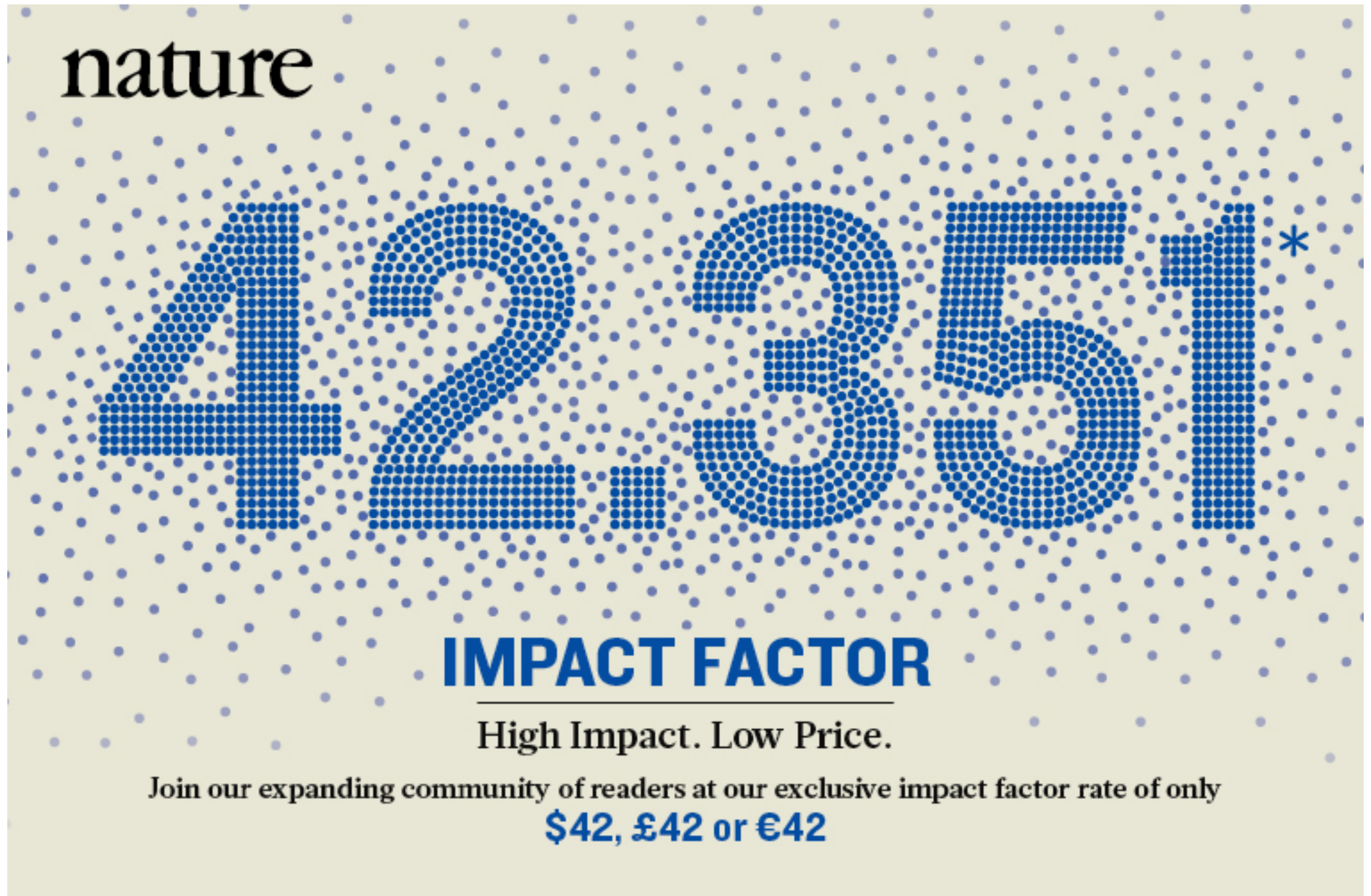
Metrics often stem from human judgements,

Have all the same subjectivity

Citation metrics

A more objective approach?





Not all citations are endorsements!

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

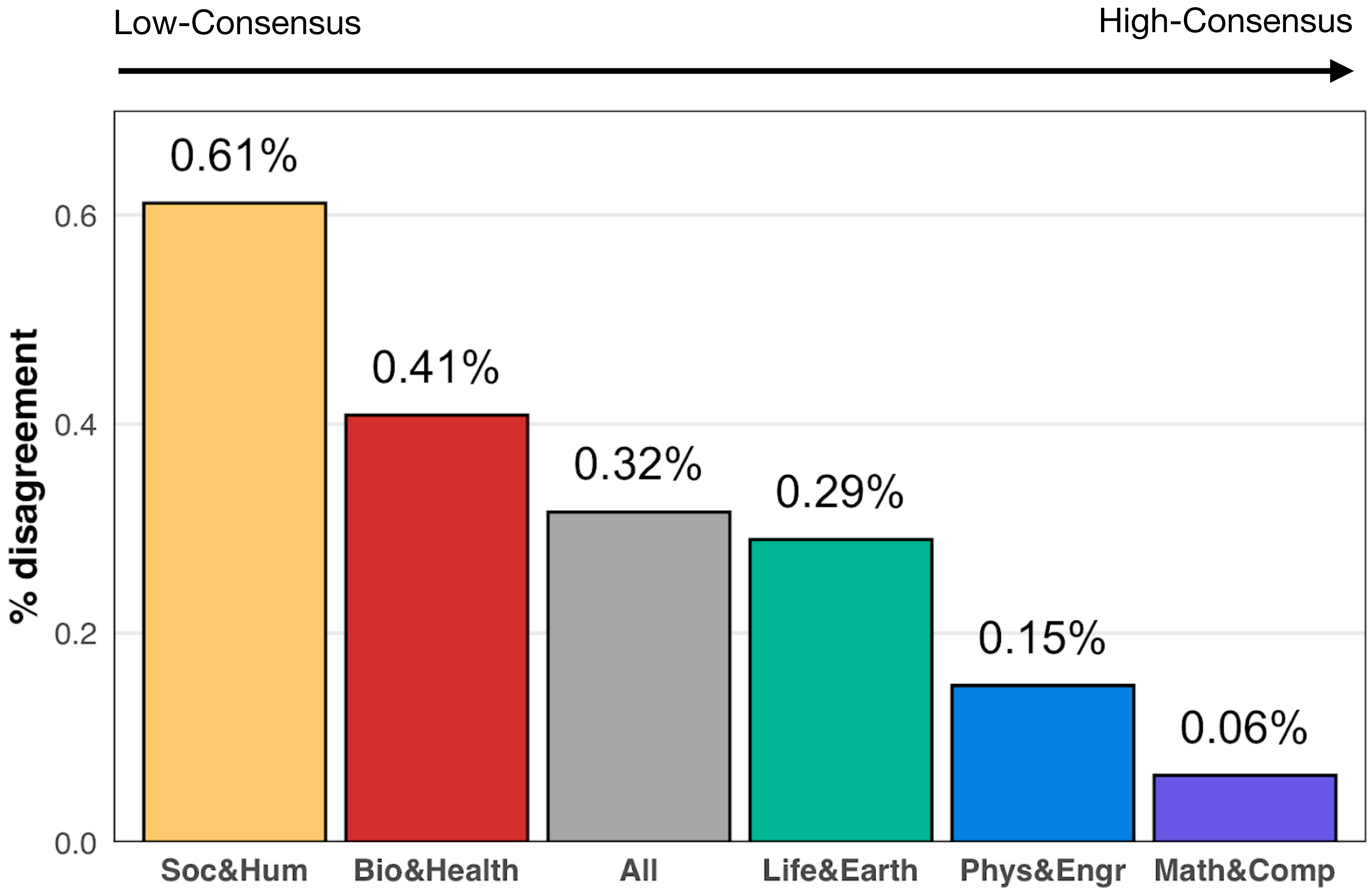
- Develop a cue-word based approach to identify 450,000 instances of disagreement across 3 million articles

- 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?

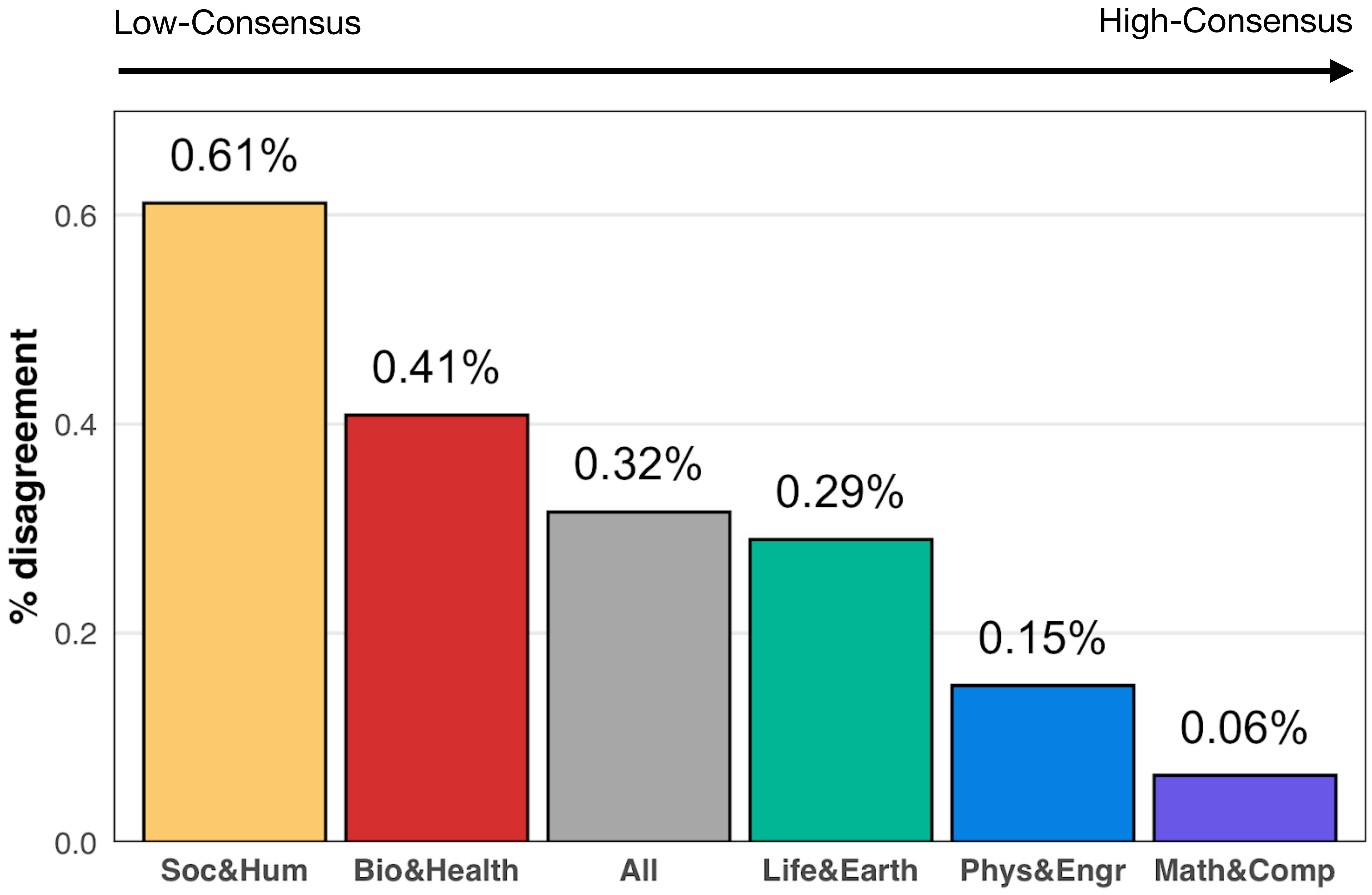


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation*).

Disagreement citations

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

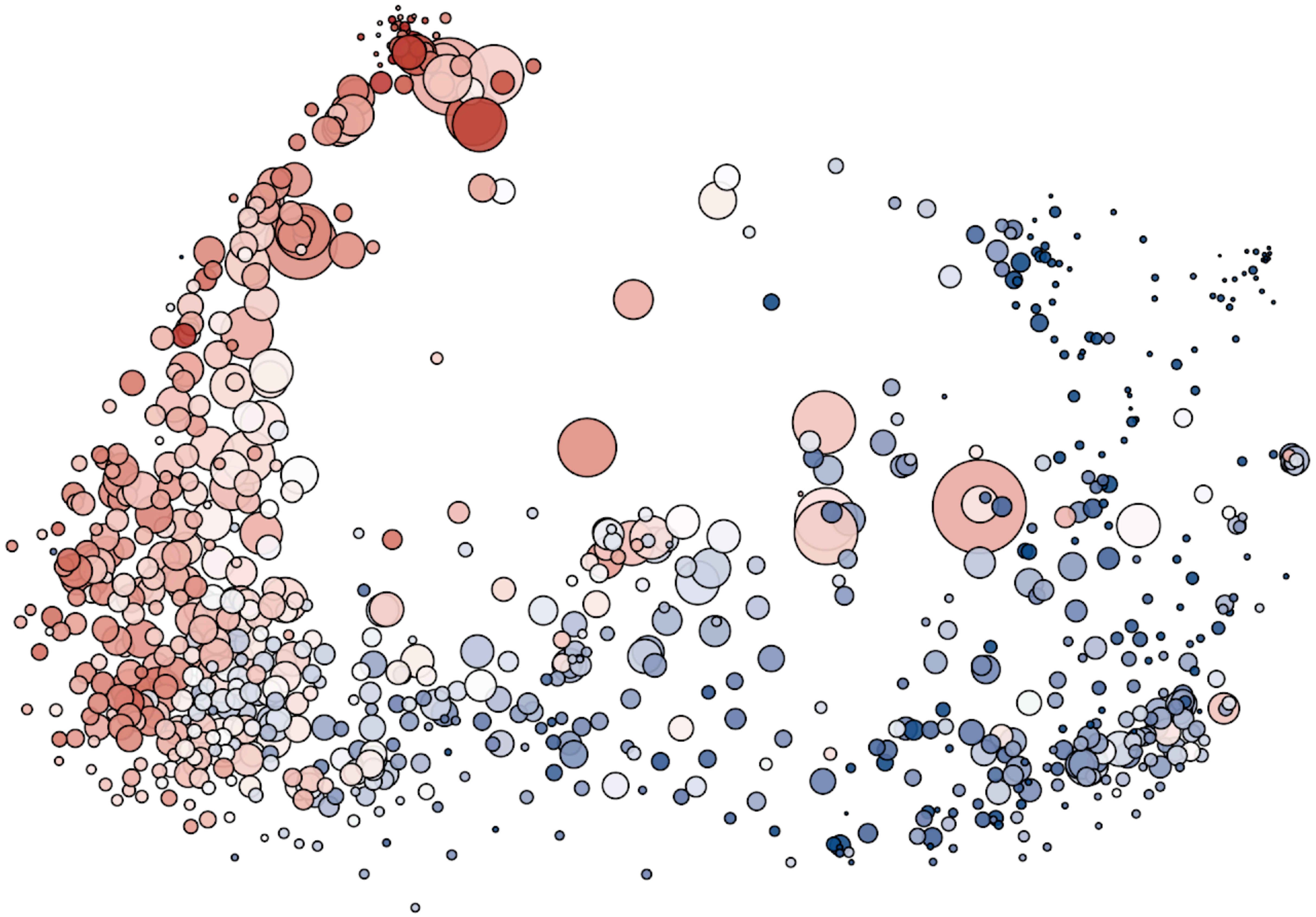
Too coarse?



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

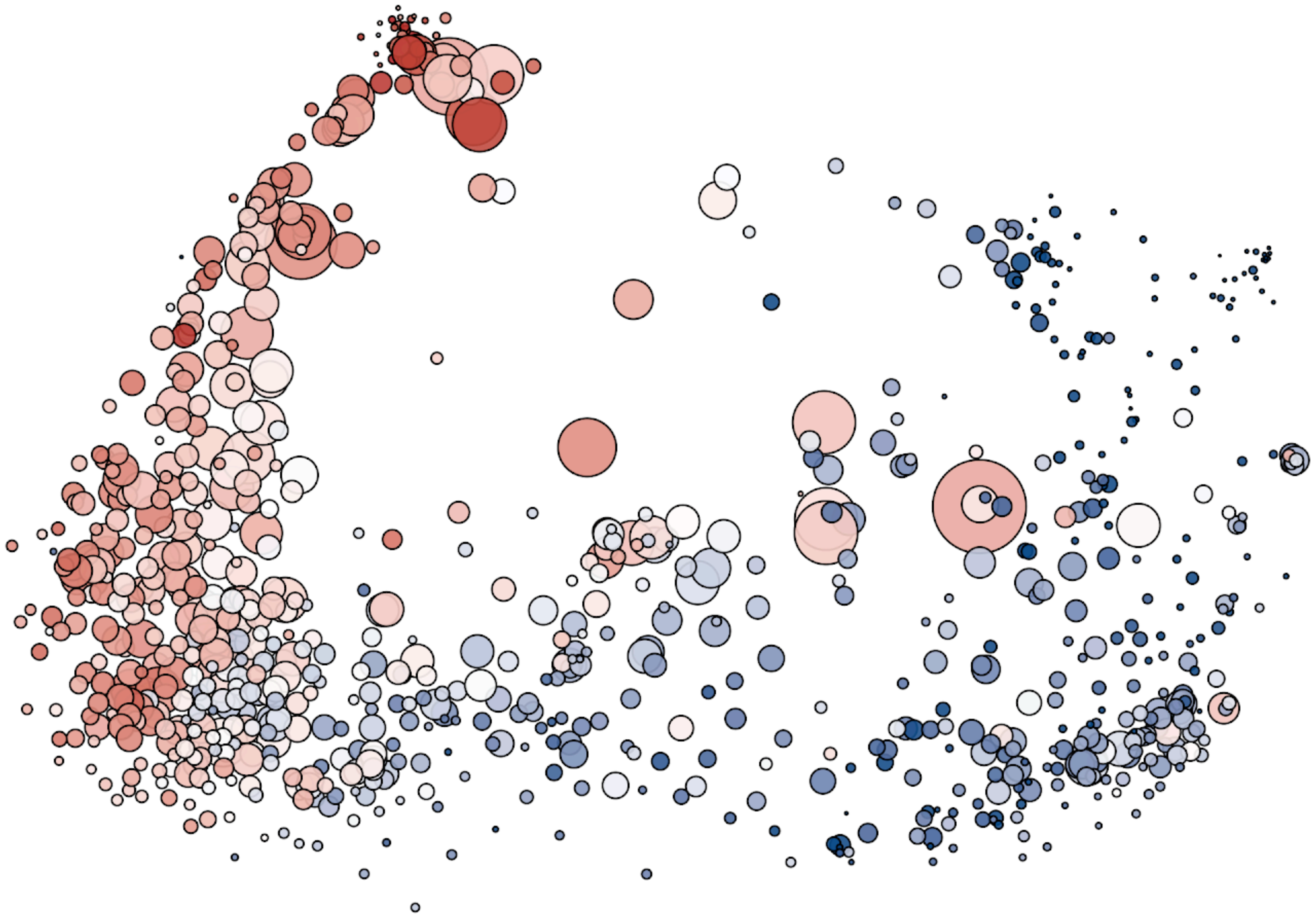


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation*).

Digging deeper

886 meso-level fields

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

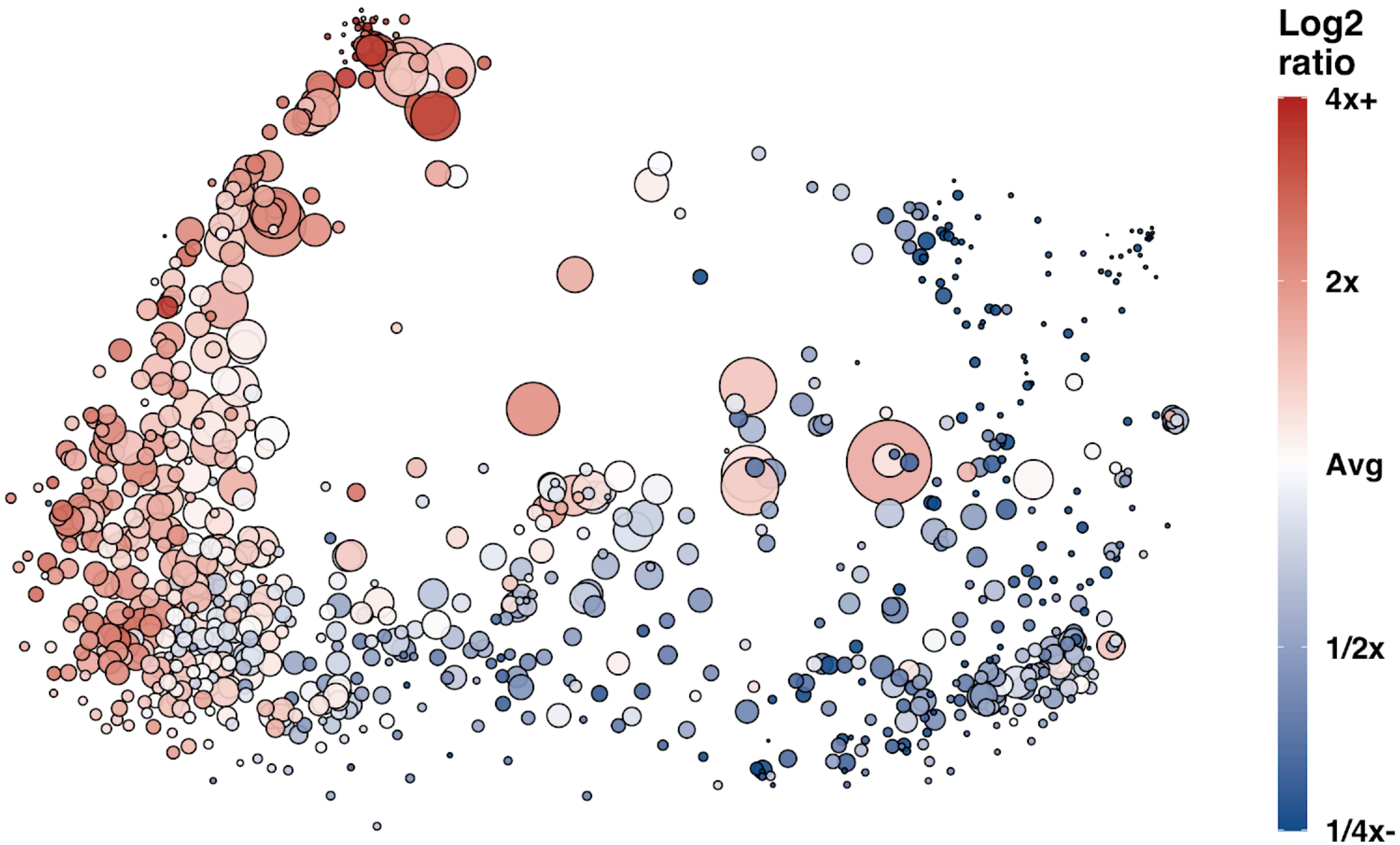


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation*).

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

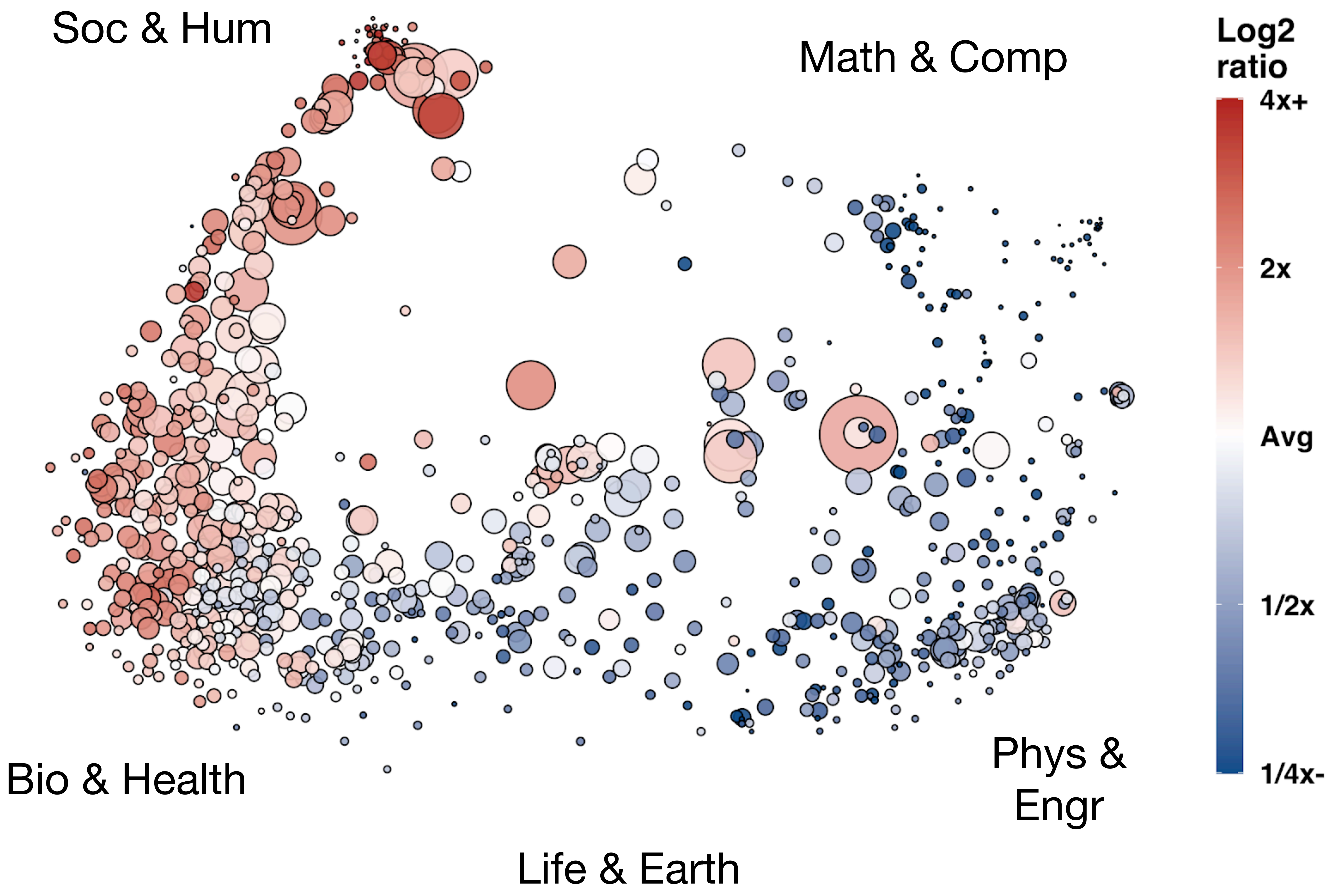


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation*).

Digging deeper

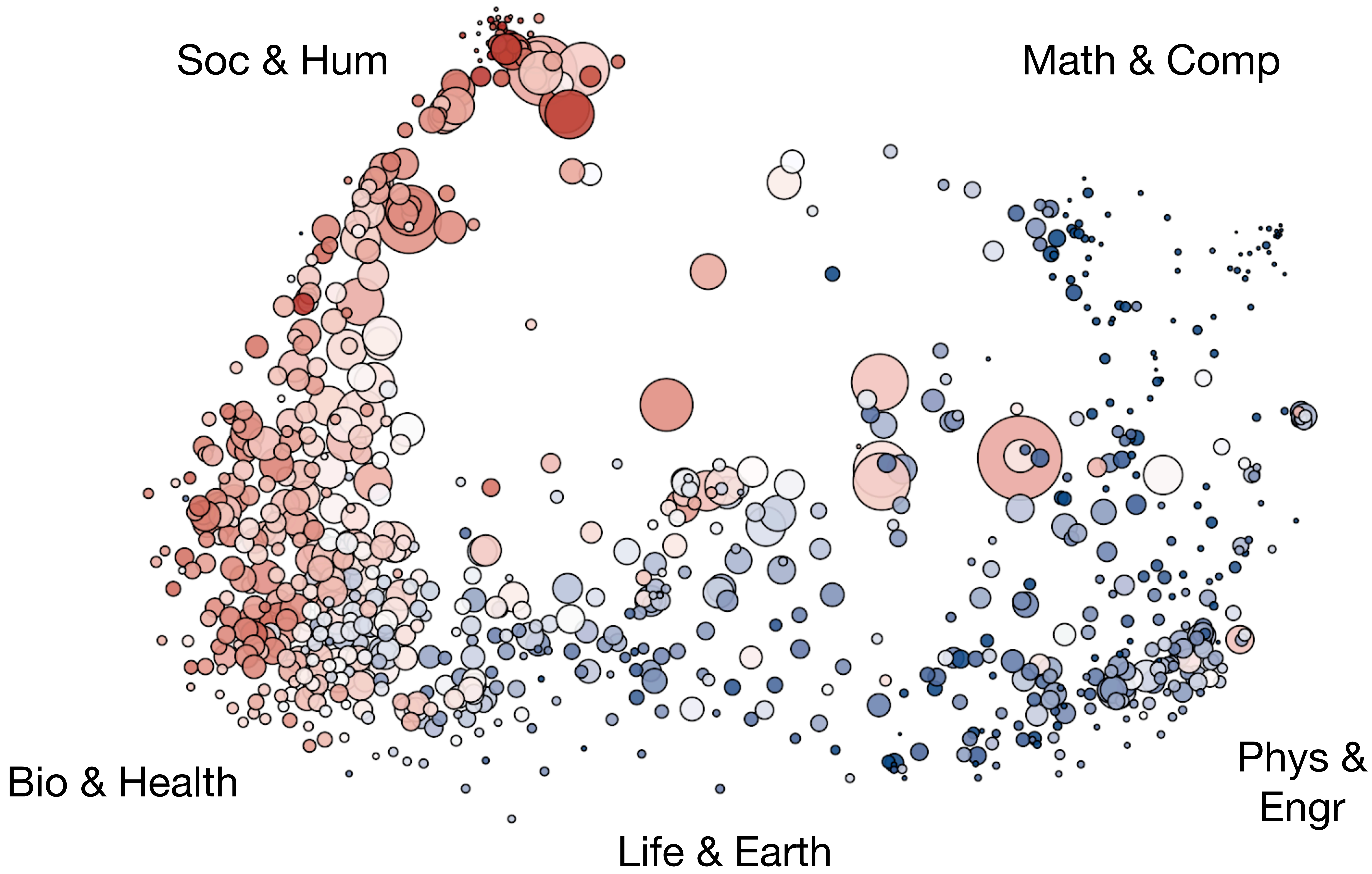
886 meso-level fields

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

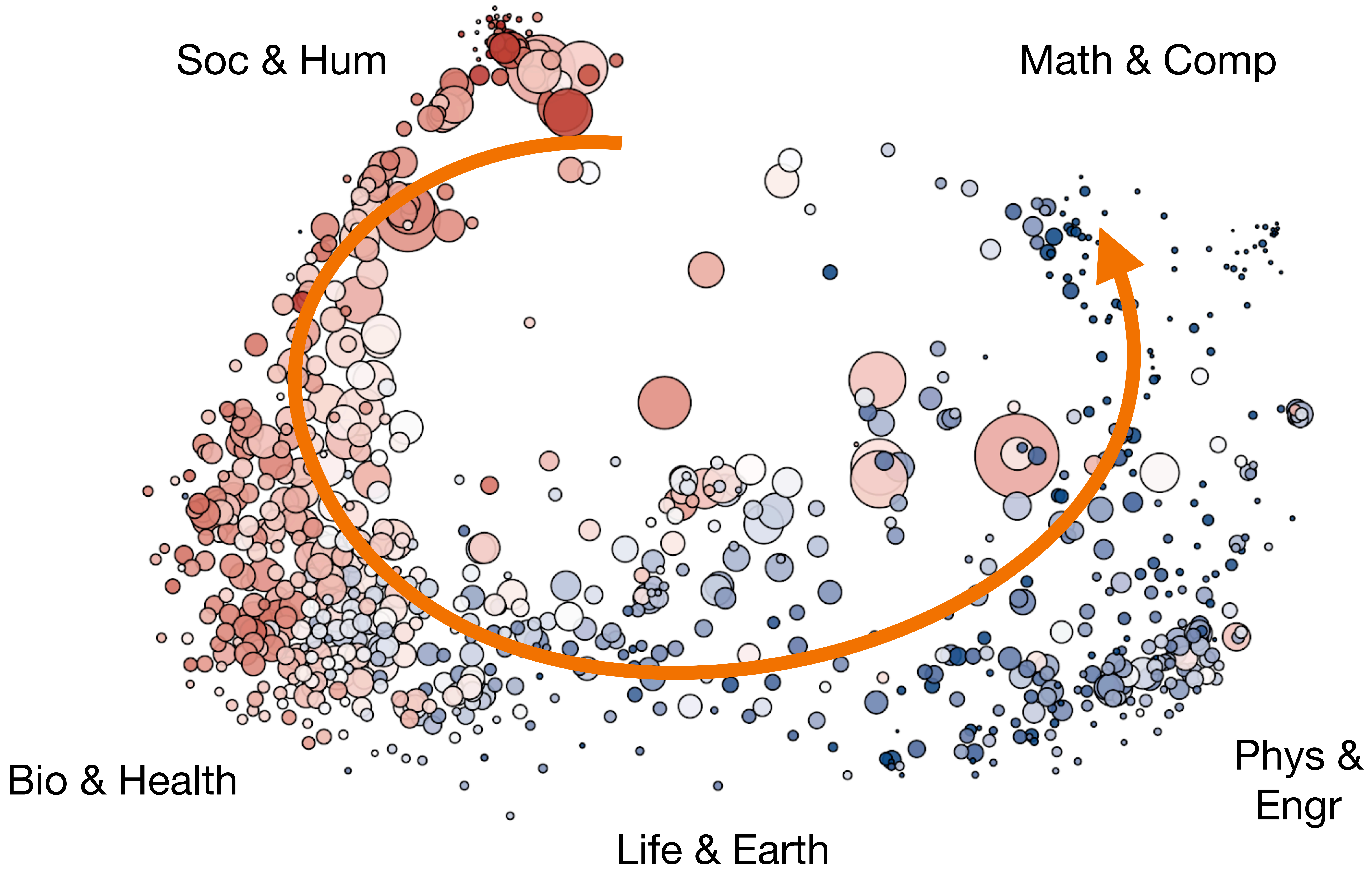


Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., & Waltman, L. Measuring disagreement in science. (*In Preparation*).

The pattern repeats

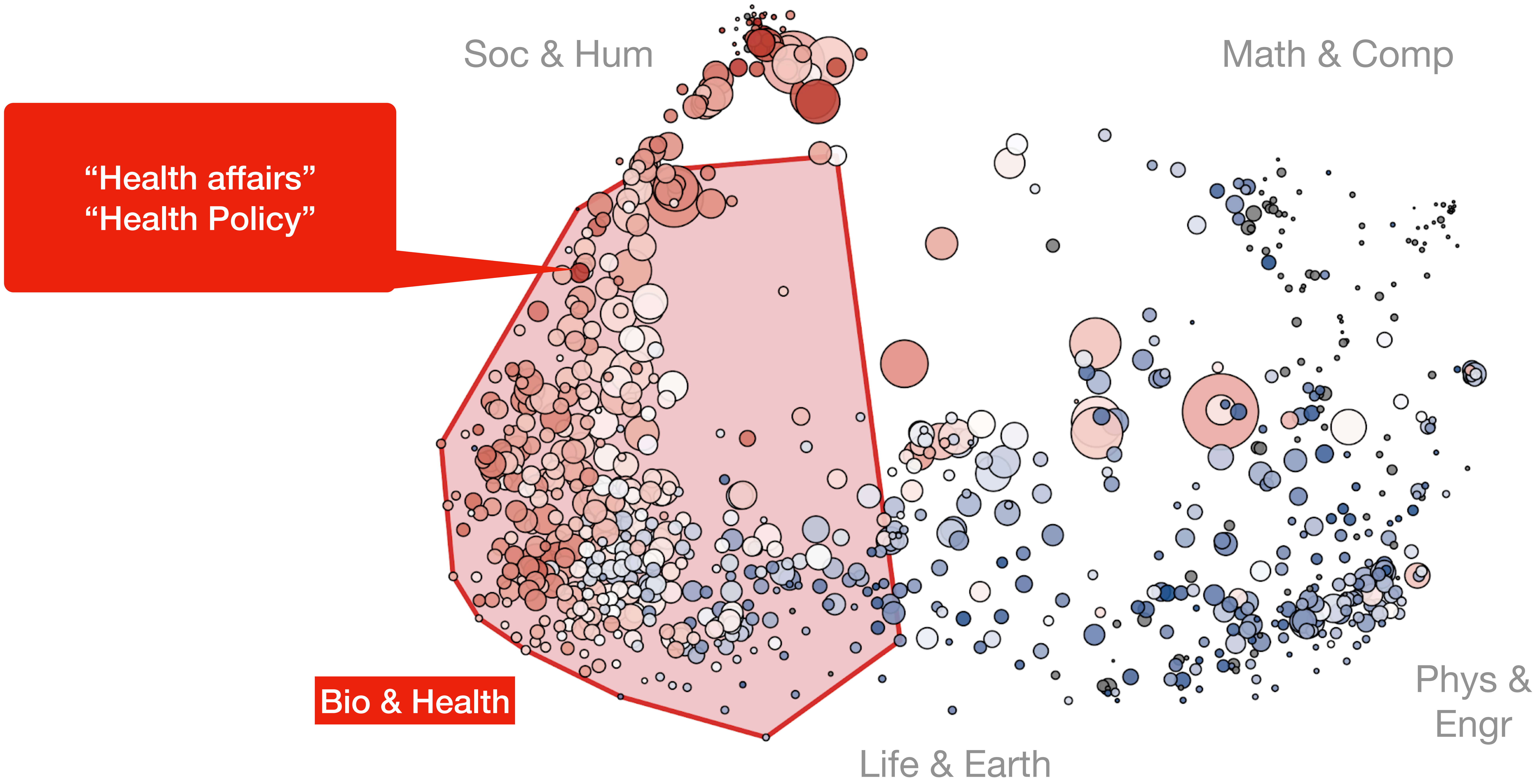


The pattern repeats



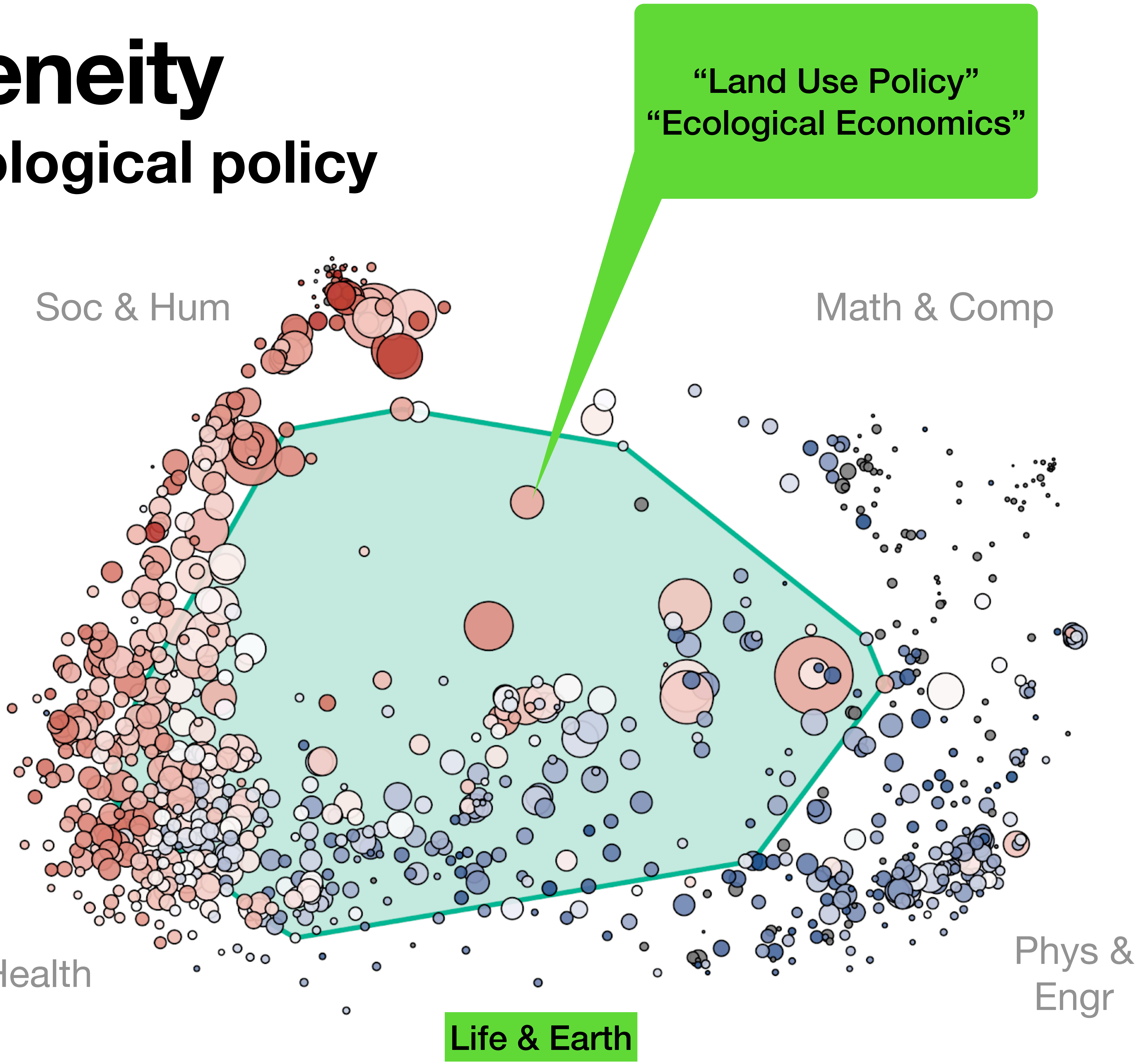
Heterogeneity

Health policy has high-disagreement relative to surrounding fields



Heterogeneity

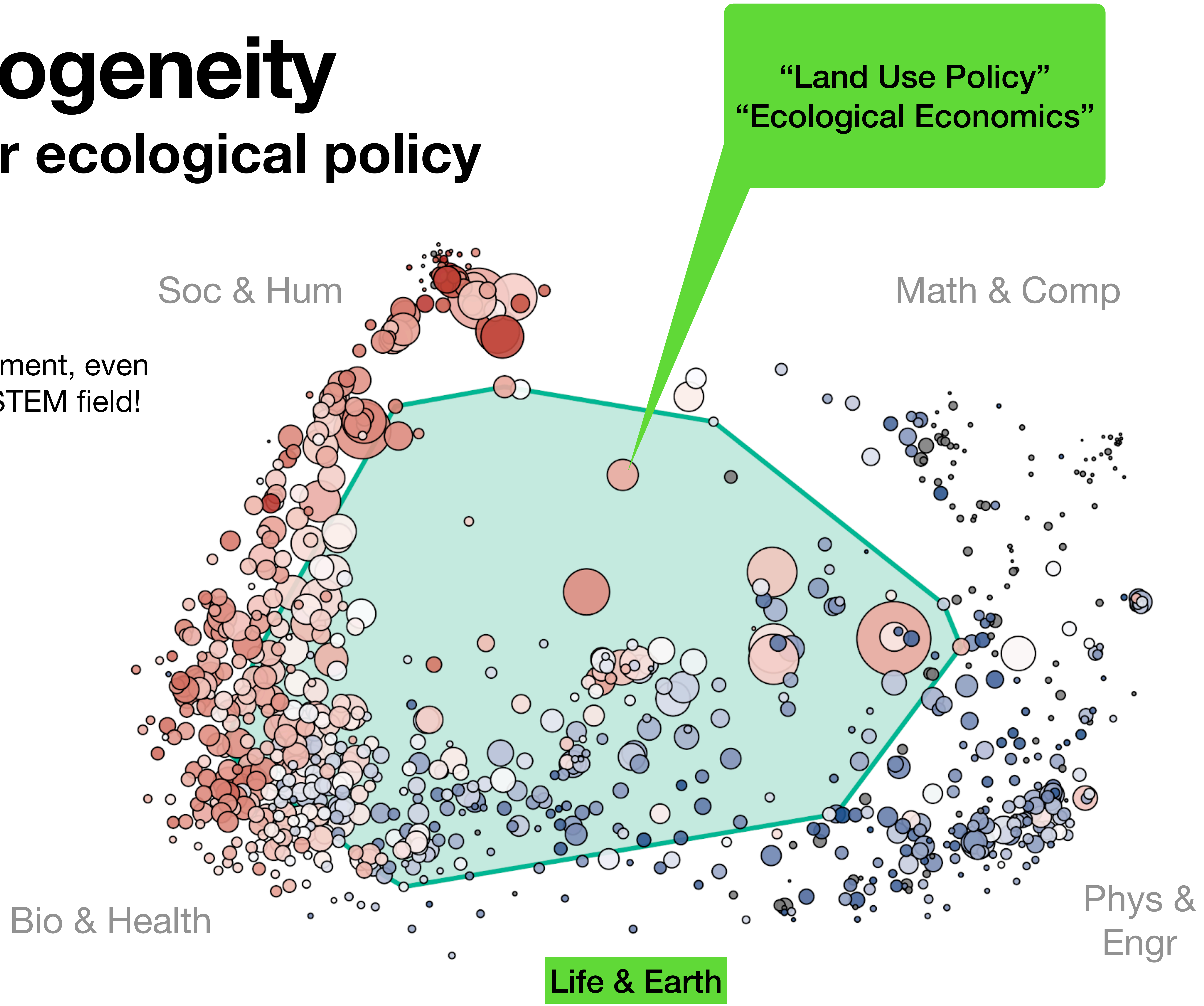
Same for ecological policy



Heterogeneity

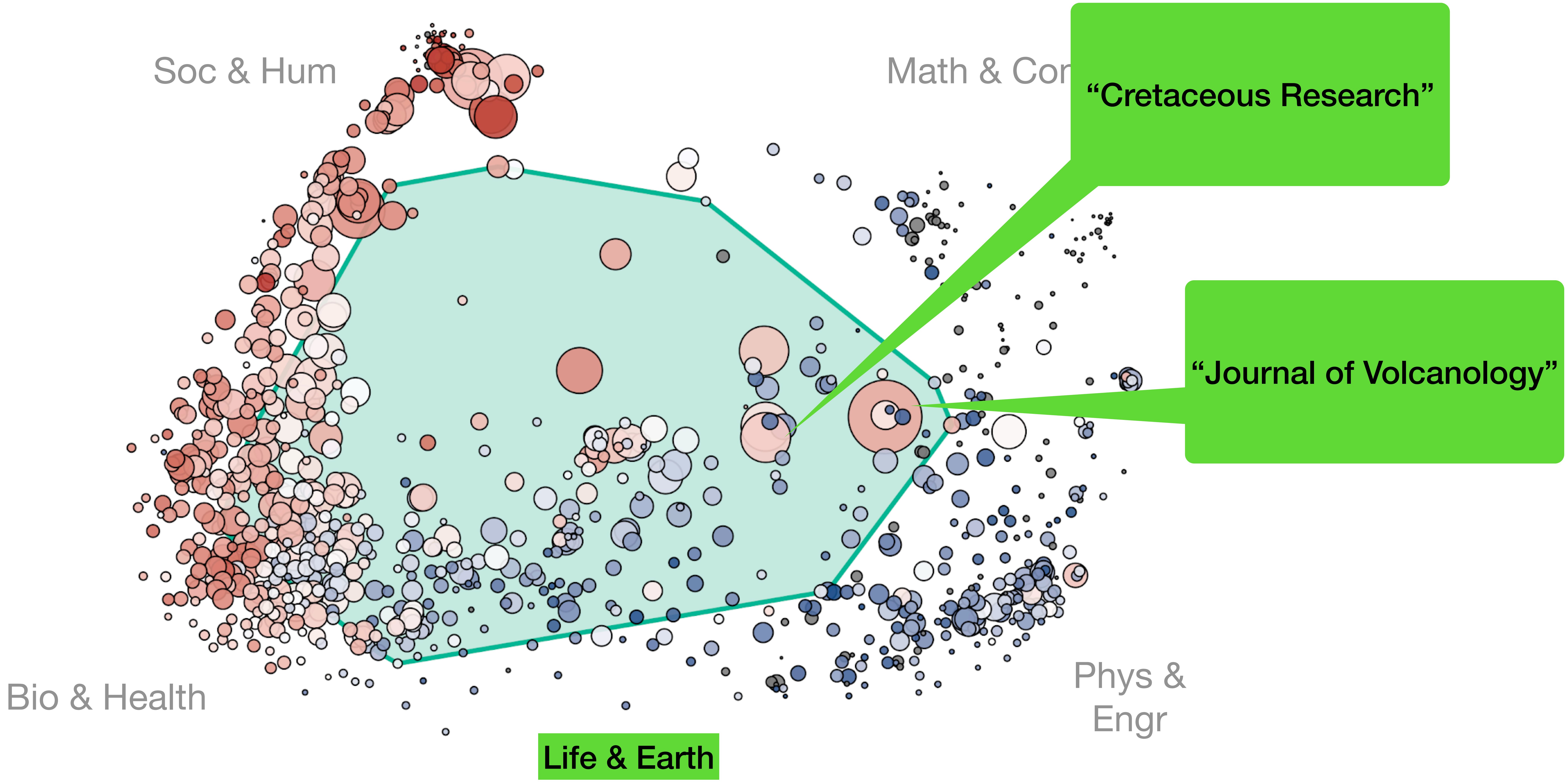
Same for ecological policy

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



Heterogeneity

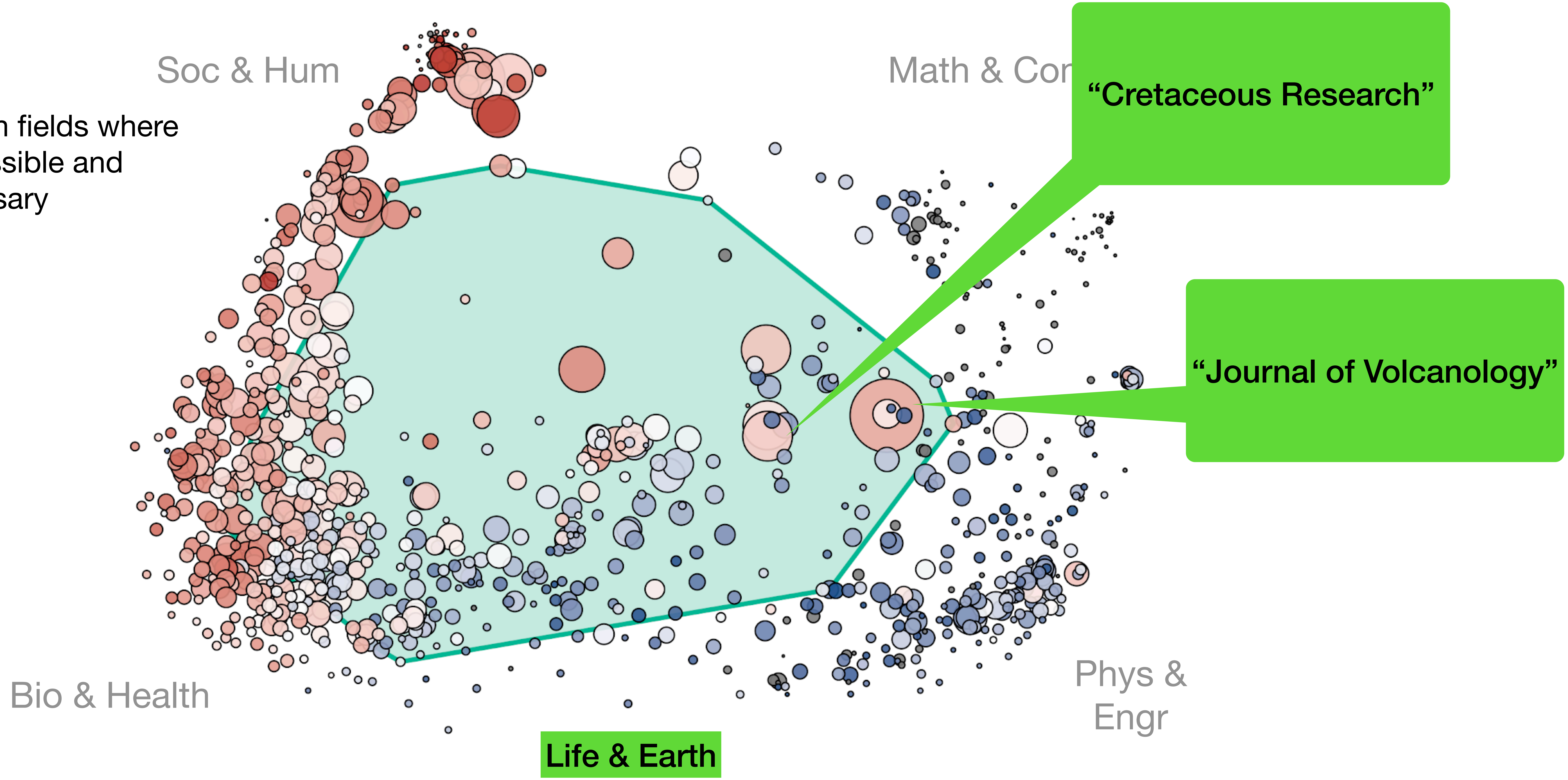
Fields depending on historical records



Heterogeneity

Fields depending on historical records

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



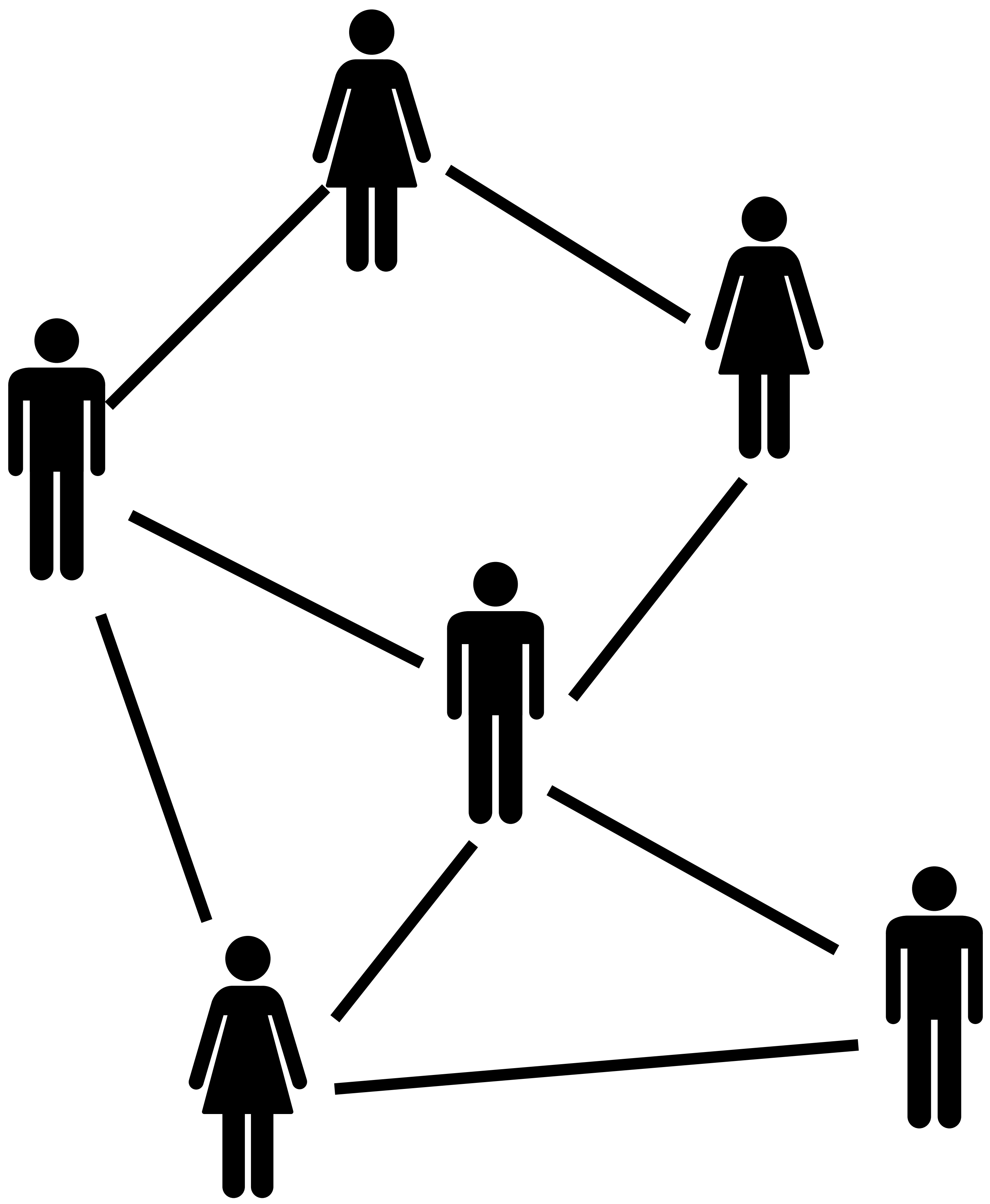
Disciplines have different norms, expectations, and cultures that shape how and why they cite

In metrics…a number is just a number

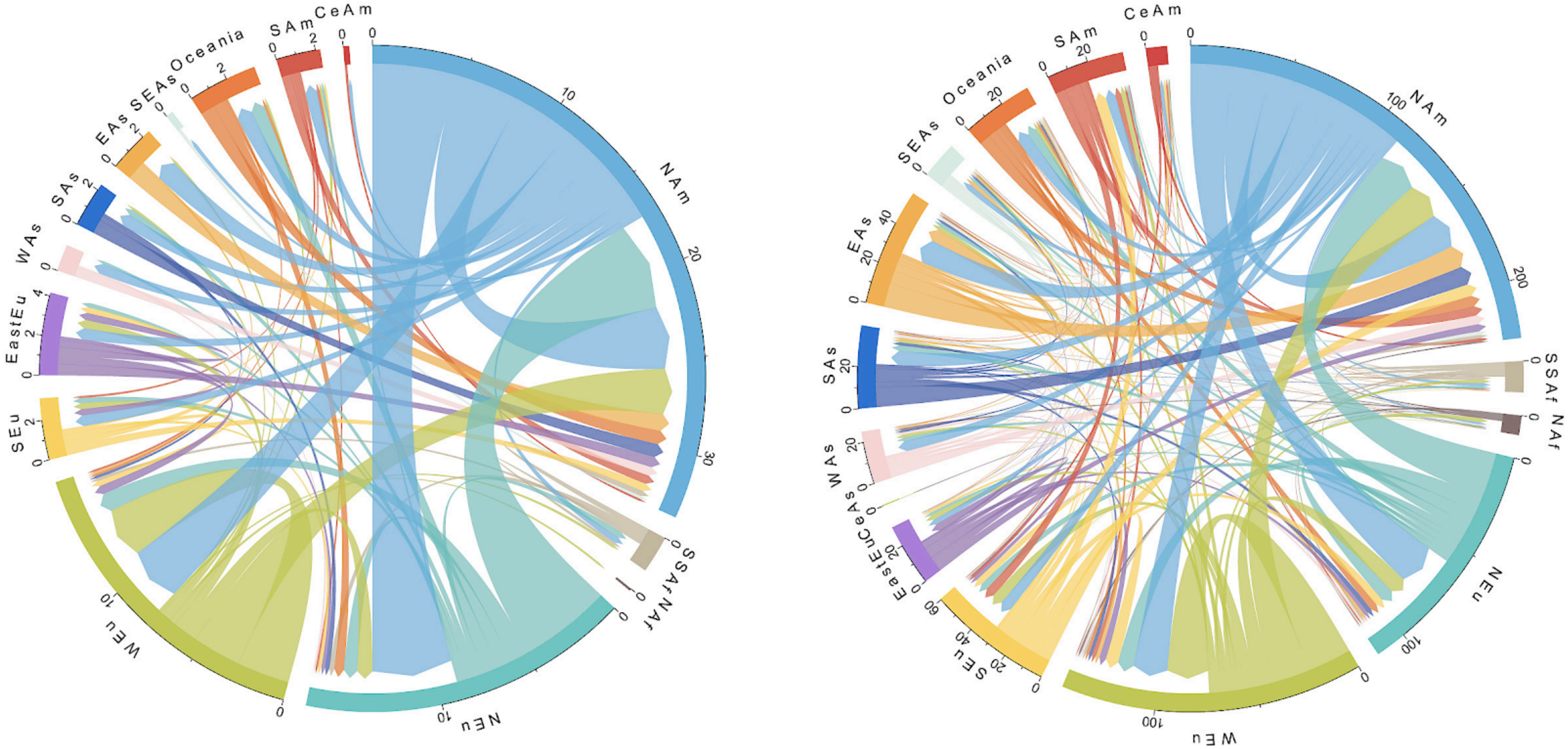
Context is needed to interpret it

Peer-to-peer judgements

(networks)



Scientific mobility

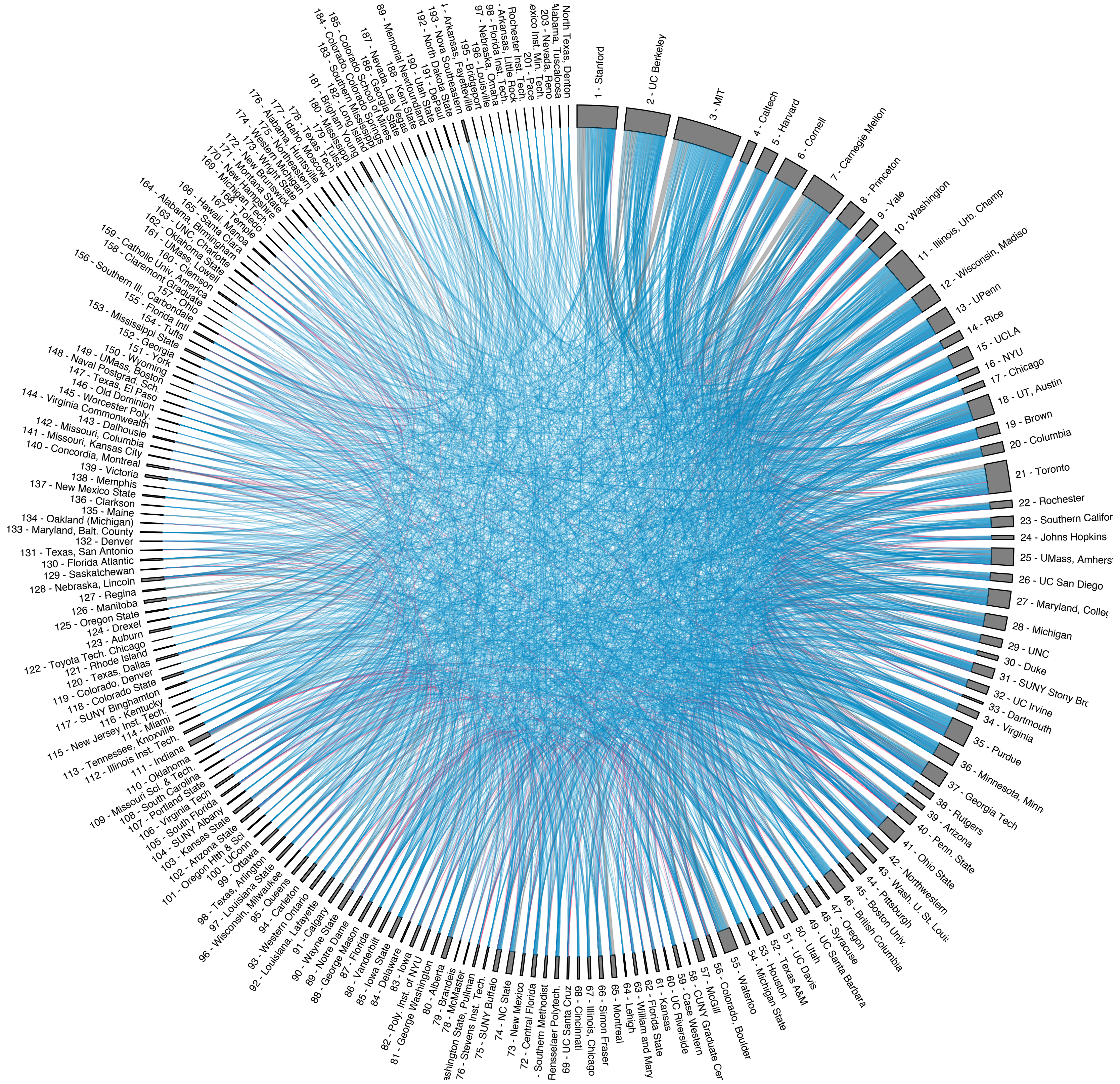


Czaika, M., & Orazbayev, S. (2018). The globalisation of scientific mobility, 1970–2014. *Applied Geography*, *96*, 1–10.

Mobility is intimately tied to evaluation and success

Mobility is intimately tied to evaluation and success

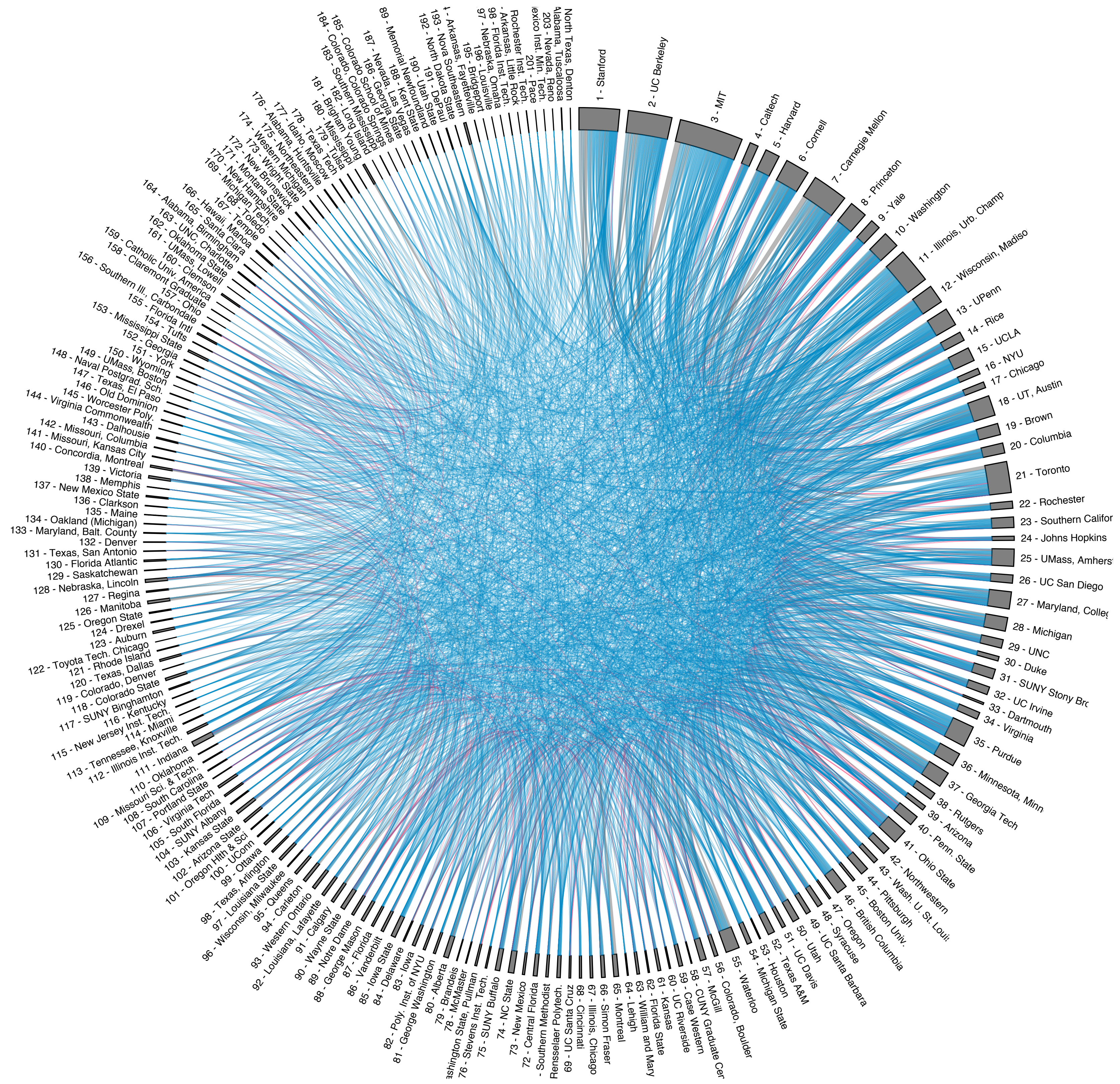
Mobility an output of evaluation in faculty hiring



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

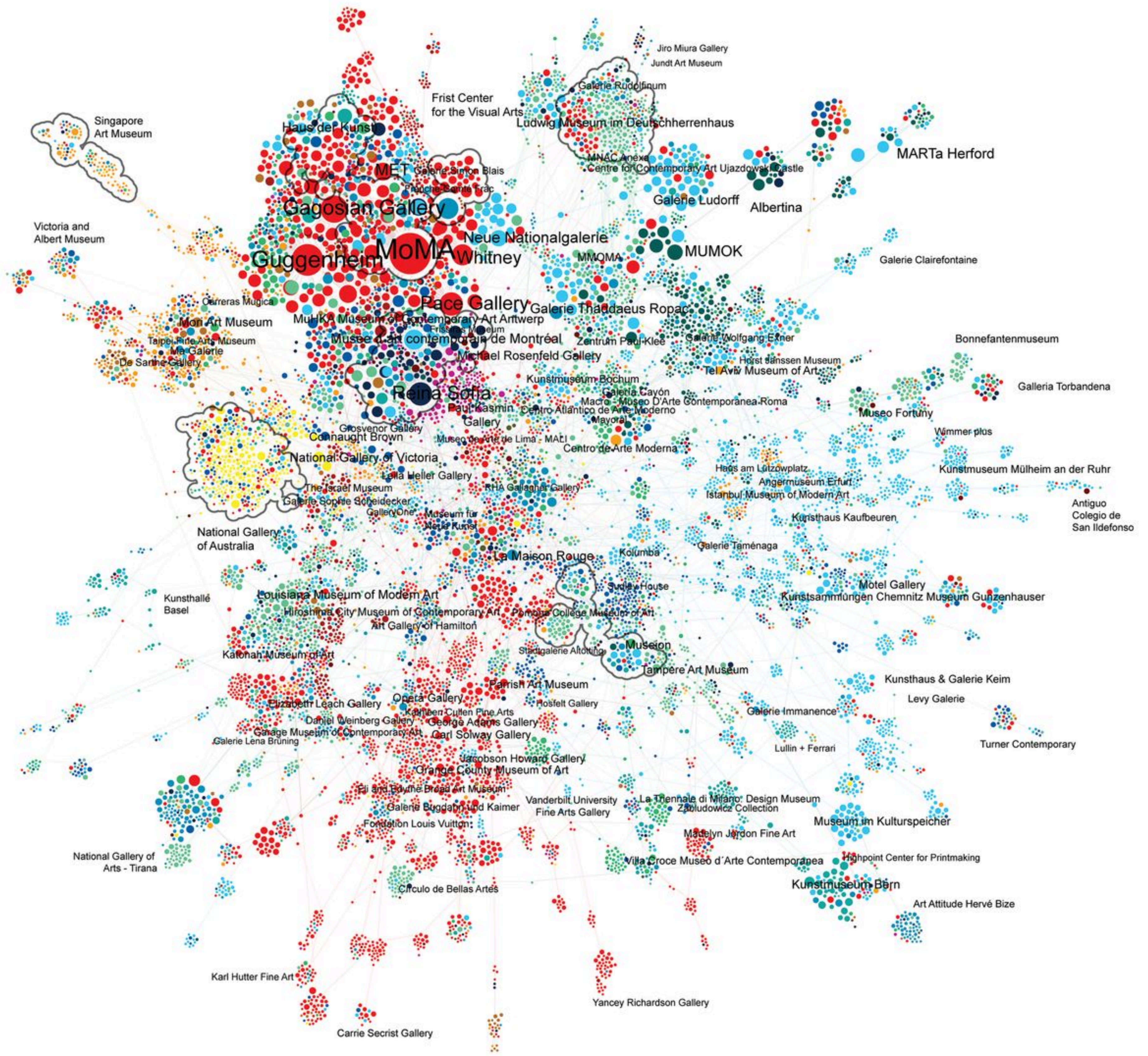
Mobility is intimately tied to evaluation and success

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.

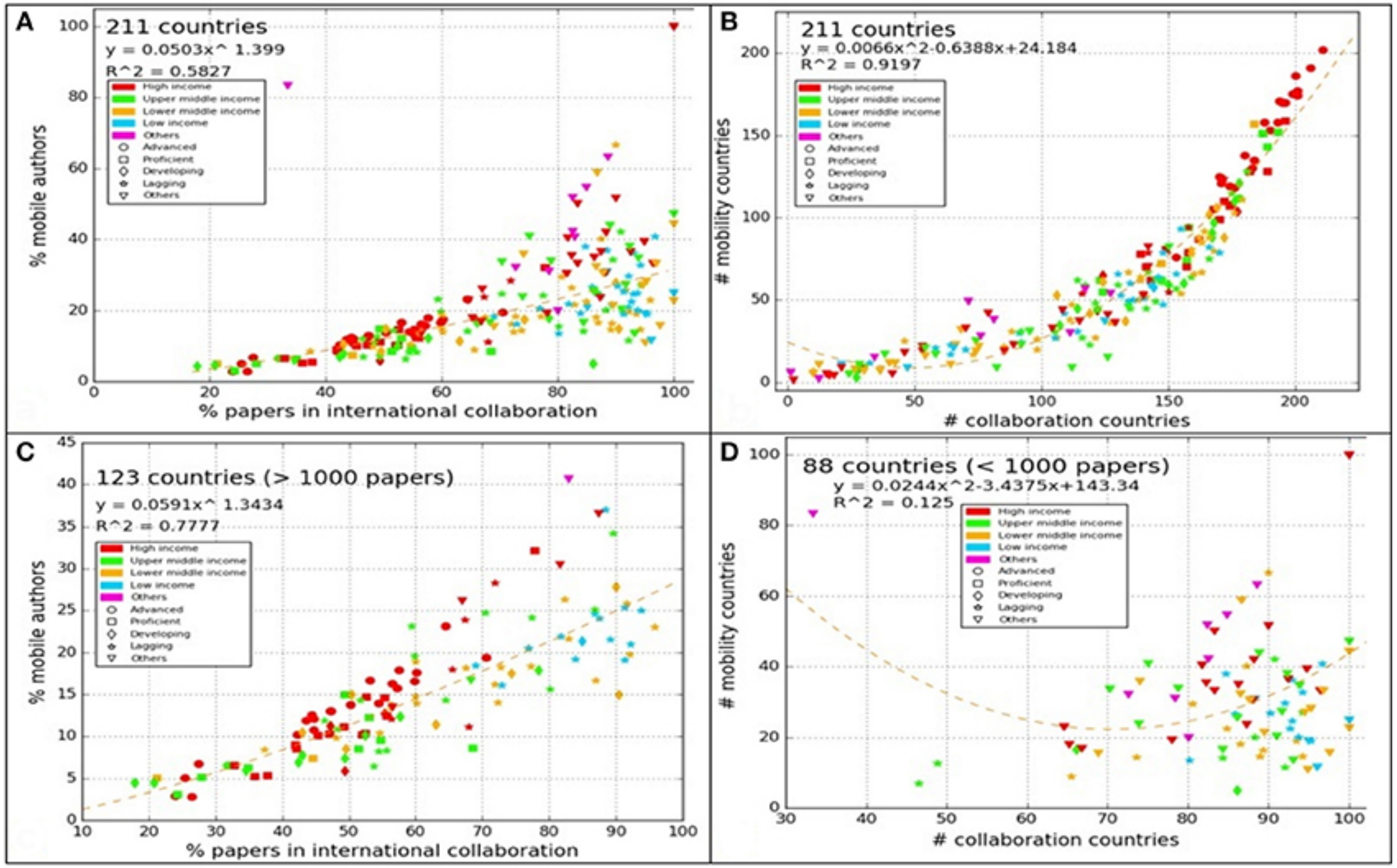
Mobility an input to success in the art world



Fraiberger, S. P., Sinatra, R., Resch, M., Riedl, C., & Barabási, A.-L. (2018). Quantifying reputation and success in art. *Science*, *362*(6416), 825–829.

Mobility expands our professional networks

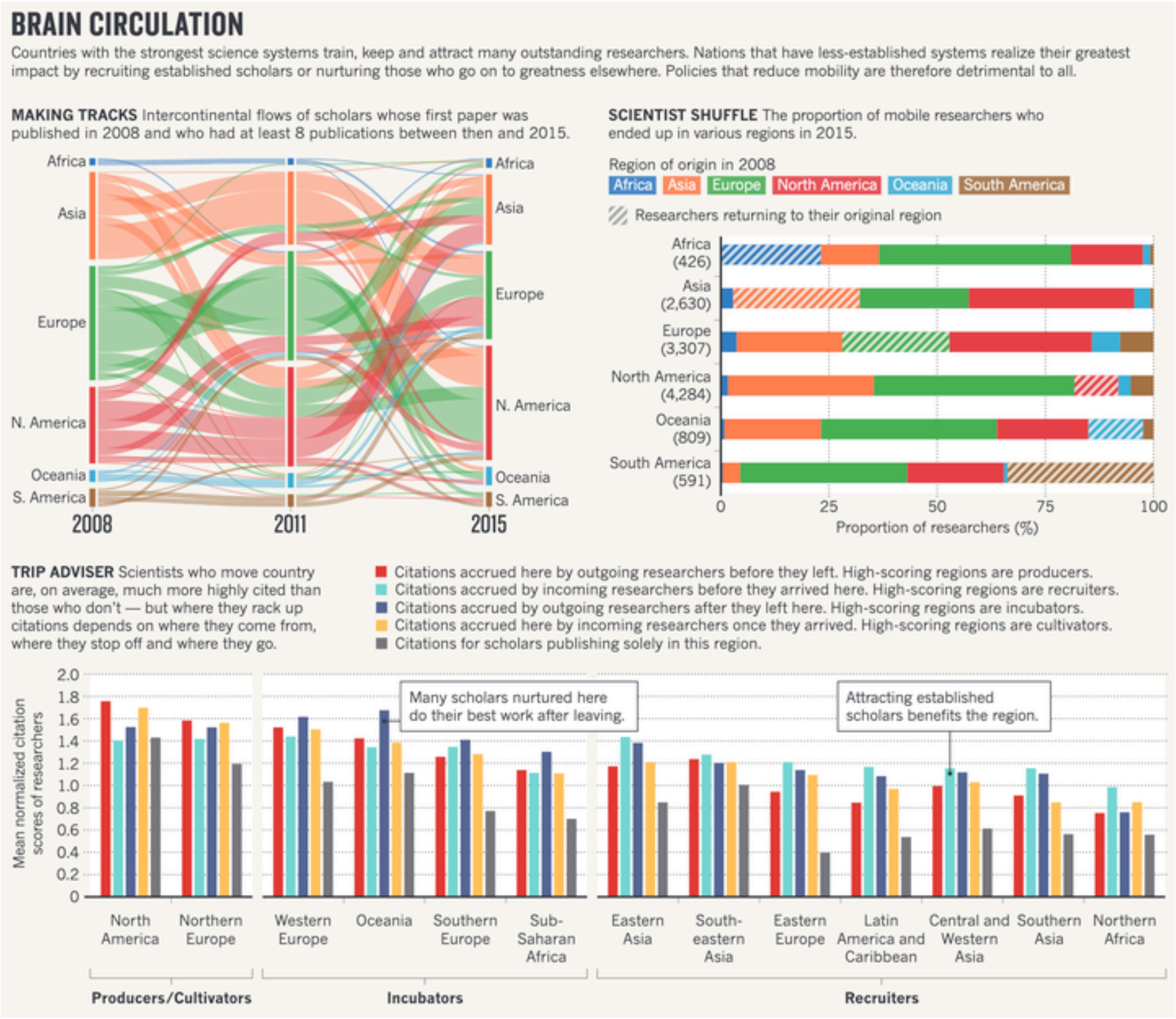
Countries with more mobility between them, also have more collaboration!



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!



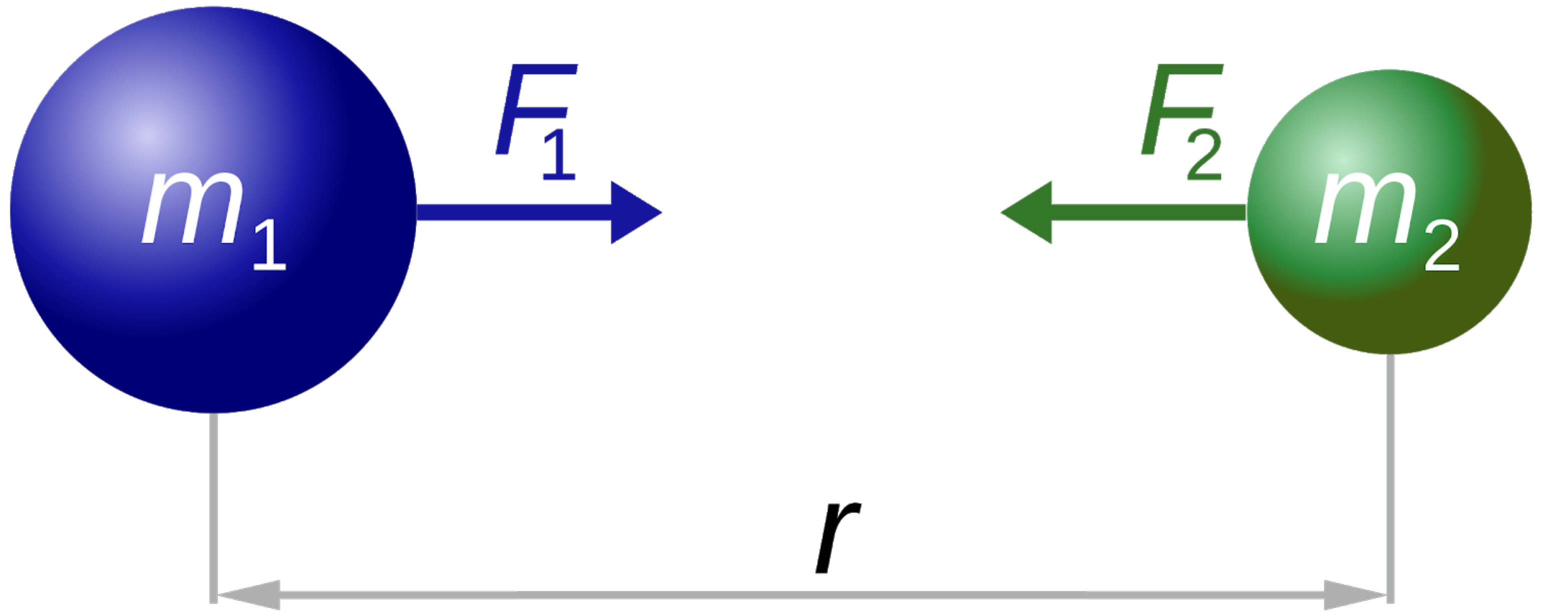
Sugimoto, C. R., Robinson-Garcia, N., **Murray, D. S.**, Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, *550*(7674), 29–31.

Mobility is deeply contextual

Not all mobility is the same

Gravity model

Ubiquitous and intuitive



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

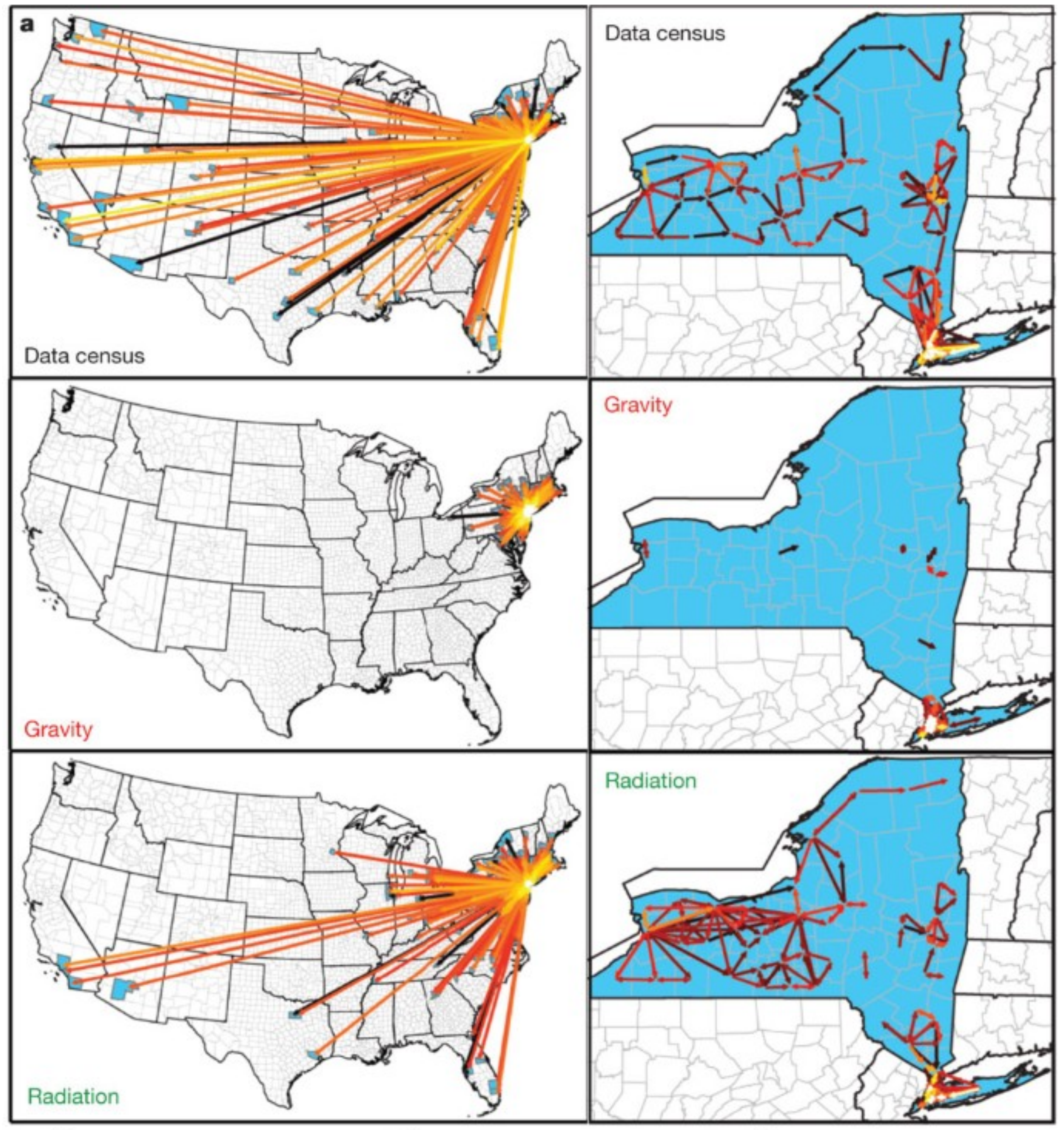
Wikipedia user Dennis Nilsson

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

Flu *Populati* *Decay*

Radiation model

Improve upon the gravity model

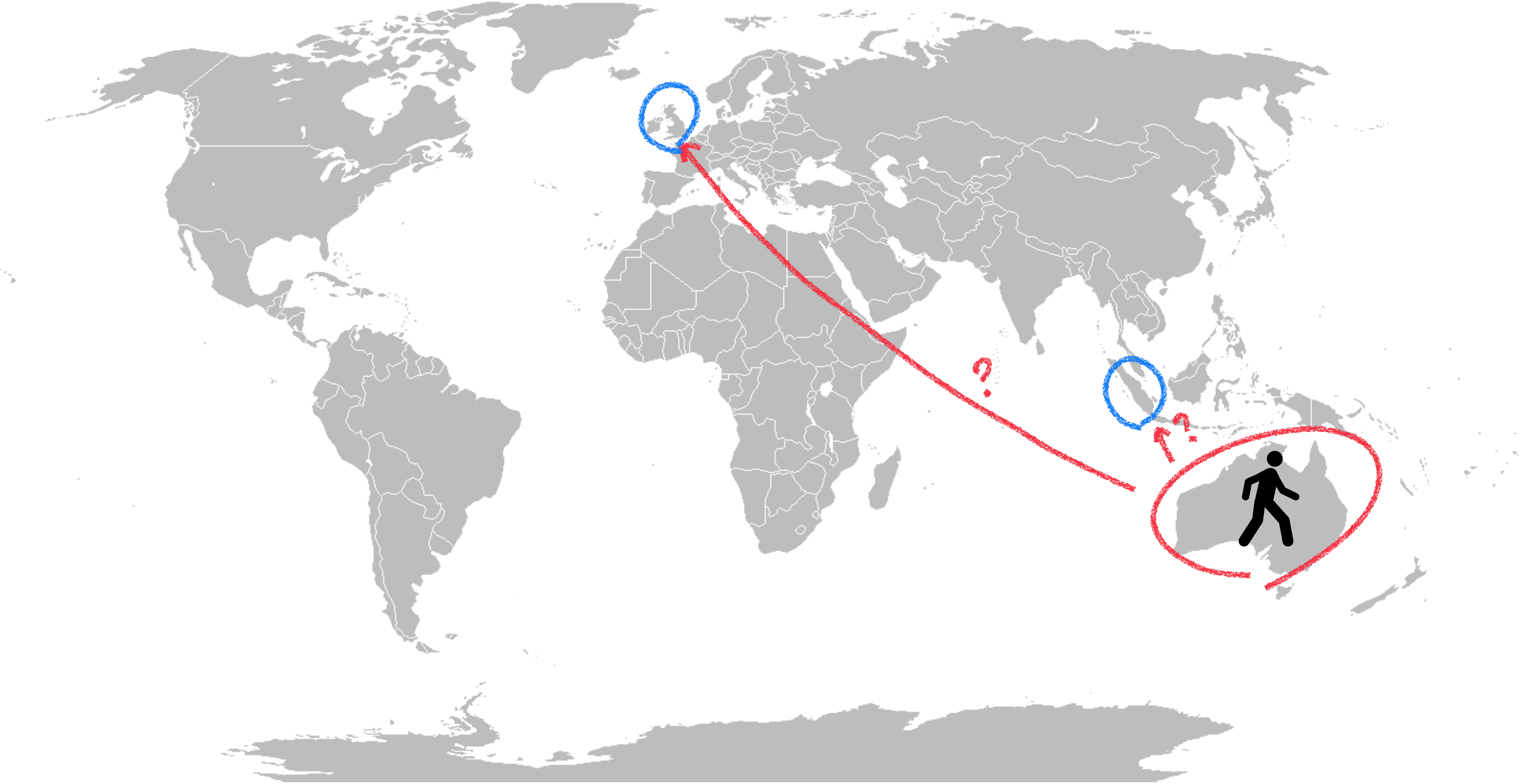


$$[T_{ij}] = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})}$$

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

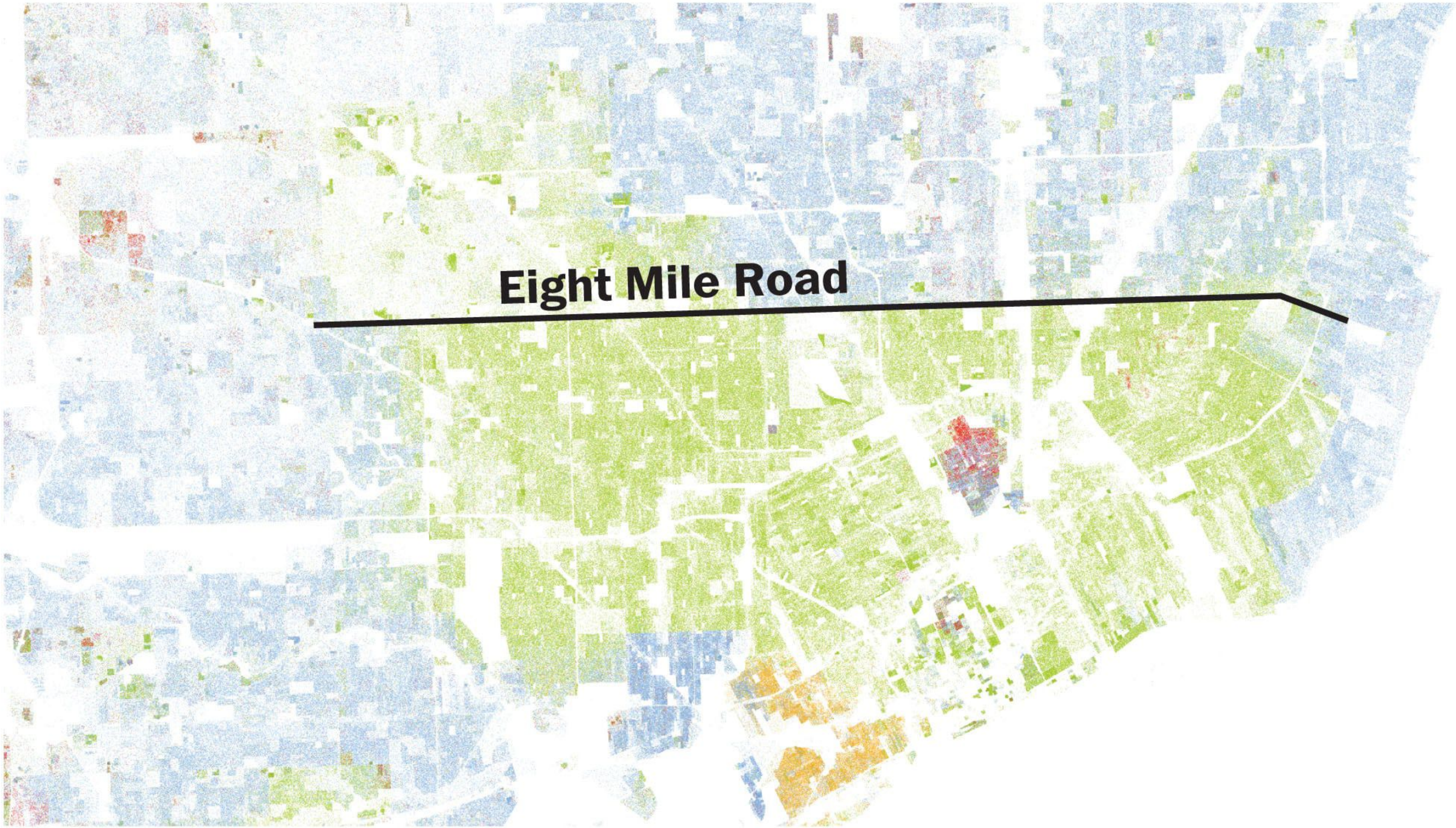
Mobility models are effective…

But geographic distance is not always appropriate



Lines of segregation in Detroit

1 dot = 1 person ● White ● Black ● Asian ● Hispanic ● Other

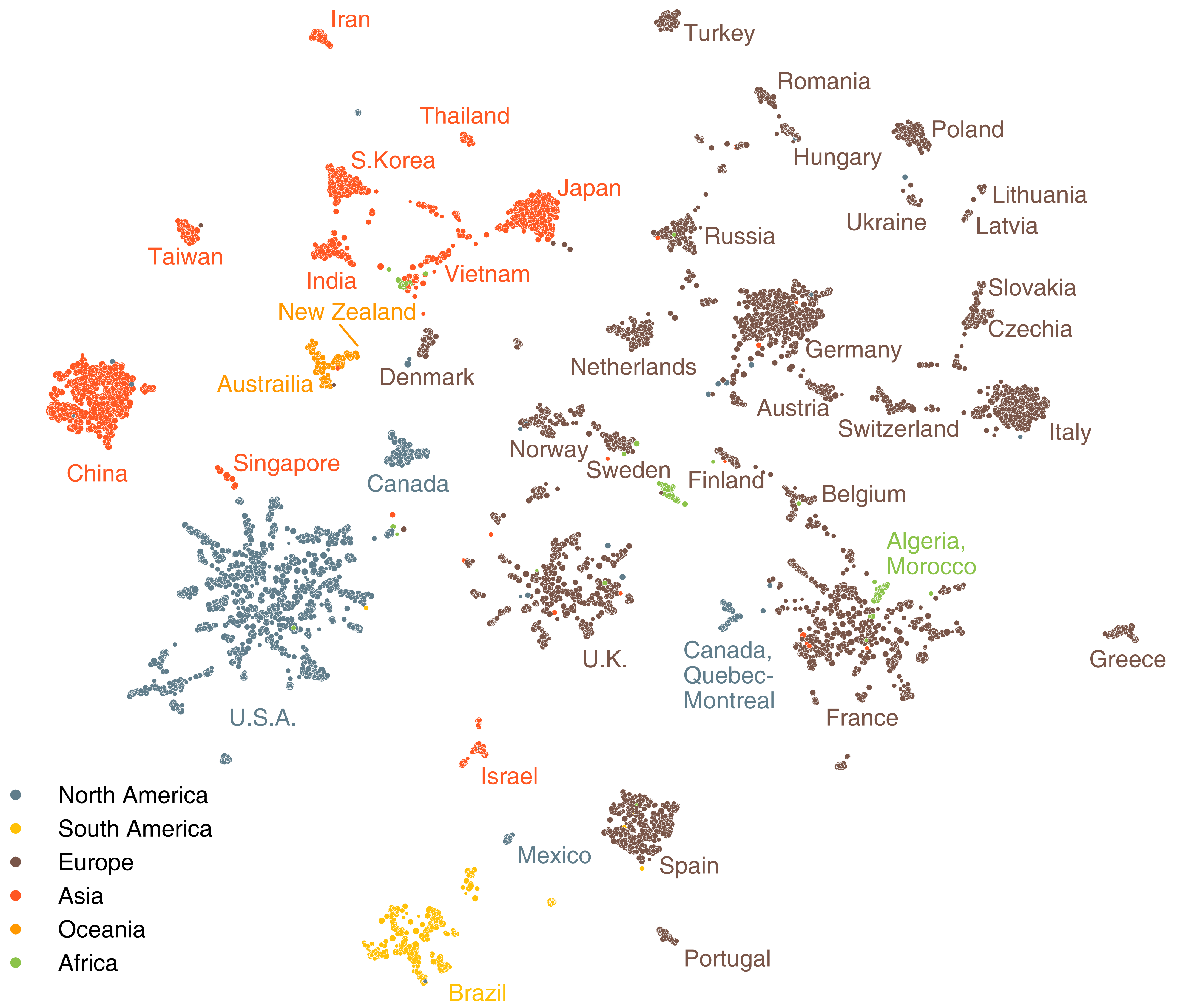


Source: U-Va. Cooper Center analysis of 2010 Census data THE WASHINGTON POST

Geography matters, but…

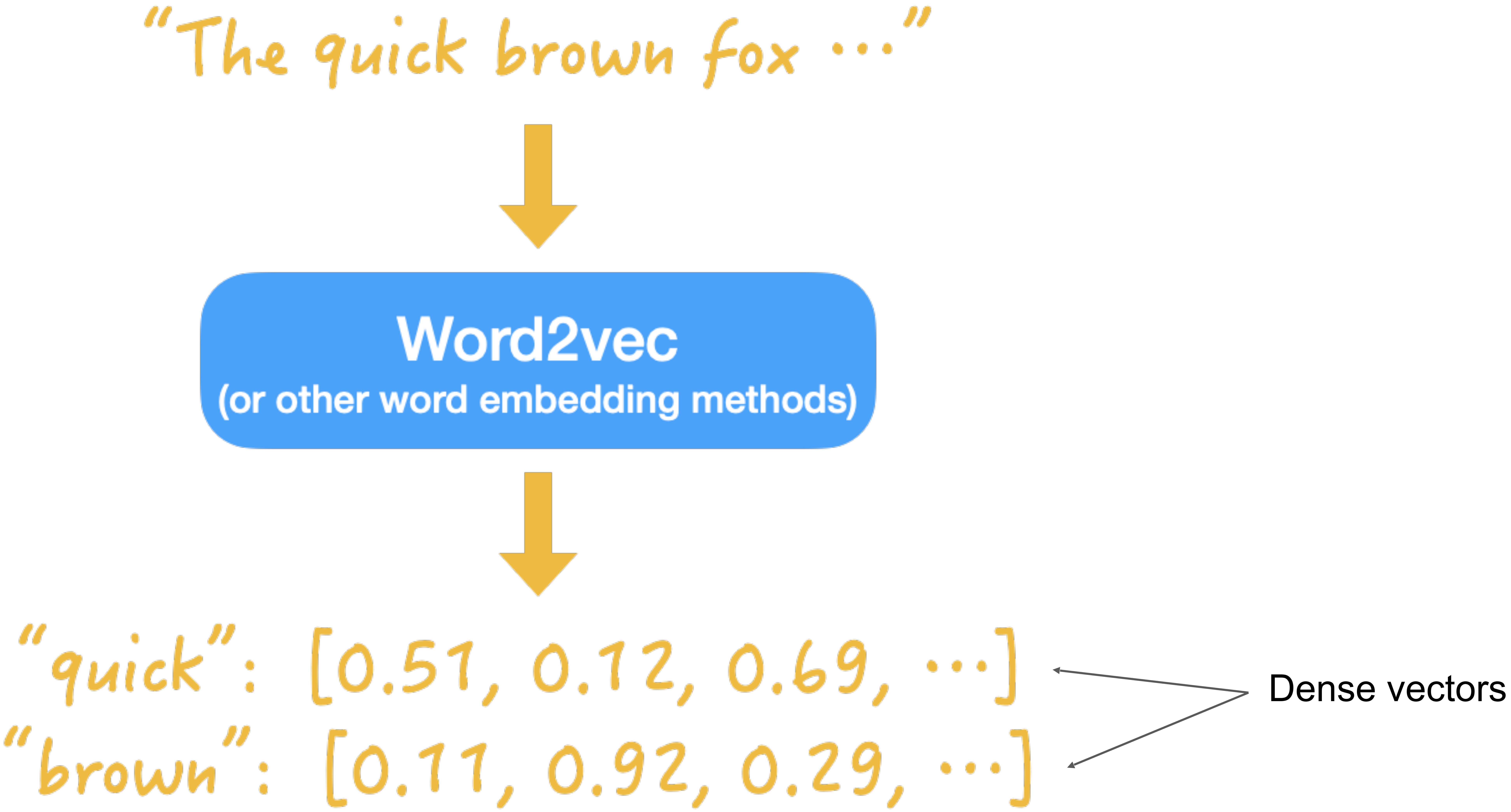
Cultural, linguistic, economic, and political distance are also important!

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. *(Under review, Nature)*

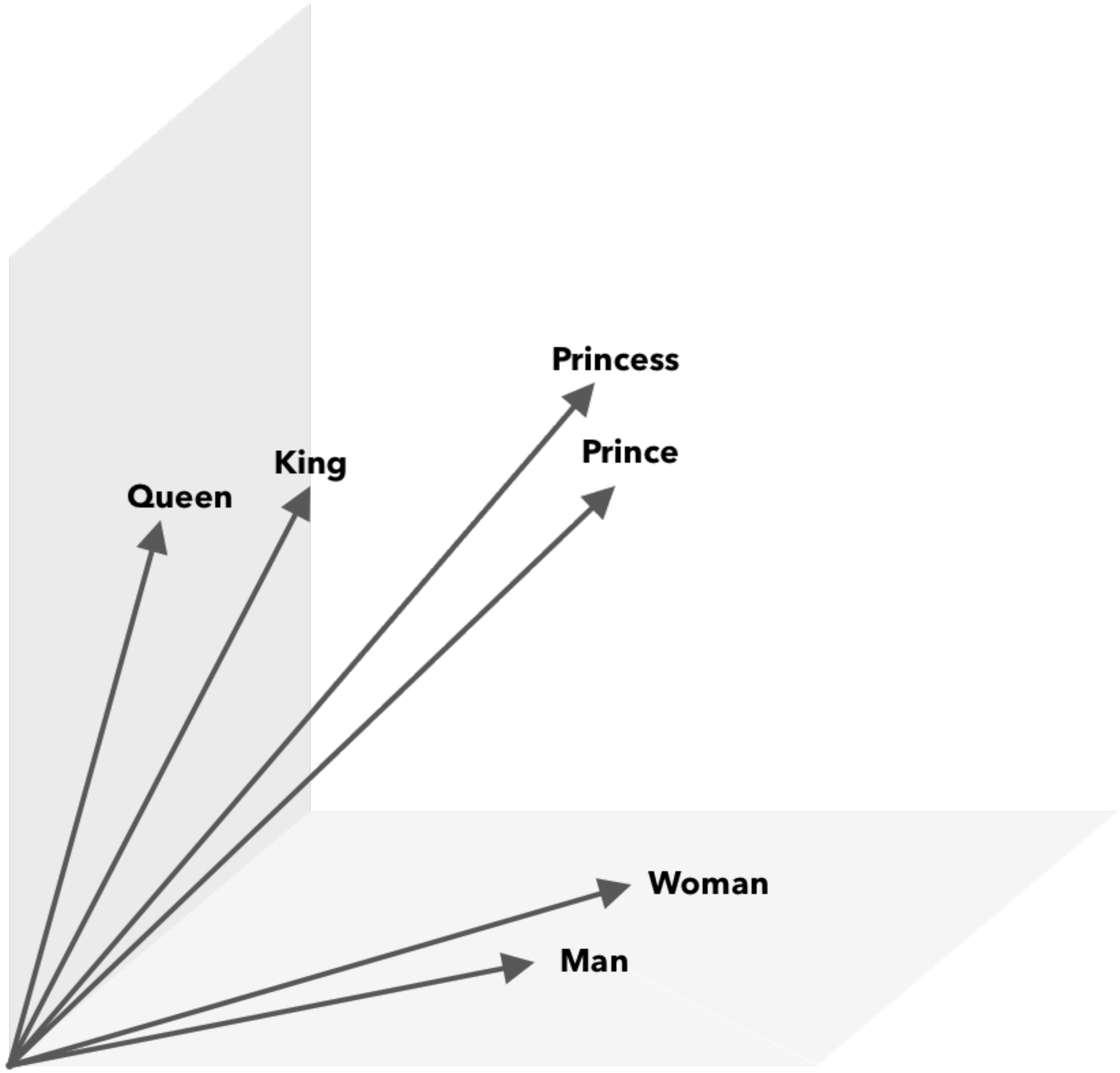
Neural embedding



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. *(Under review, Nature)*

The geometry of the vector space encodes semantic relationships

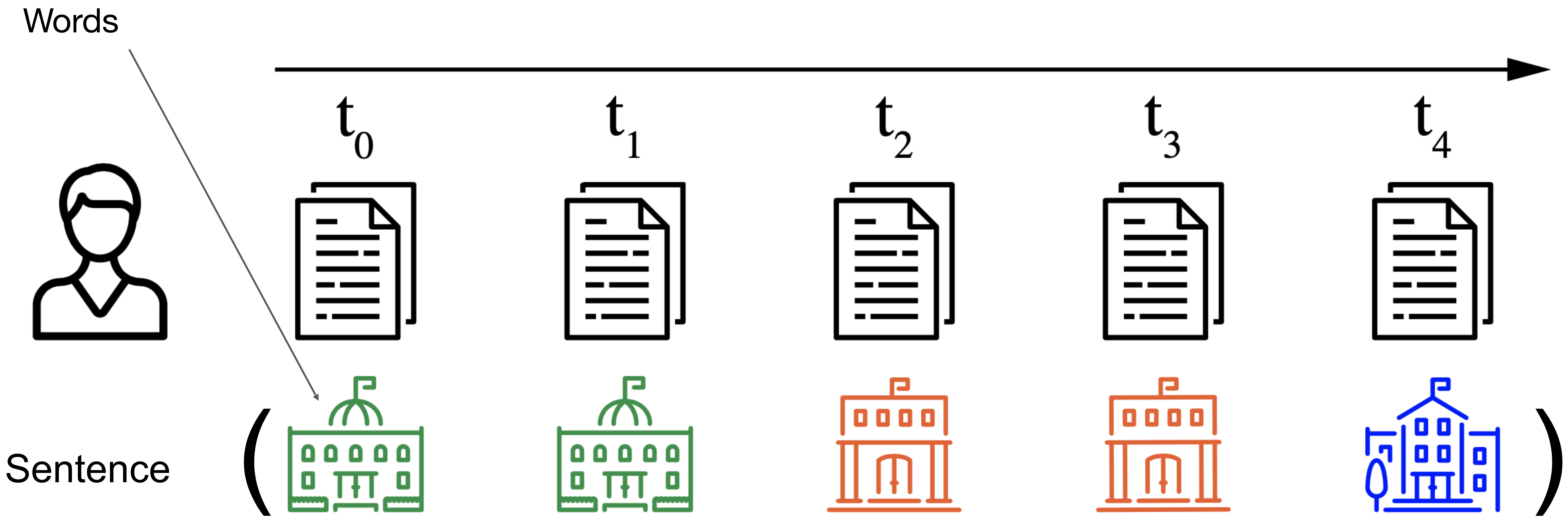
Using cosine distance



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. *(Under review, Nature)*

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

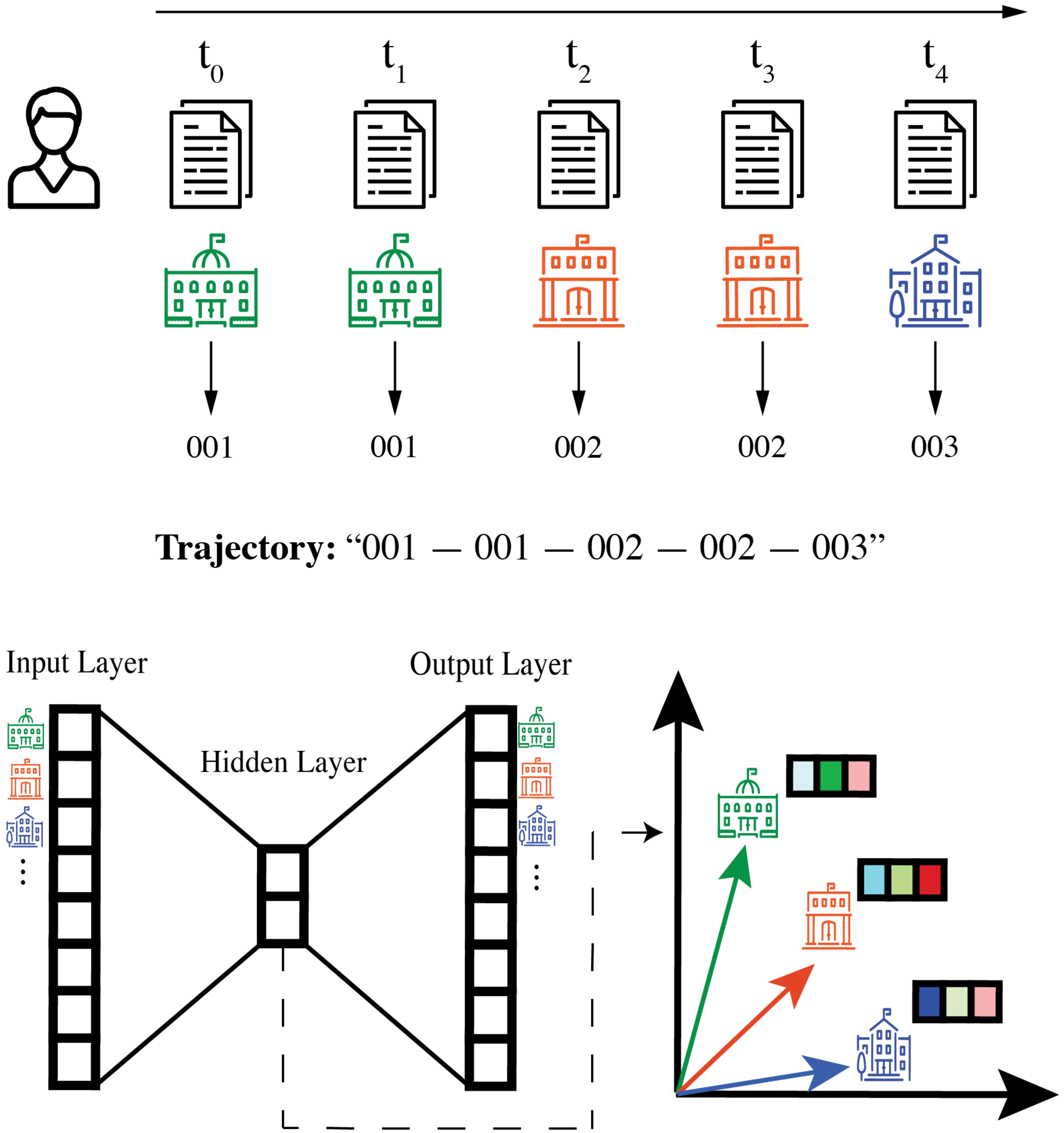
Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use **trajectories of scientists** as **sentences** and **organizations** as **words**.



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. *(Under review, Nature)*

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

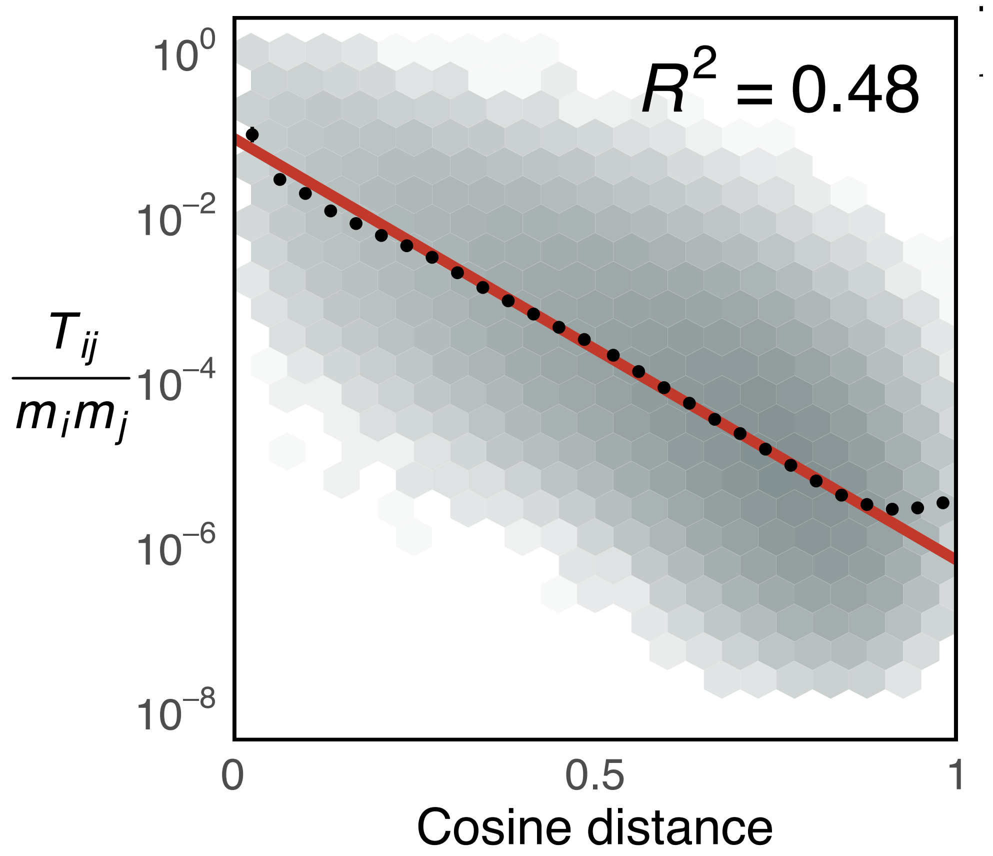


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

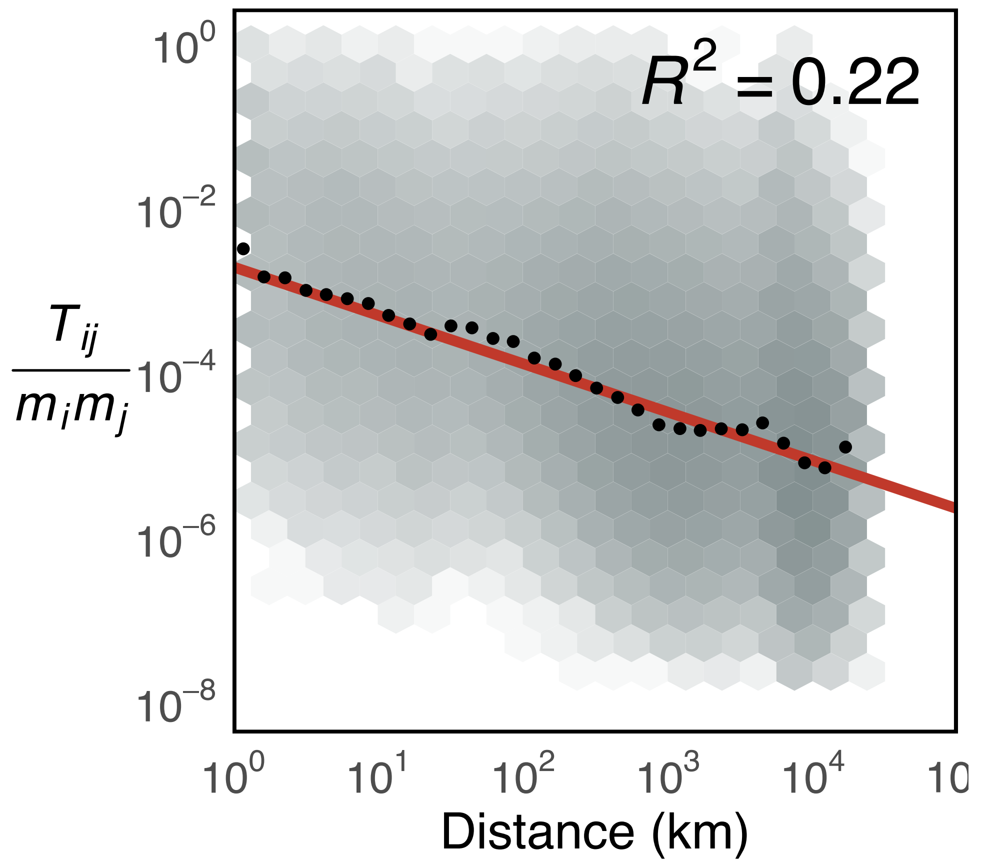
Embedding distance outperforms geographic distance

Correlation

Embedding Distance



Geographic Distance

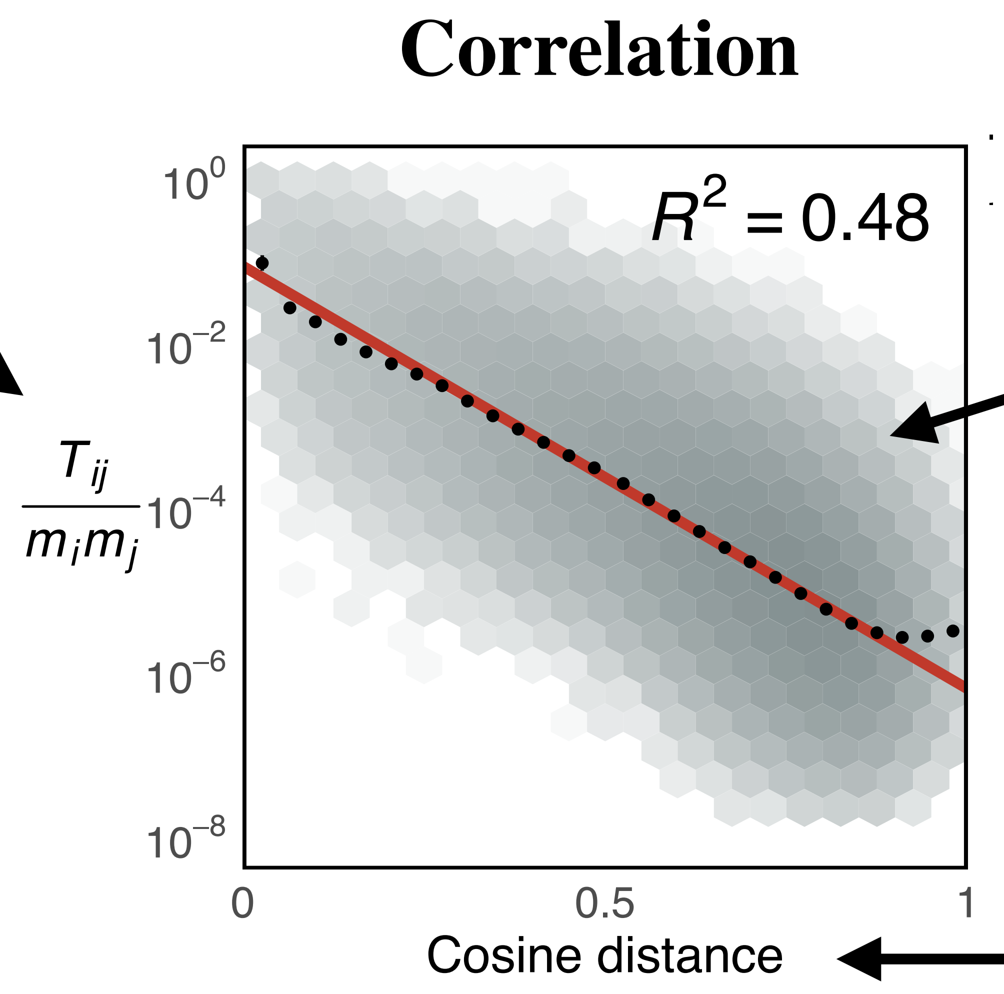


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

Embedding distance outperforms geographic distance

Flux, given the organizations sizes

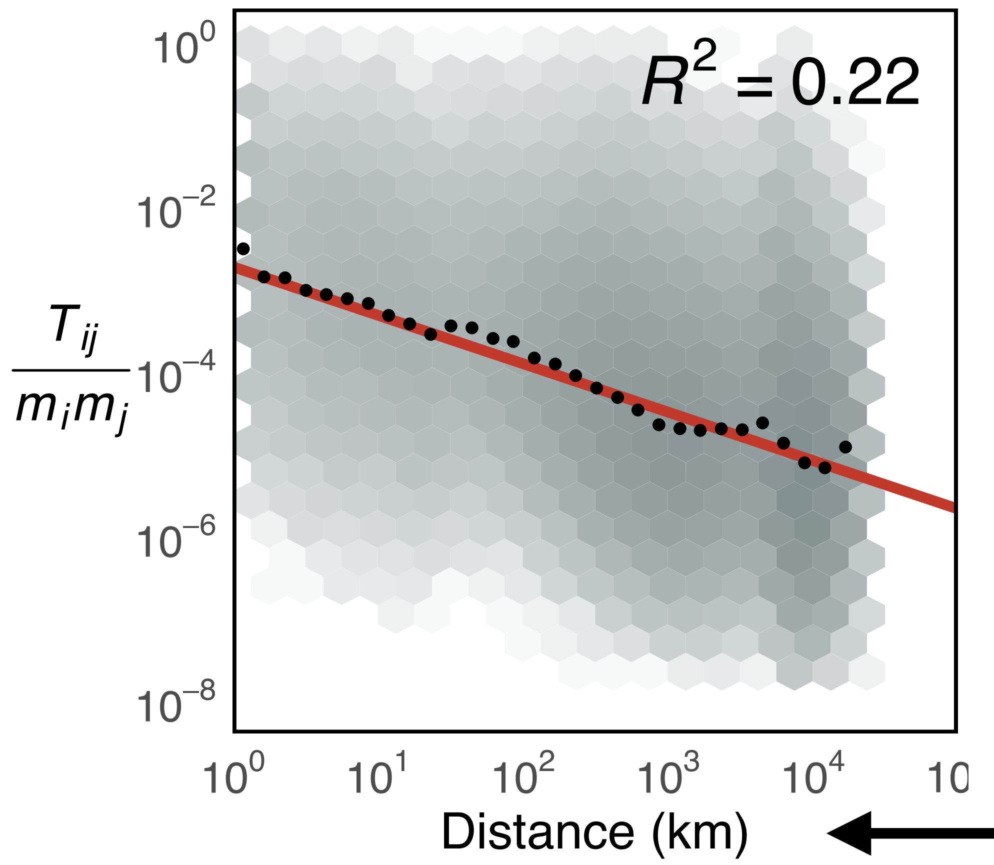
Embedding Distance



Every point is a pair of organizations (binned)

Cosine distance between organization vectors

Geographic Distance



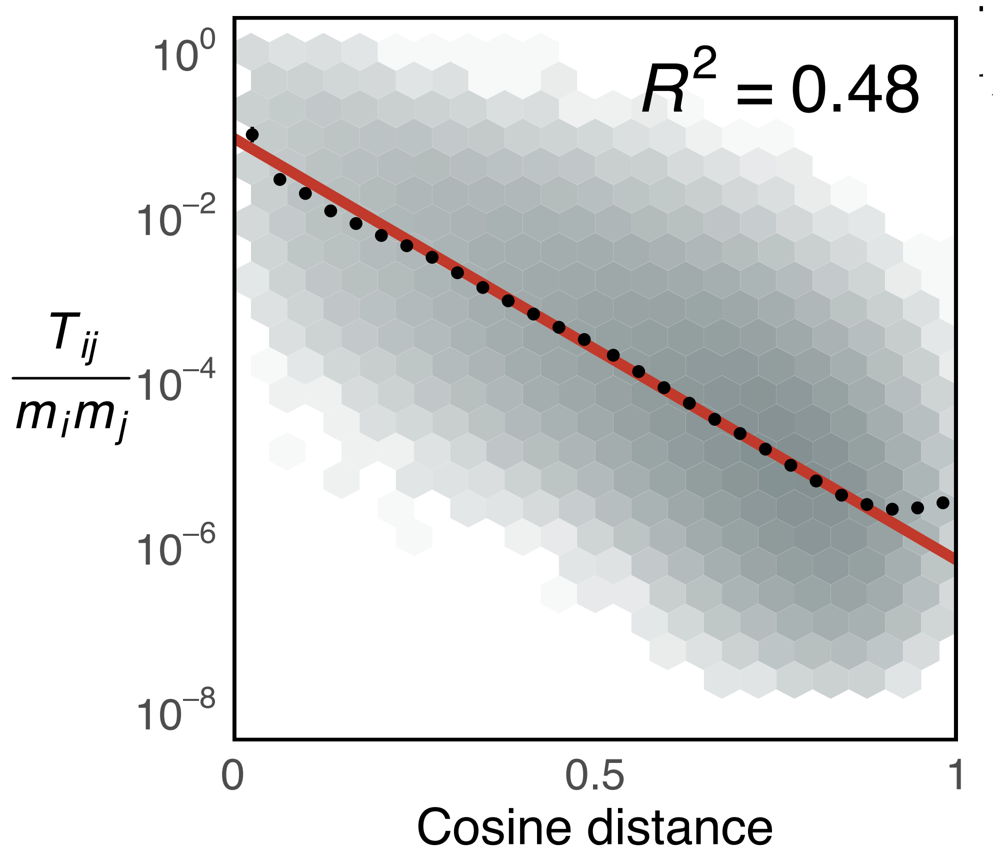
Geographic distance between organizations

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. *(Under review, Nature)*

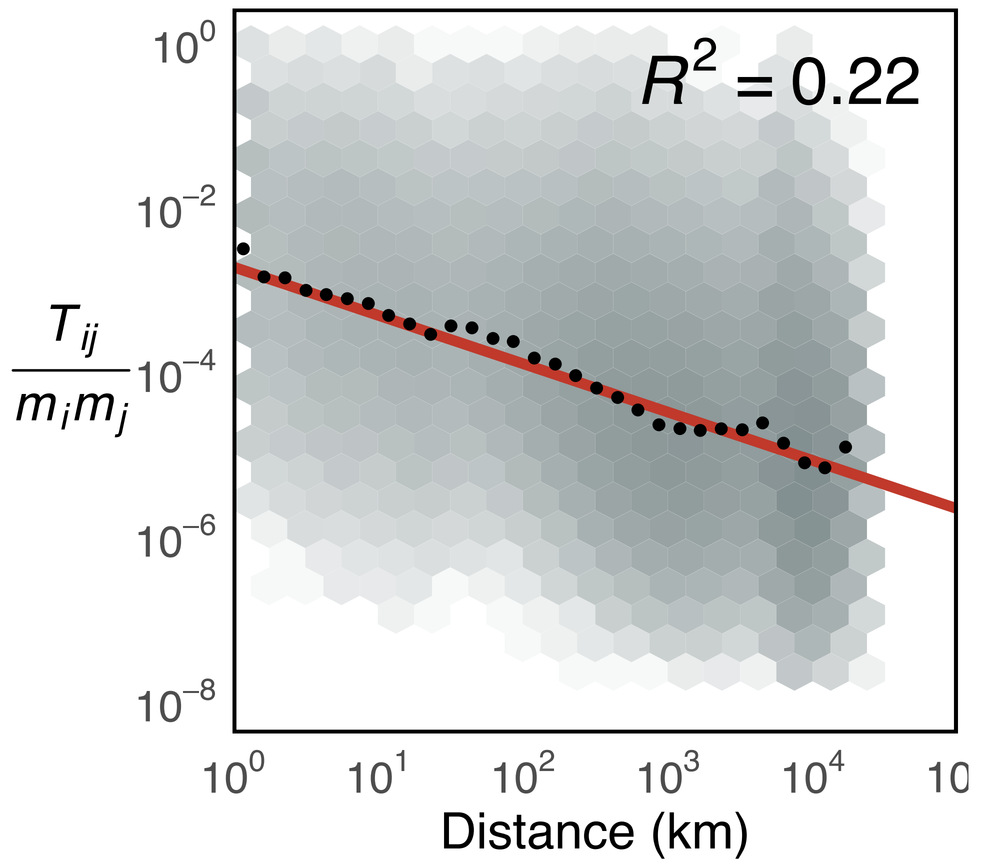
Embedding distance outperforms geographic distance

Correlation

Embedding Distance



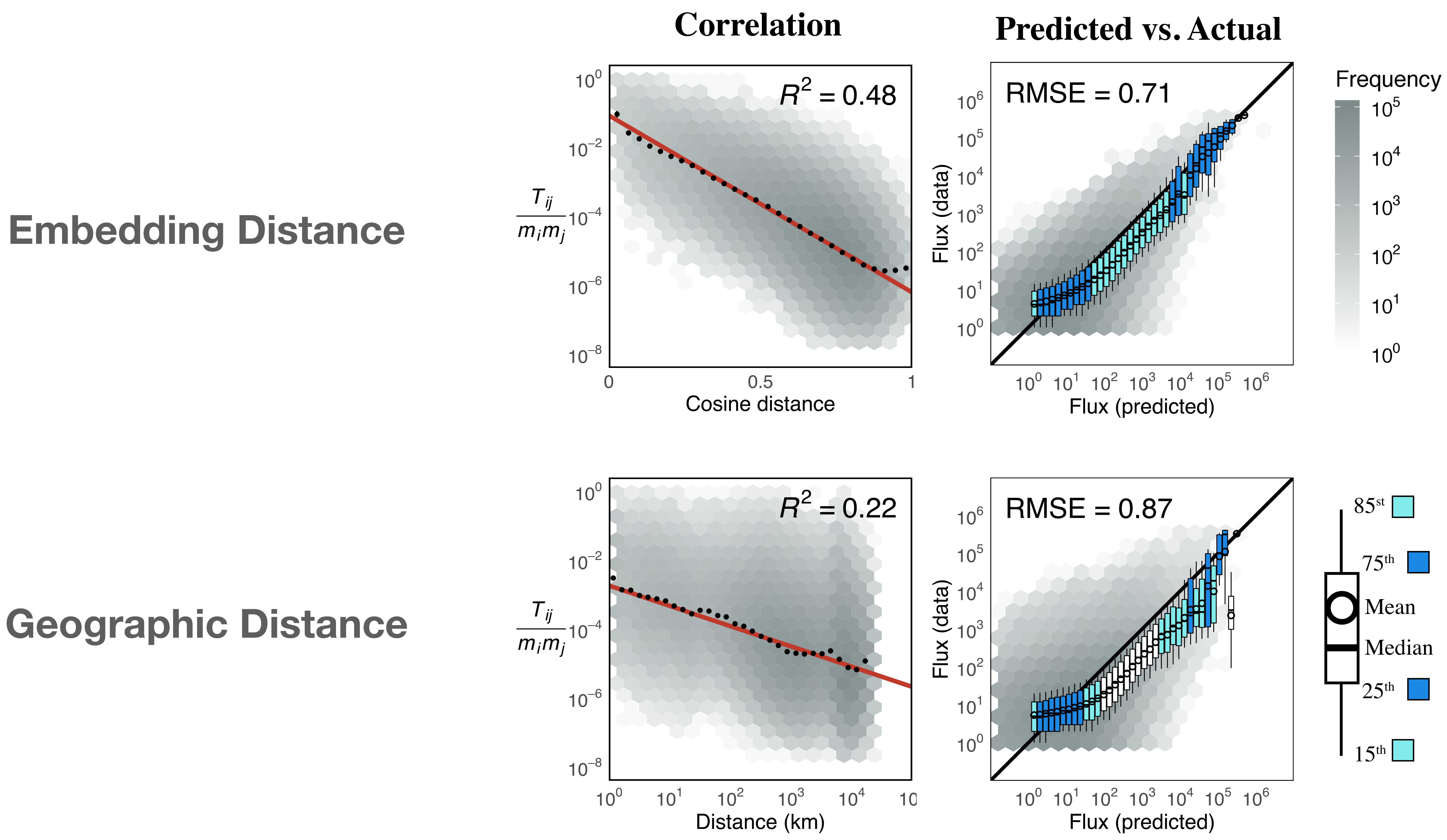
Geographic Distance



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. *(Under review, Nature)*

Embedding distance outperforms geographic distance



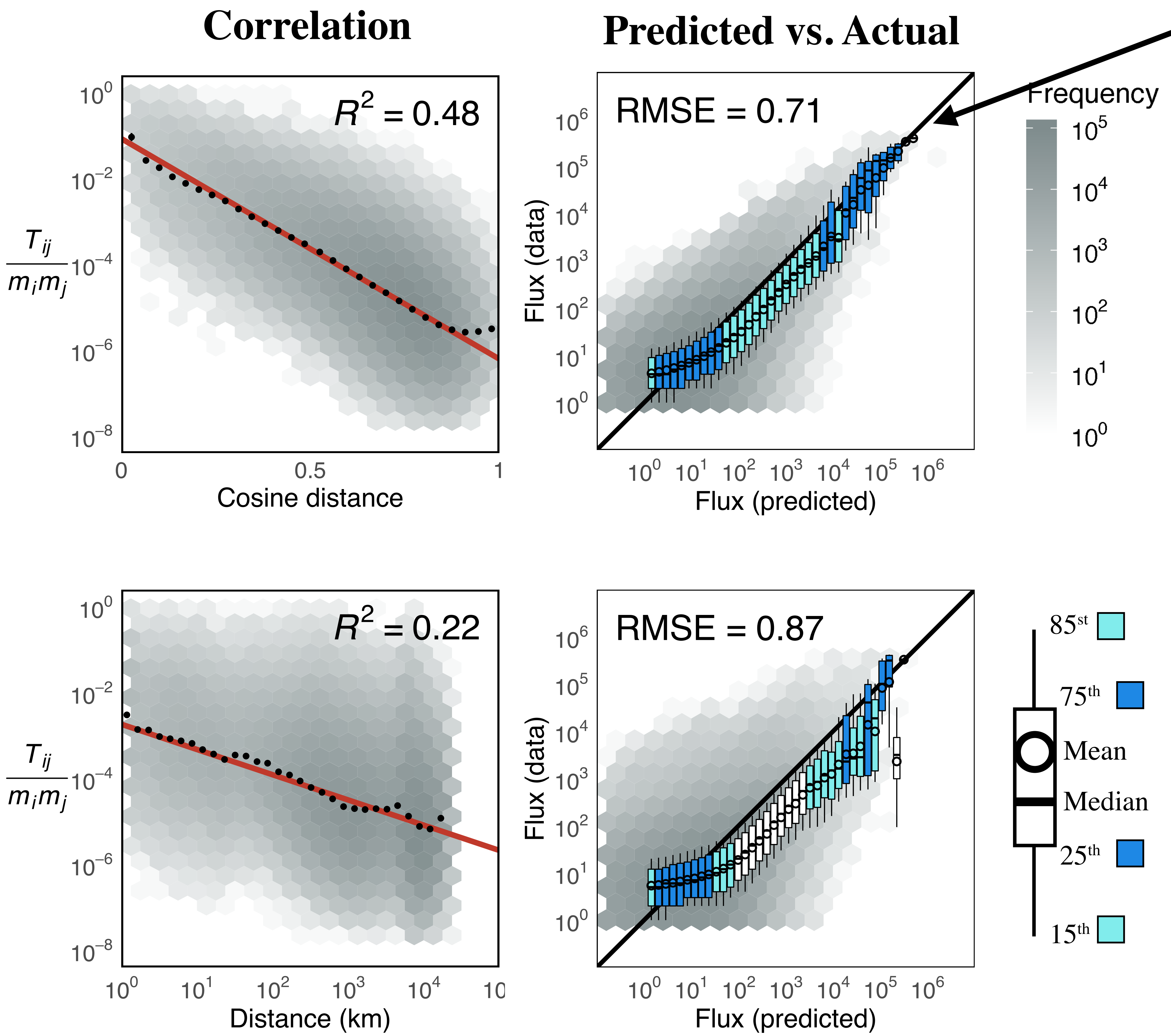
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

Embedding distance outperforms geographic distance

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

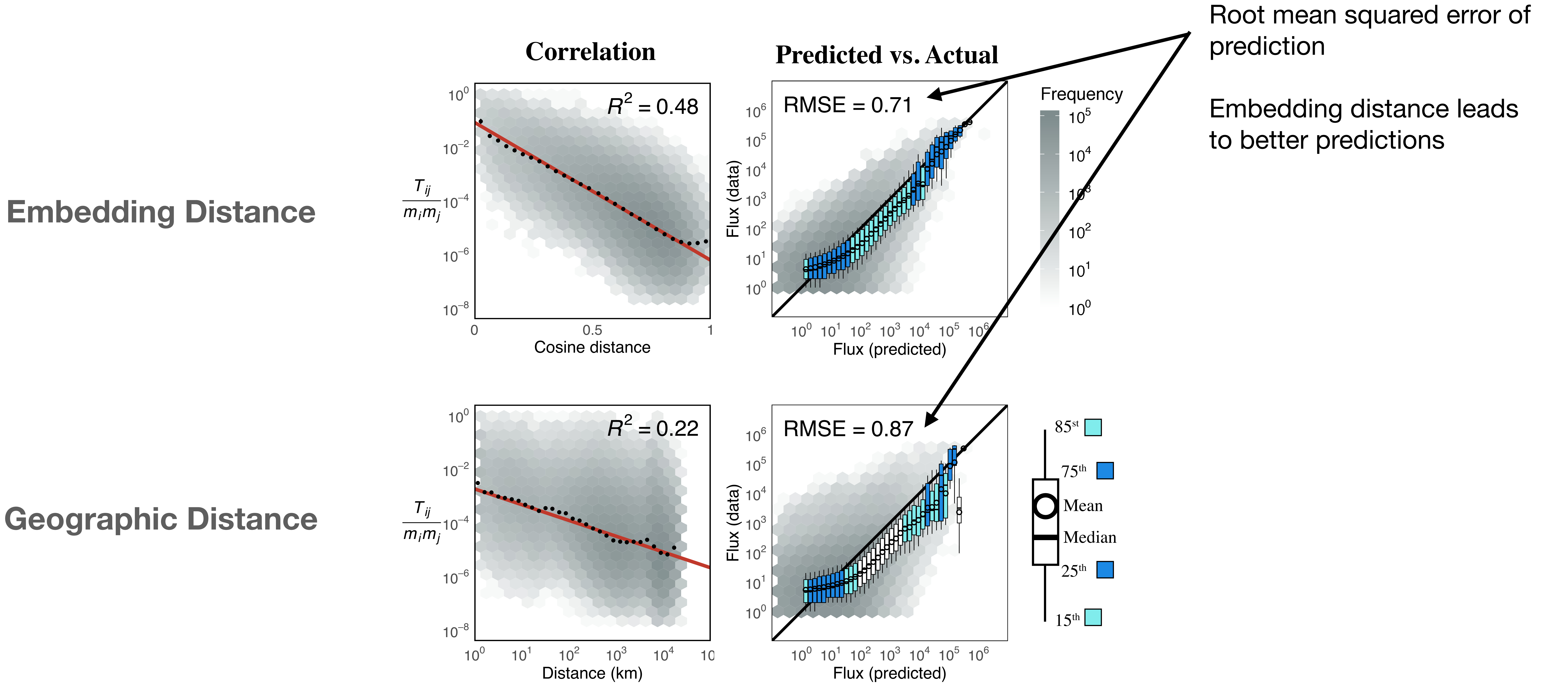
Embedding Distance

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. *(Under review, Nature)*

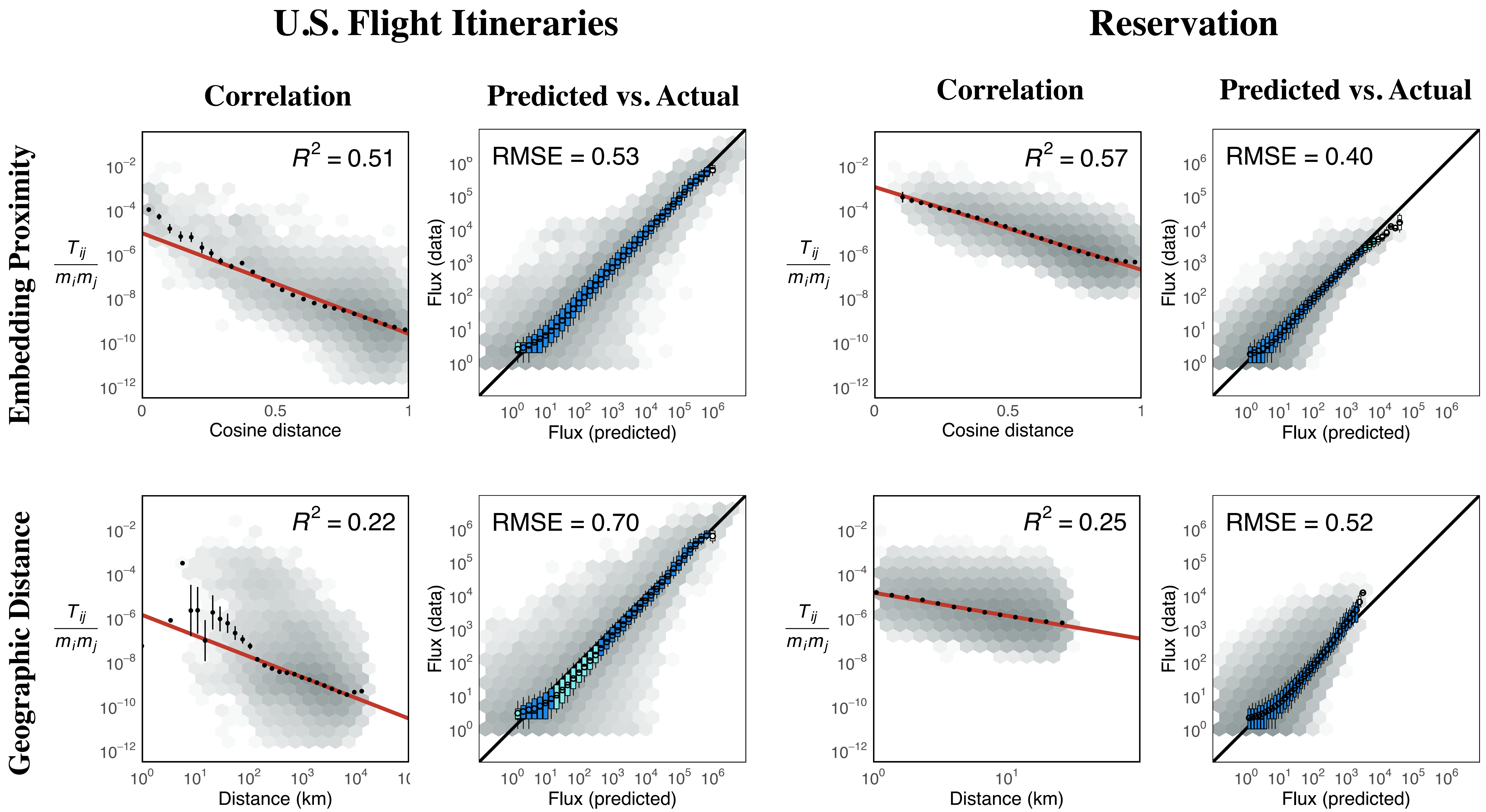
Embedding distance outperforms geographic distance



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. *(Under review, Nature)*

The embedding performs well in other domains!

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

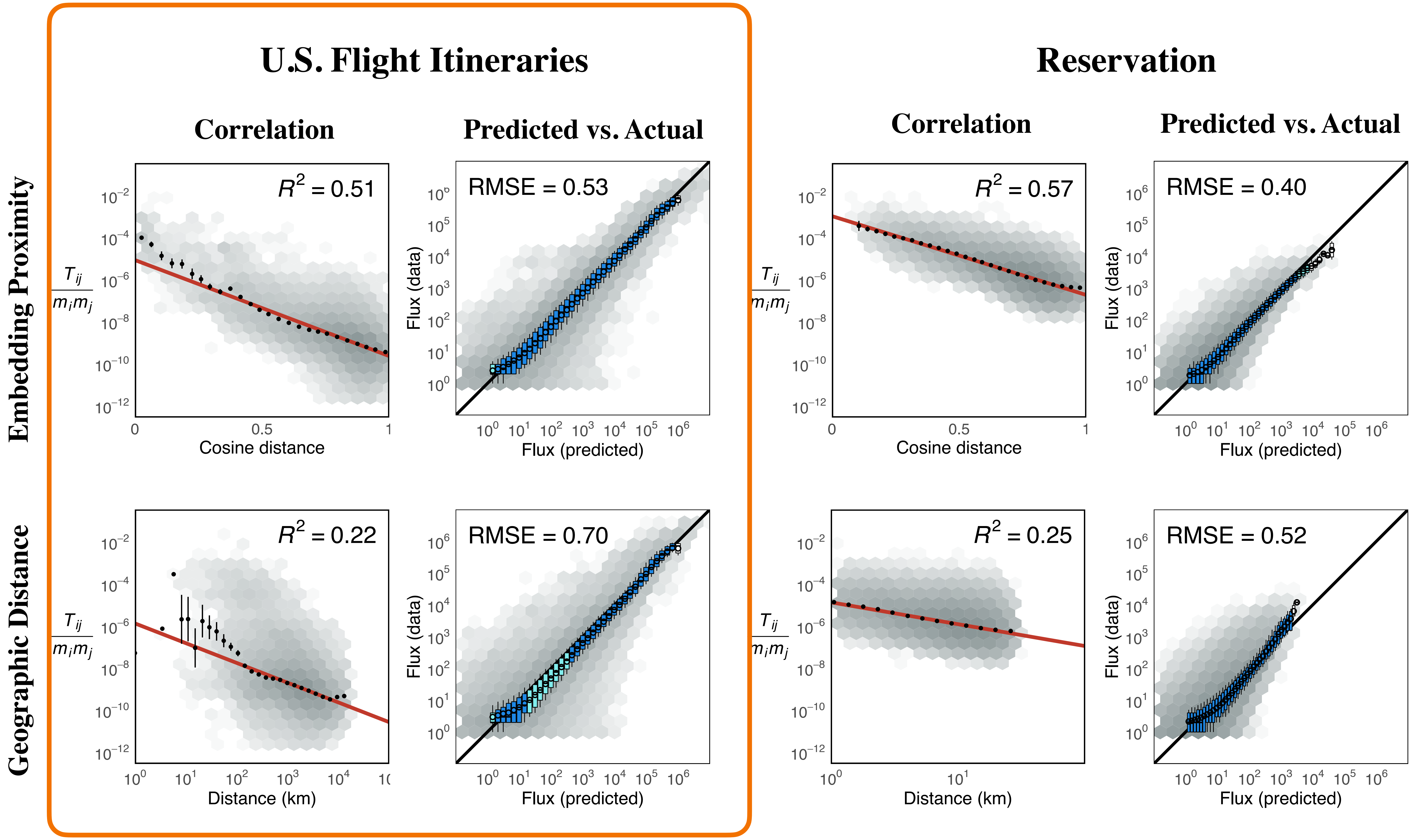


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

The embedding performs well in other domains!

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

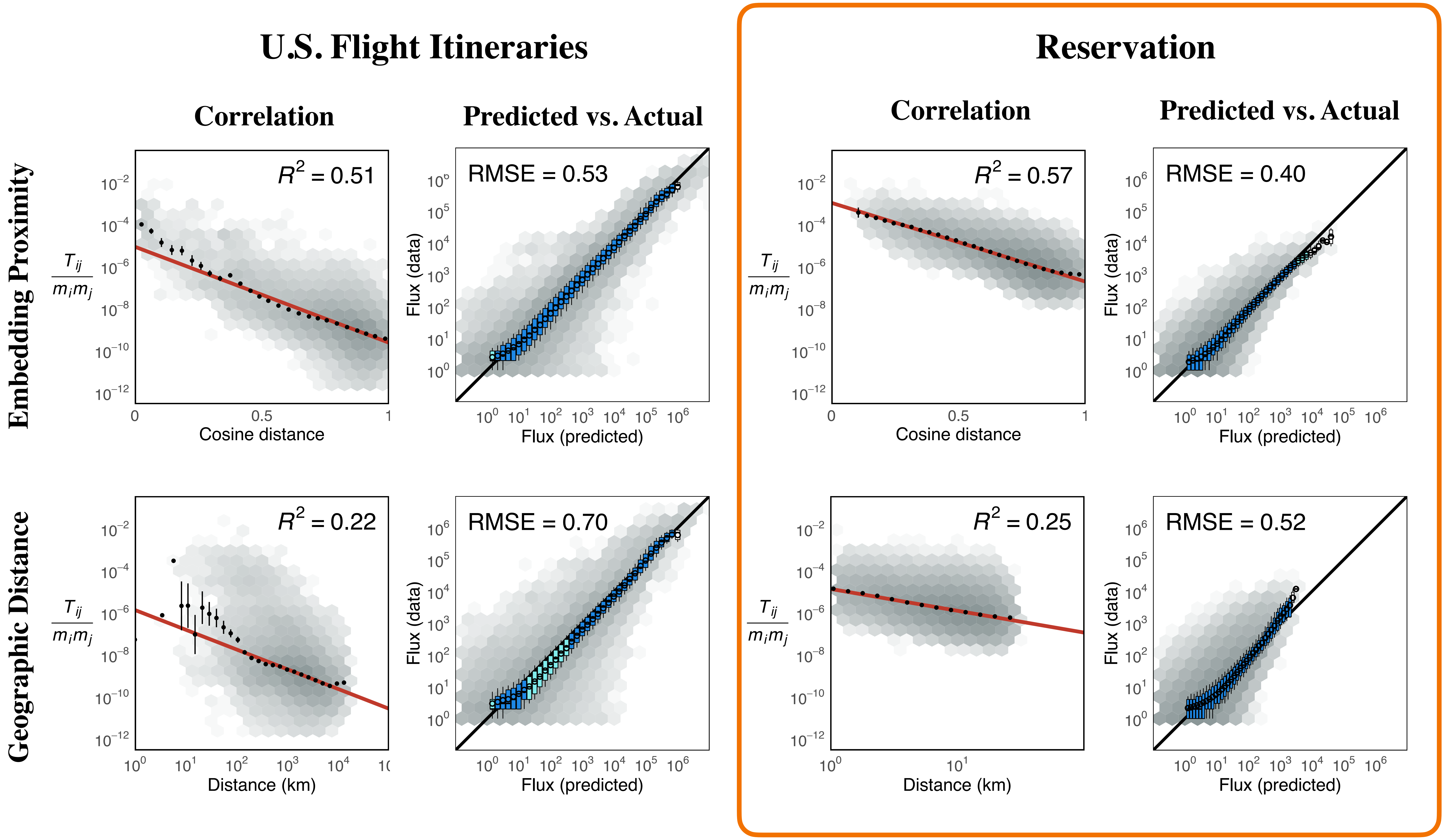
Better performance using embedding distance on trajectories derived from U.S. domestic flights



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. *(Under review, Nature)*

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. *(Under review, Nature)*

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)*

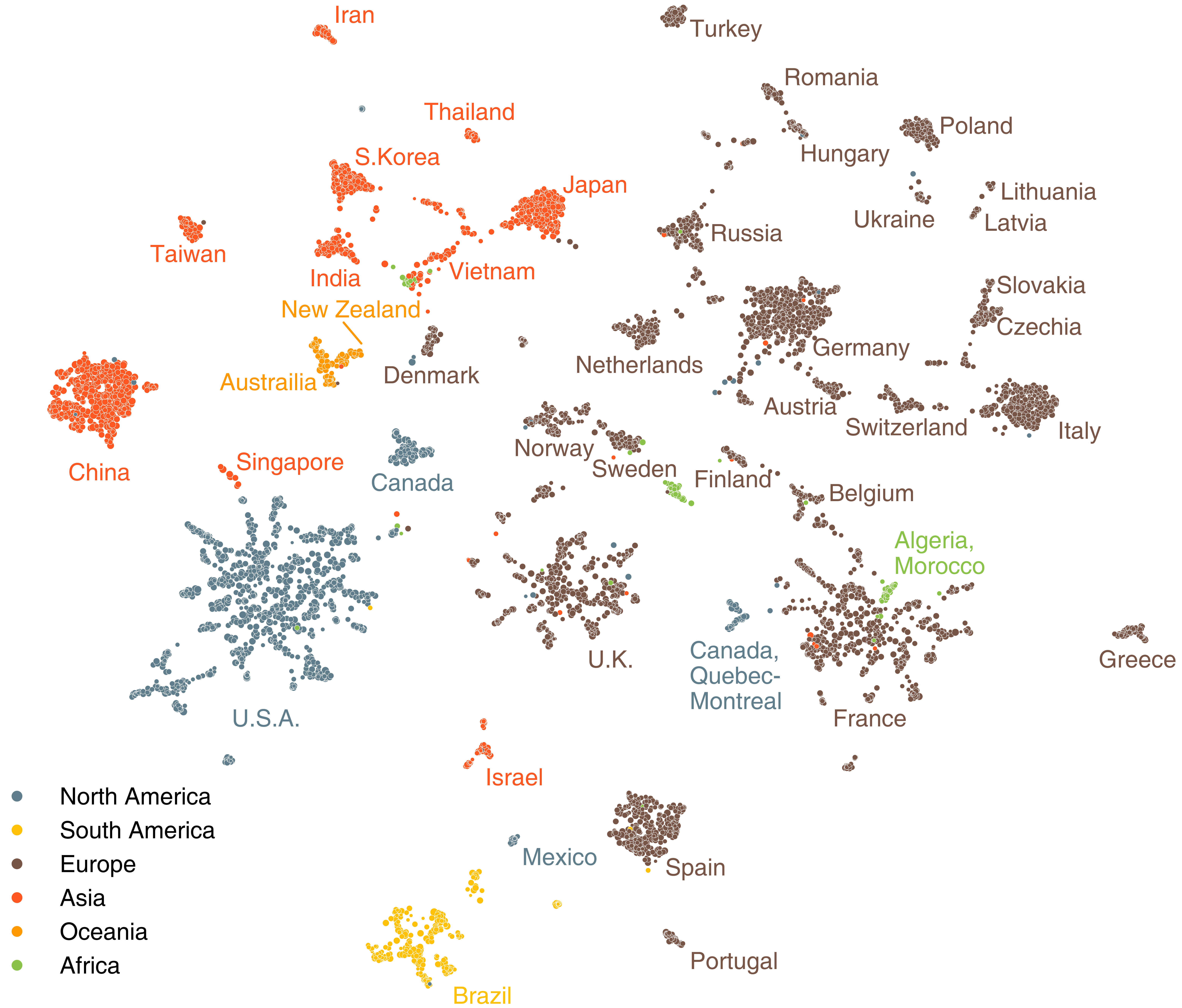
word2vec captures the latent structure of mobility

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

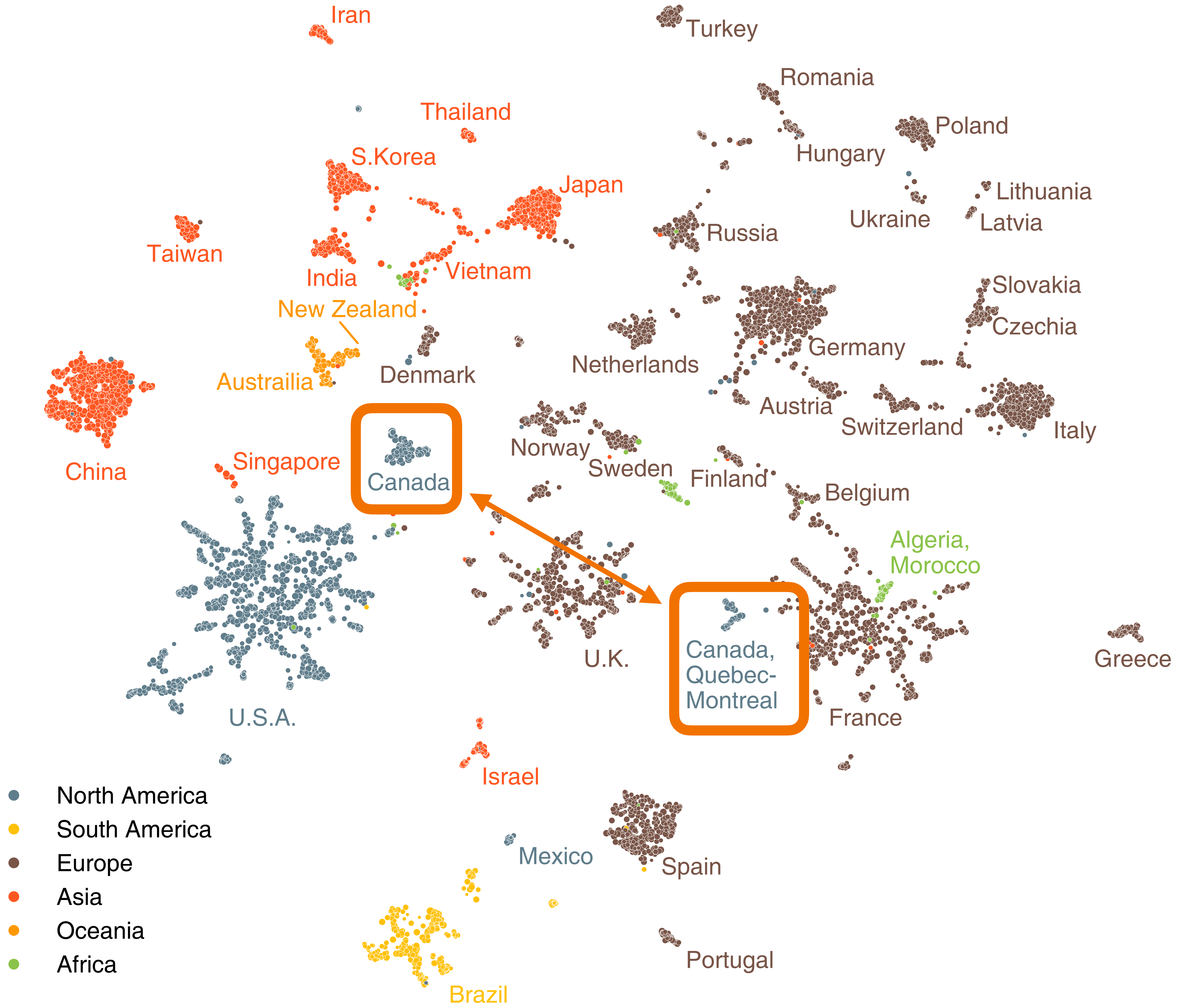
UMAP projection of organizations



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)*.

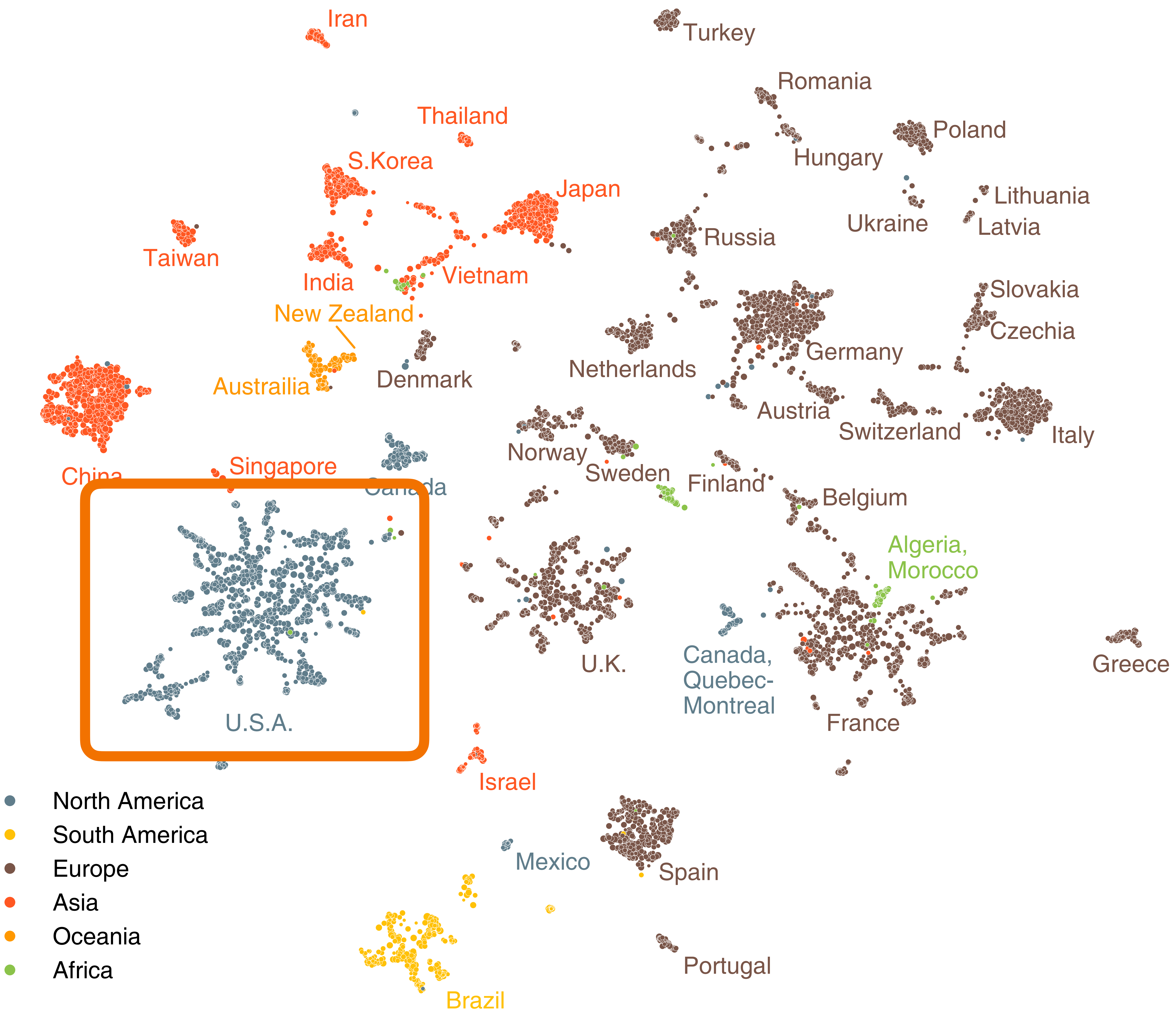
Visualizing the embedding space

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).*

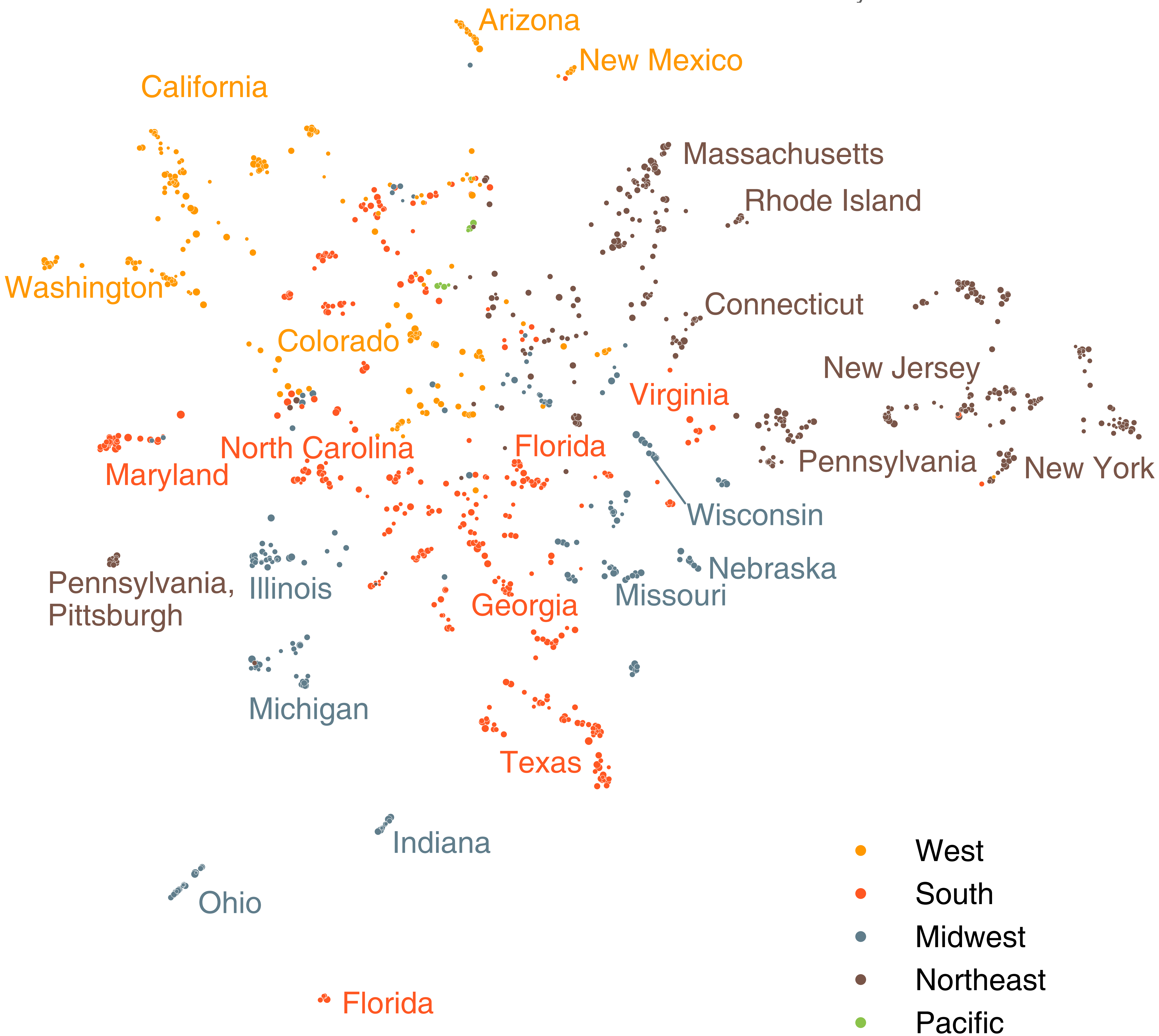
Visualizing the embedding space

We can “zoom in”



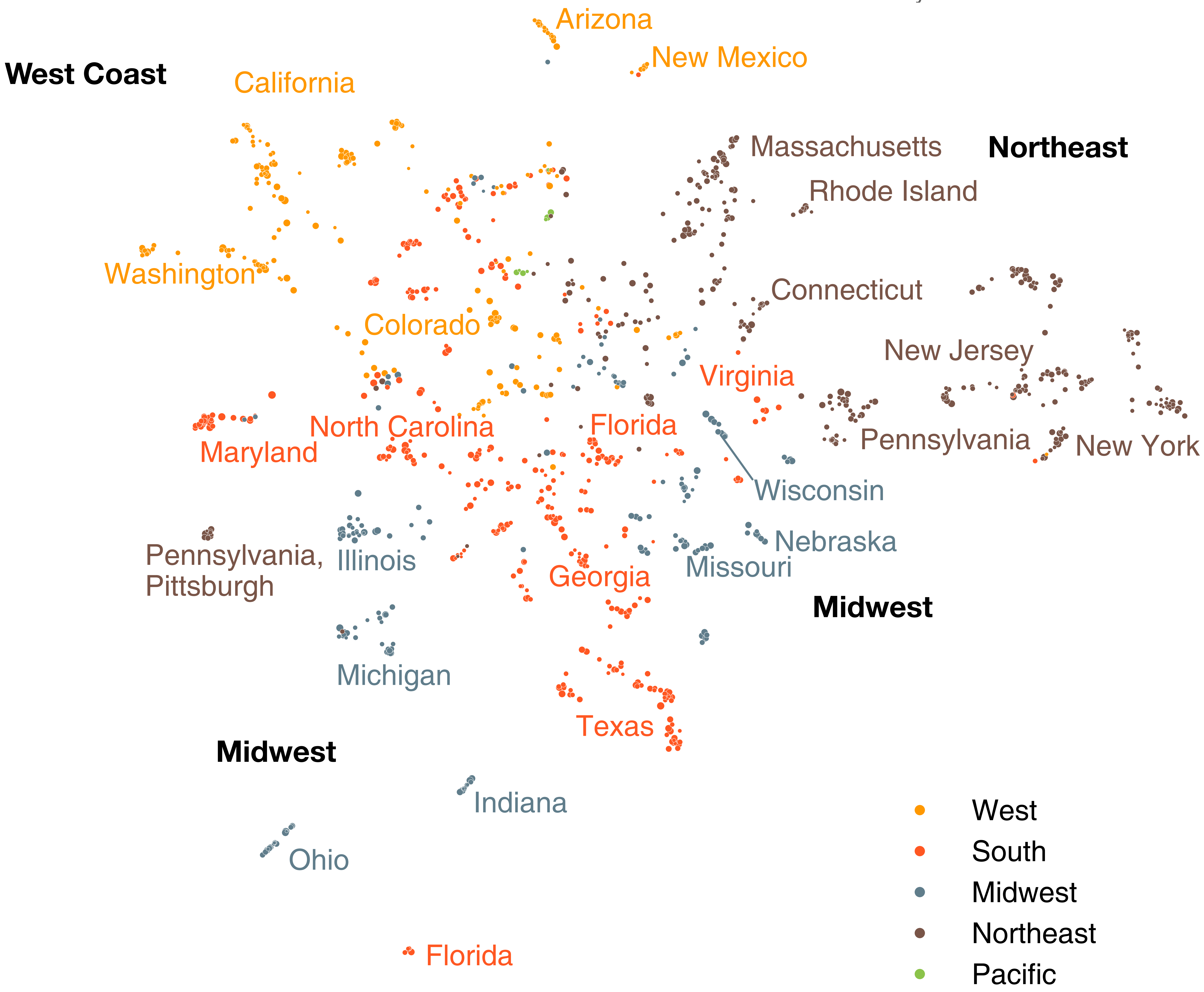
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

United States



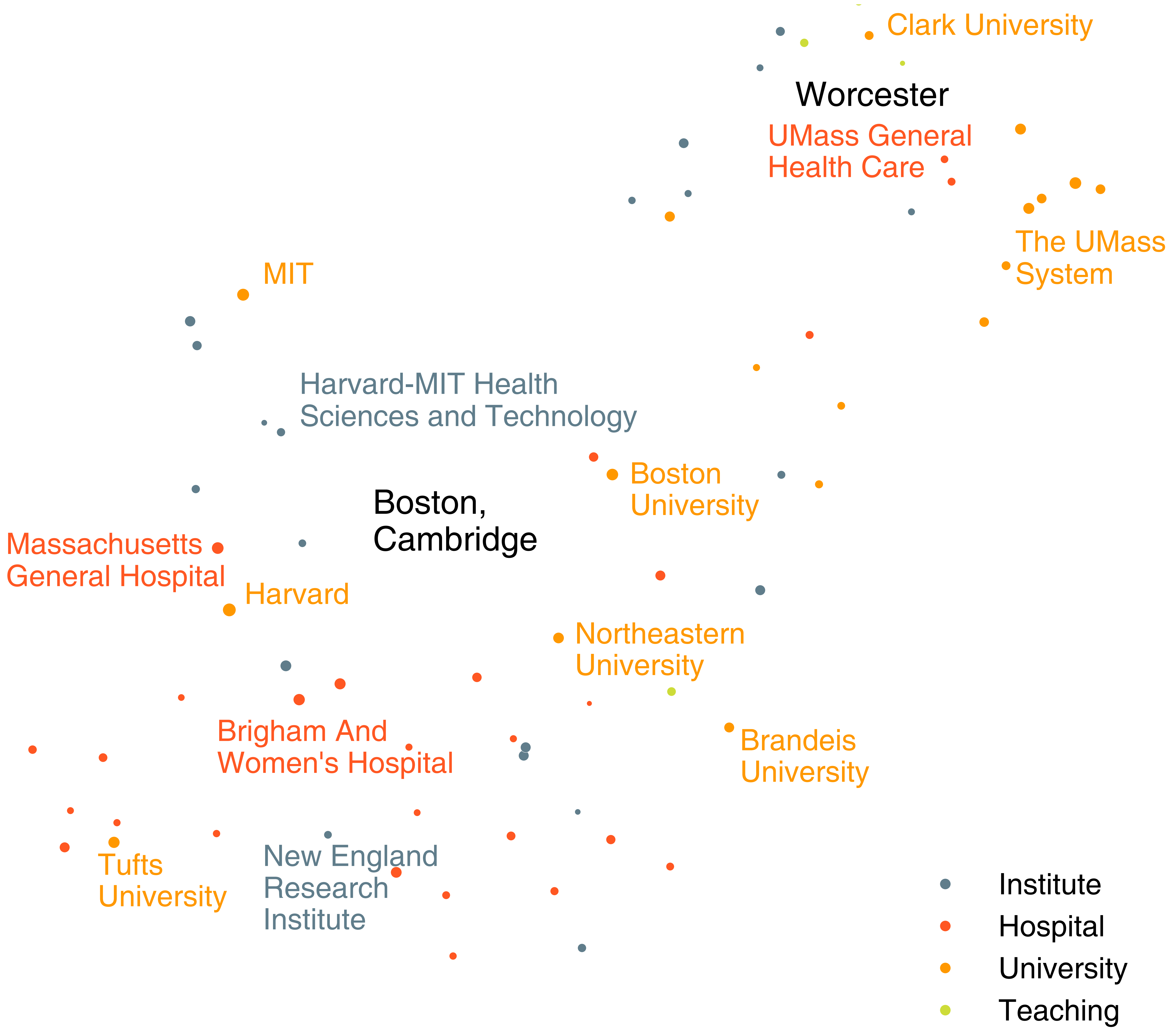
Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

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)*.

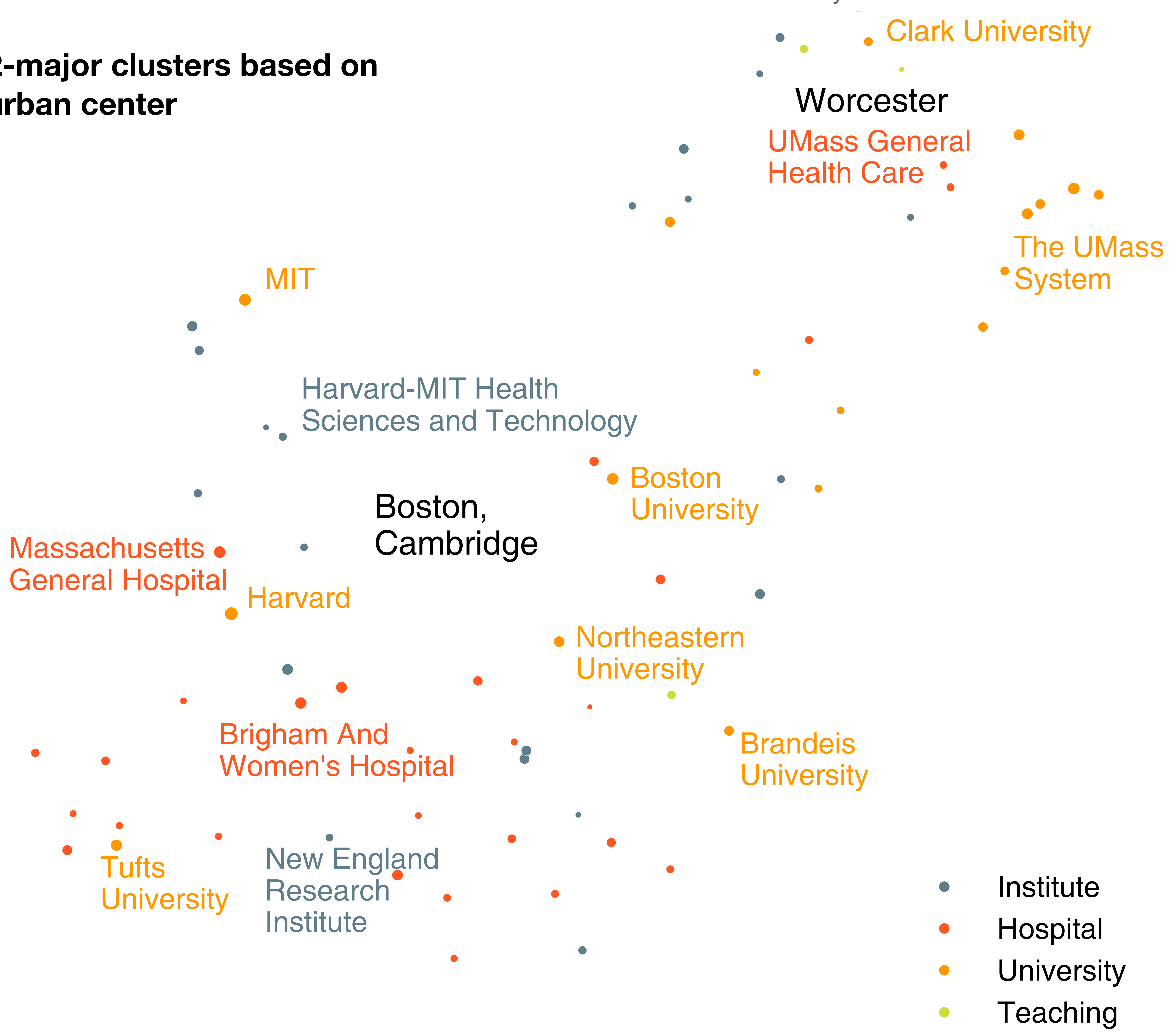
Massachusetts



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)*.

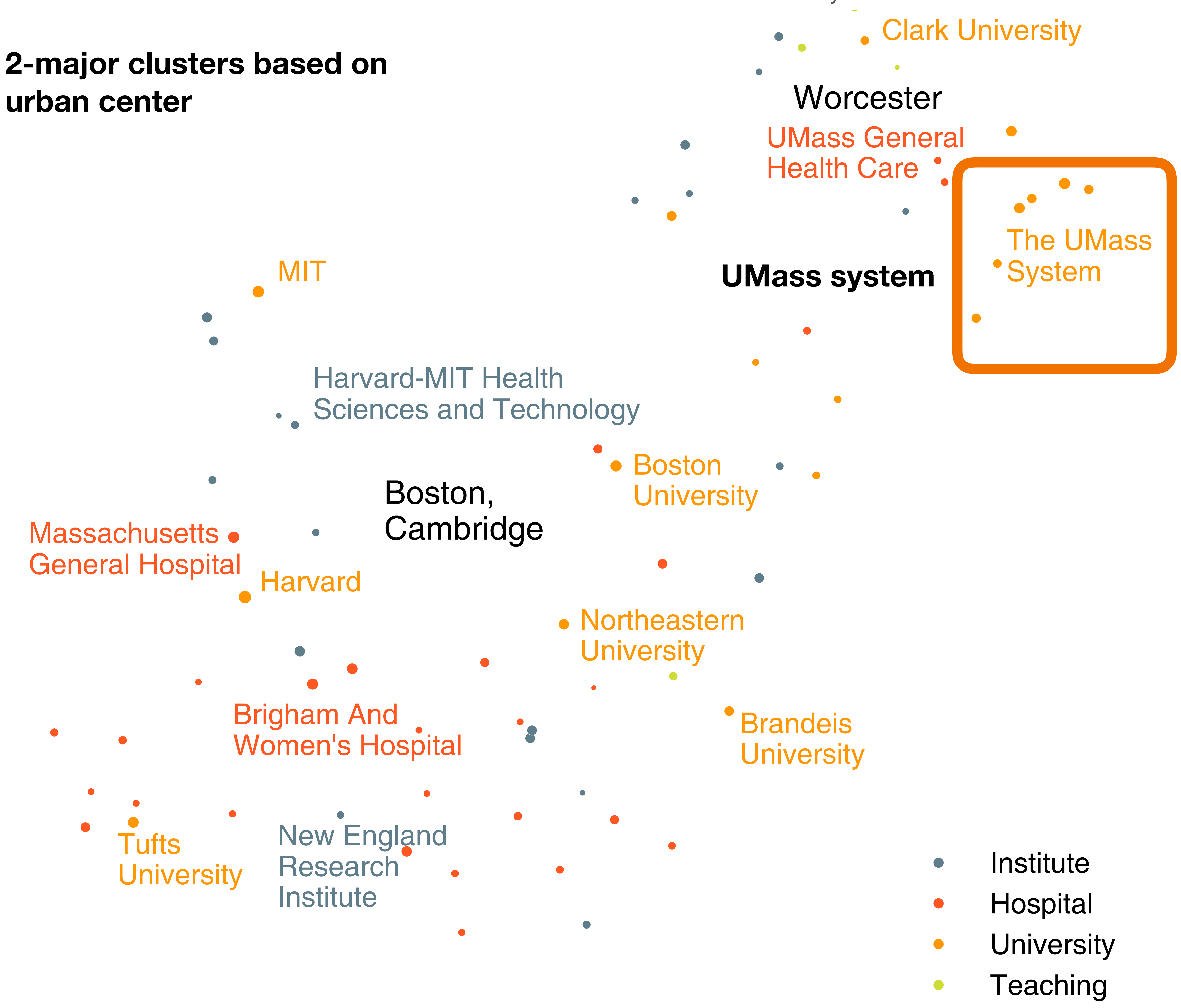
Massachusetts

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

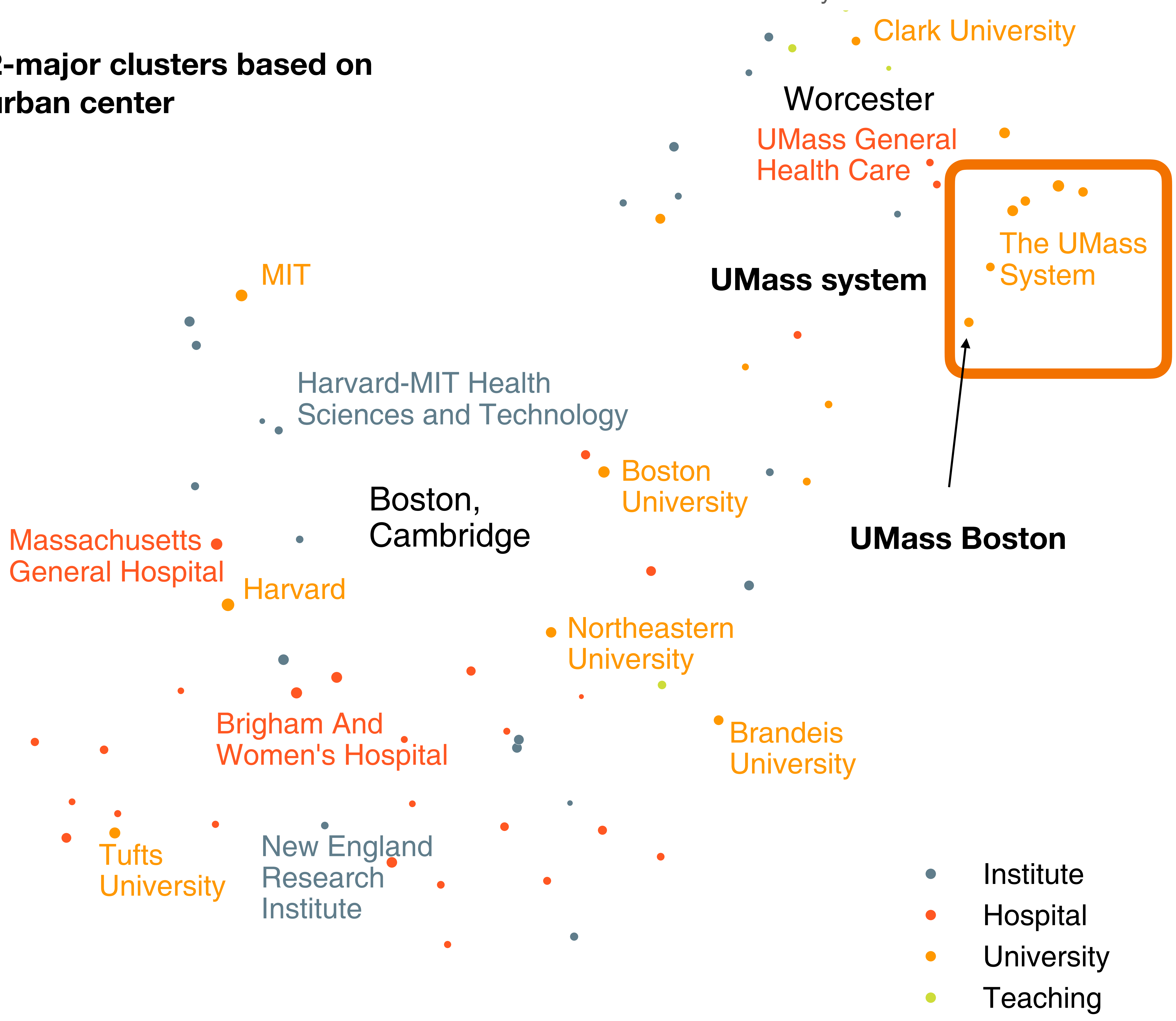


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*.

dakotamurray.me/talks/2020-northeastern/

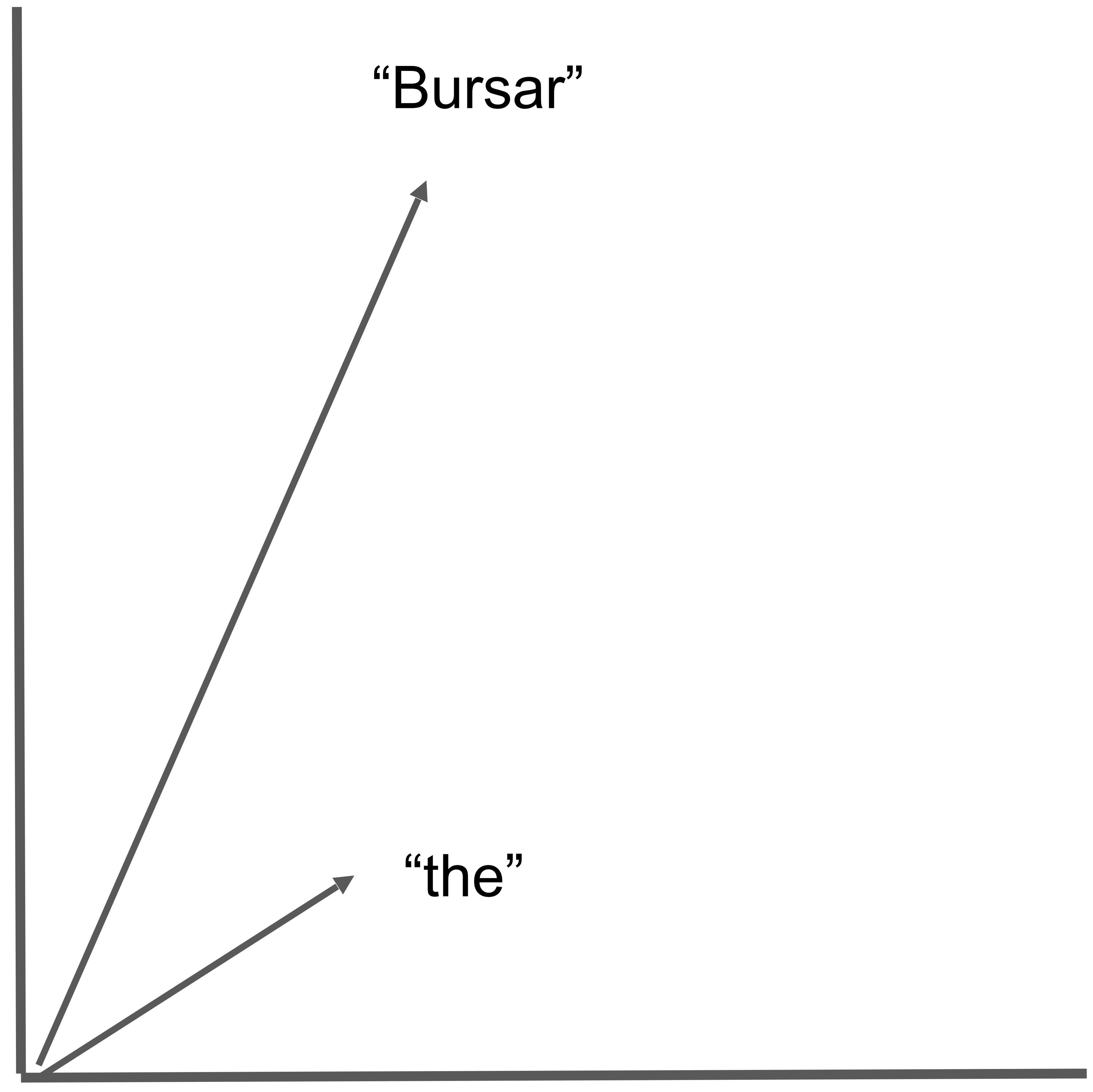
Massachusetts

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)*.

What else is encoded? Vector length

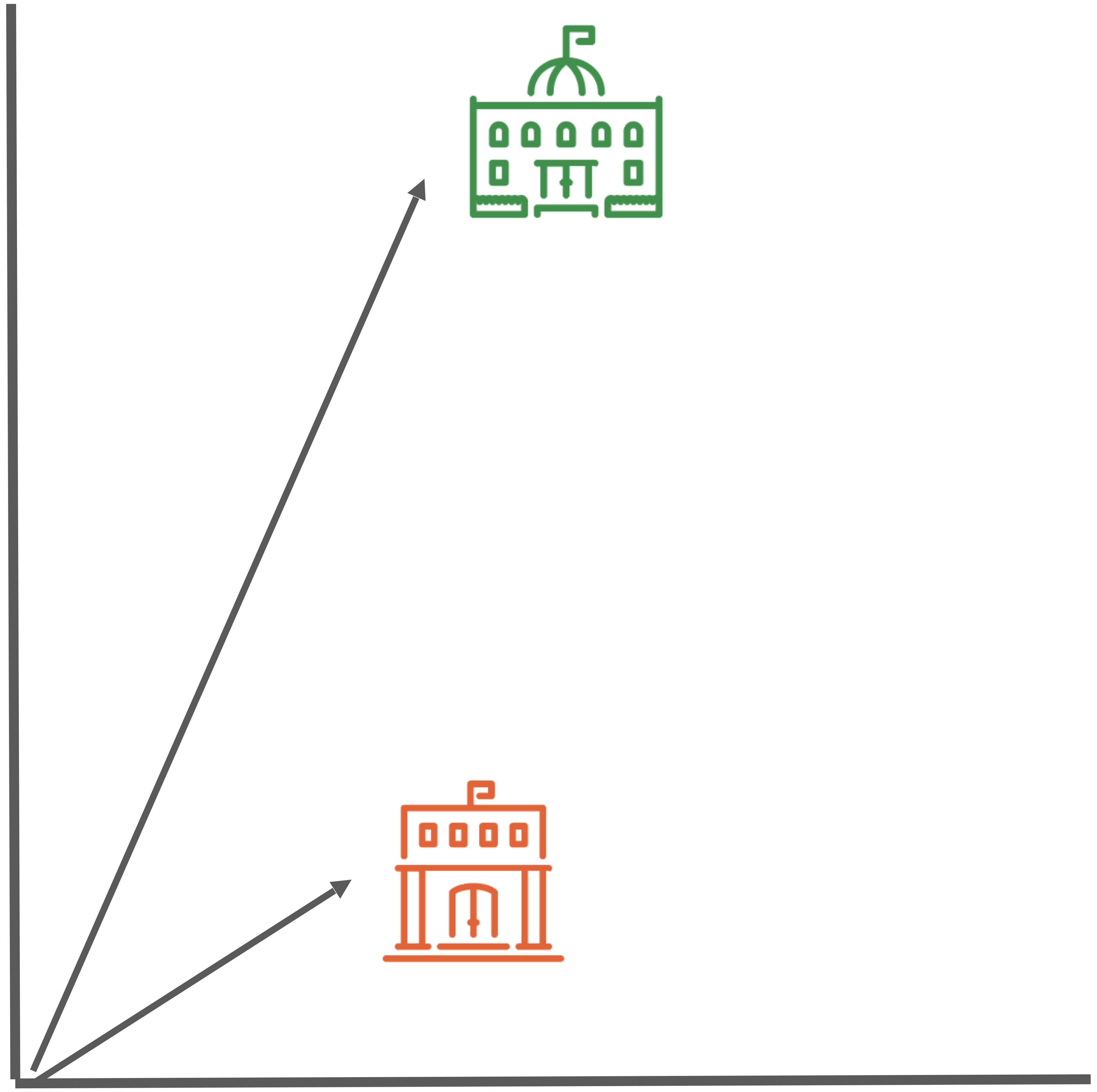


- In word embeddings, larger vectors (by magnitude) tend to appear in a single context

- 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



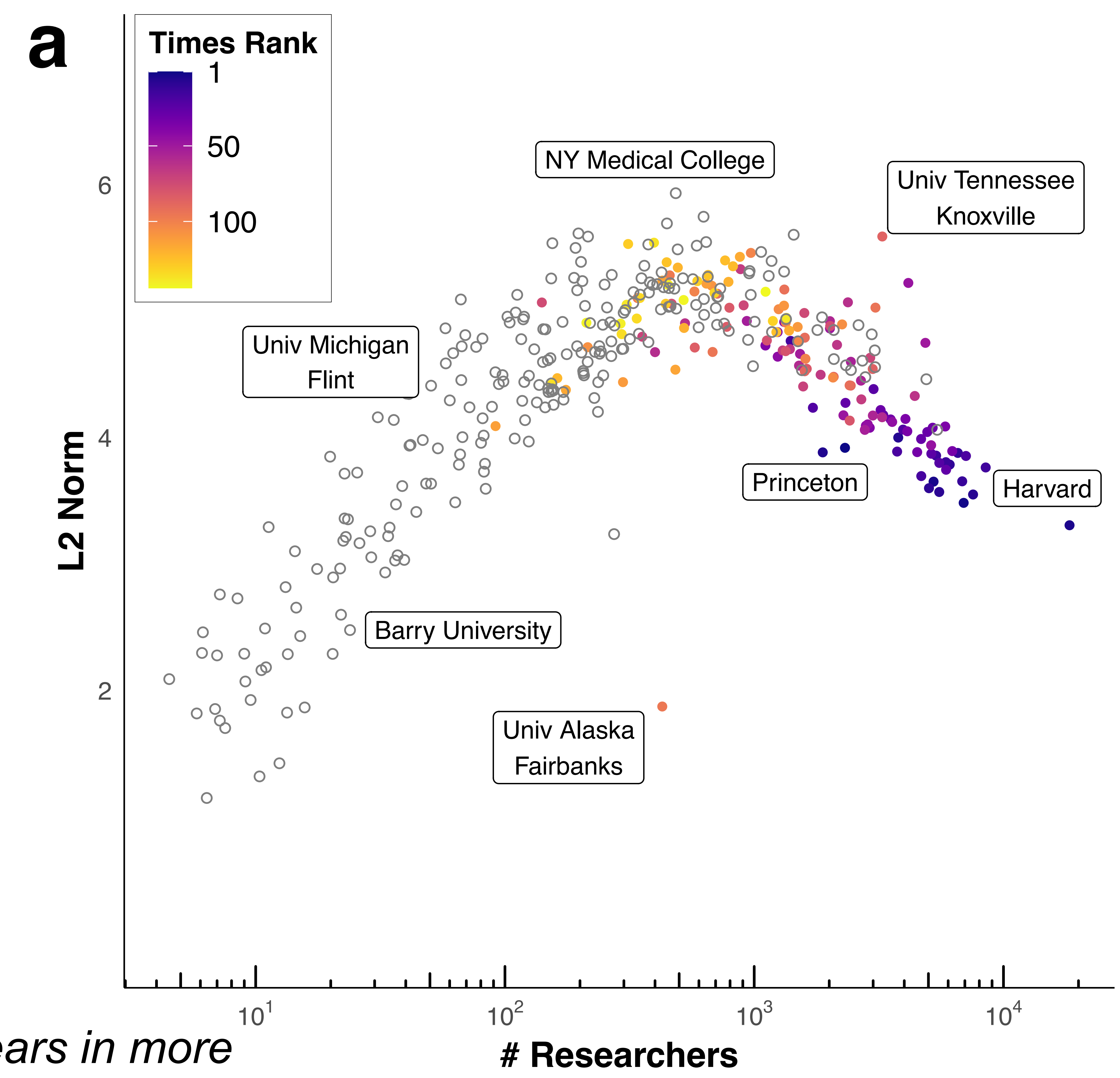
- In word embeddings, larger vectors (by magnitude) tend to appear in a single context

- 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)*.

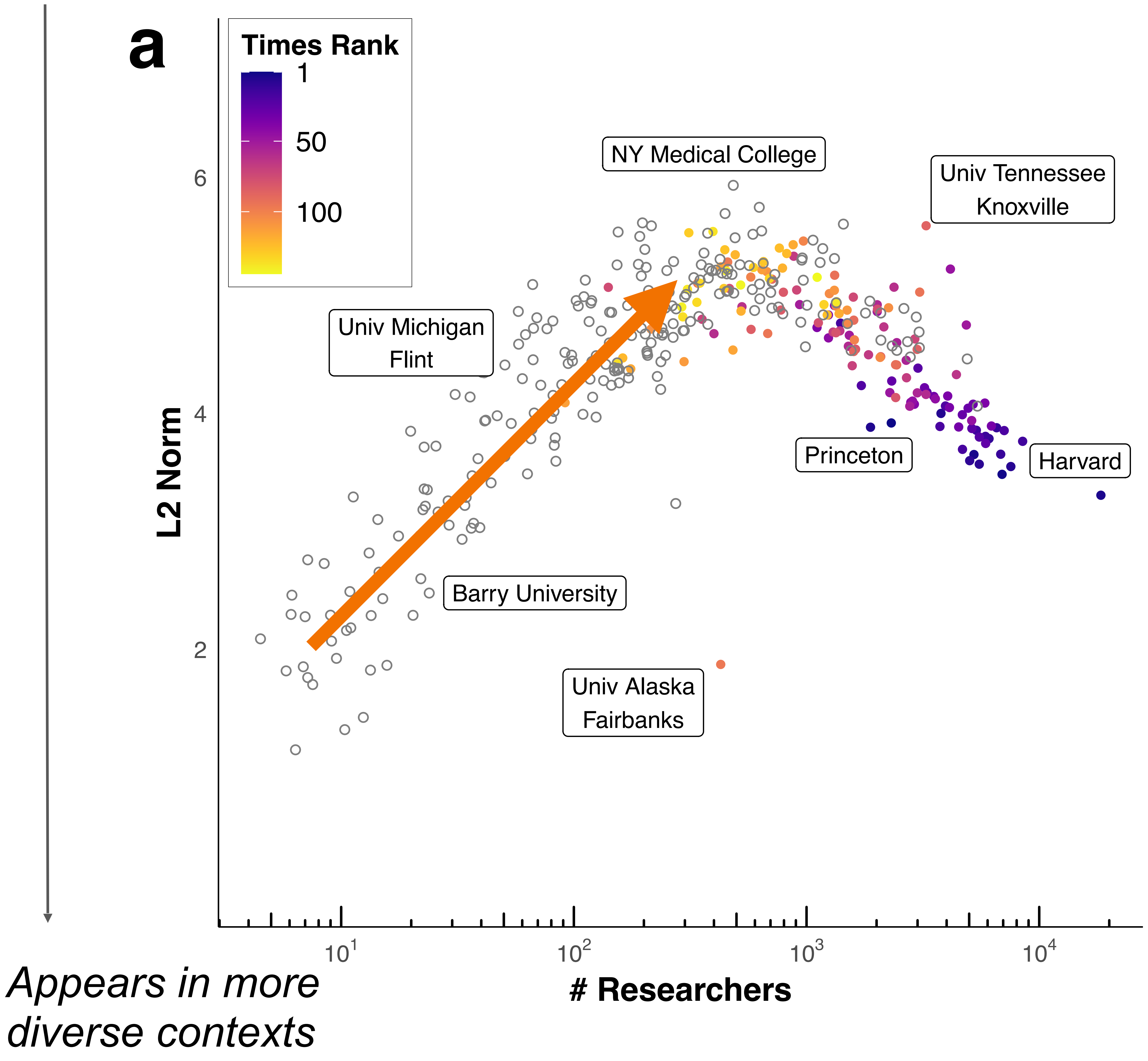
Prestigious U.S. universities appear in more diverse contexts



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)*.

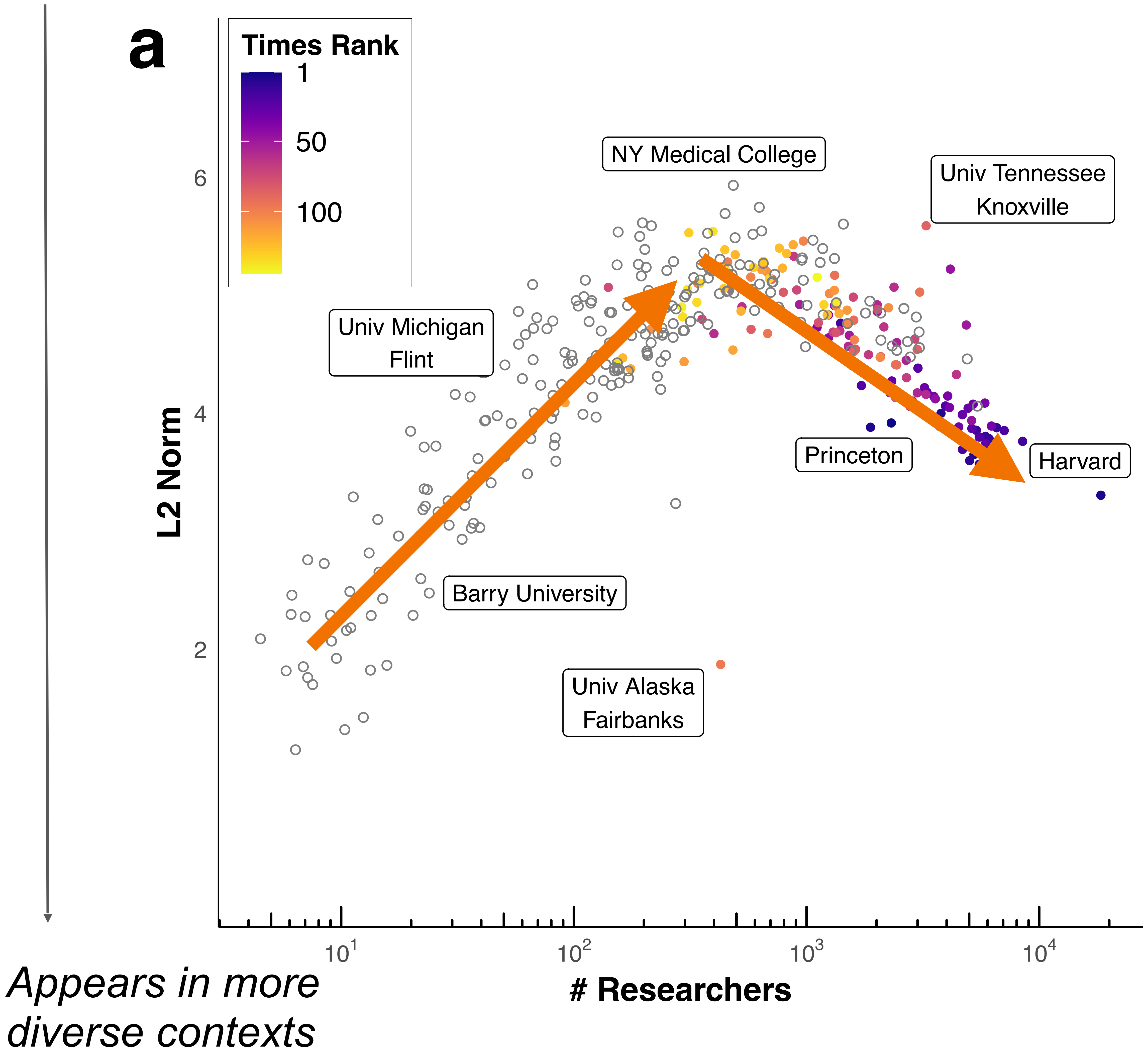
Prestigious U.S. universities appear in more diverse contexts



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



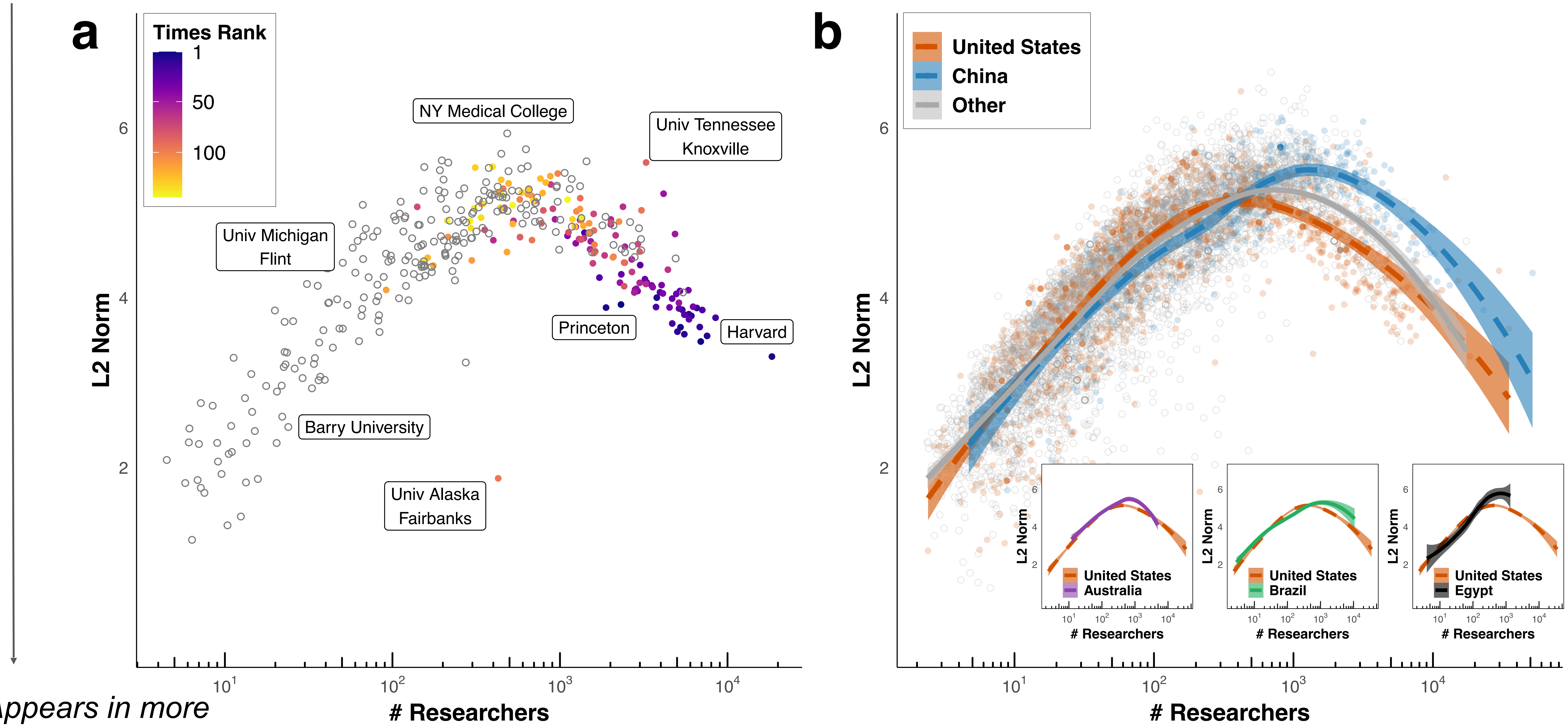
Bigger organizations are more isolated….unless they are prestigious

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

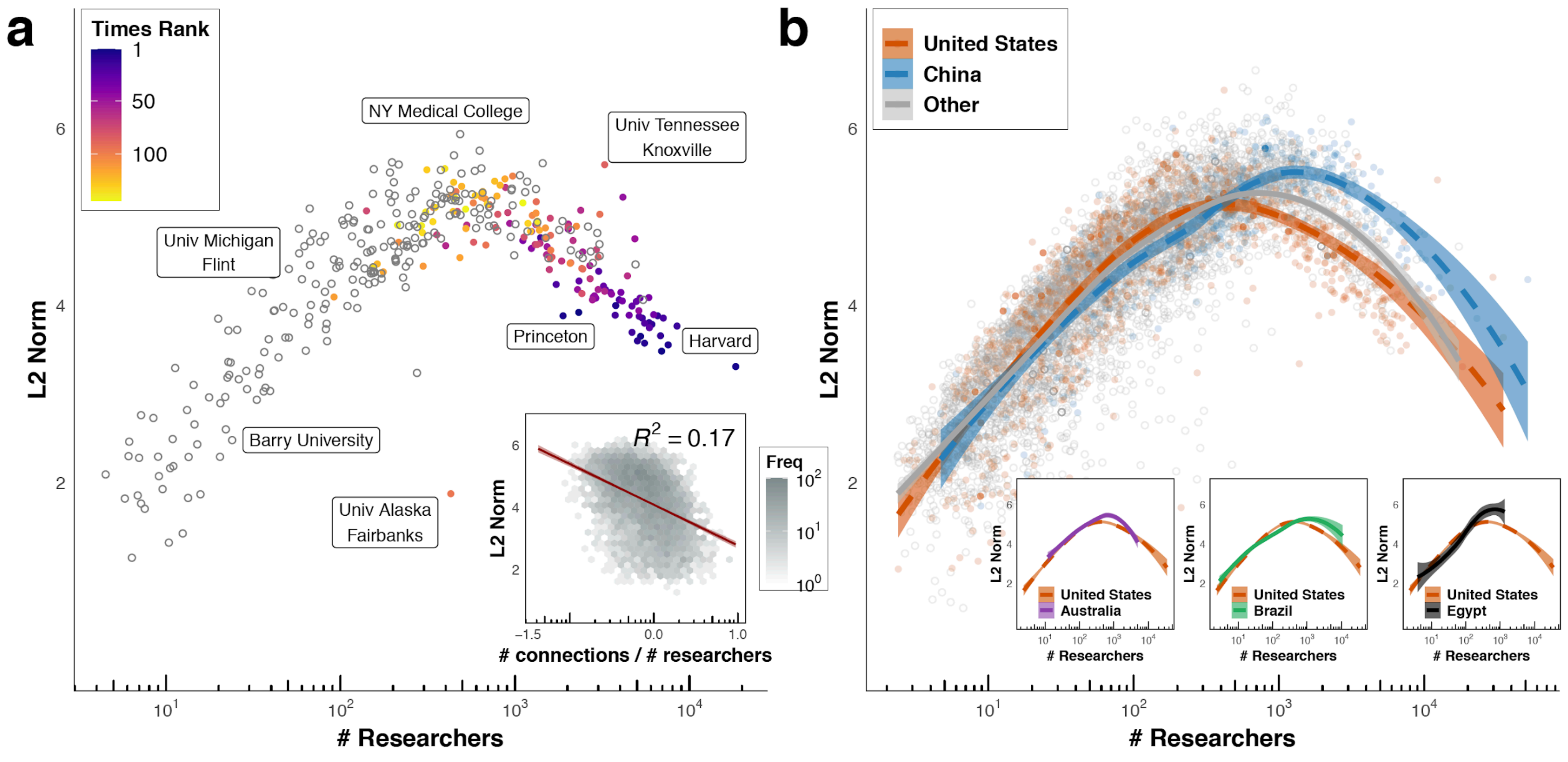
Repeats across many countries

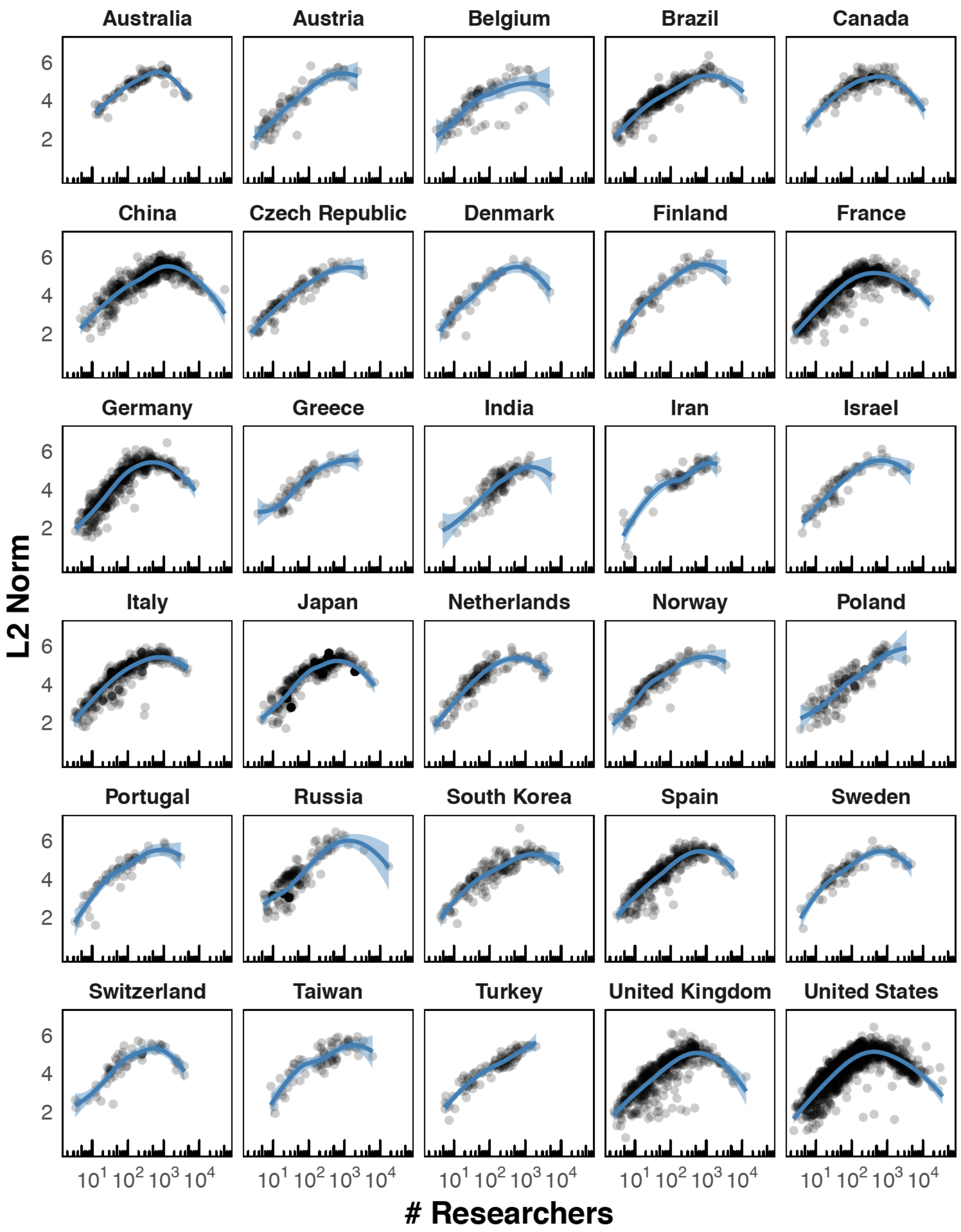


Appears in more diverse contexts

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

Conclusion

The first law of success:

“Performance drives success but when performance can’t be measured, networks drive success”

Success isn’t about you, its about us

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

“you” and “us” both exist in a context

Evaluation and success are deeply contextual

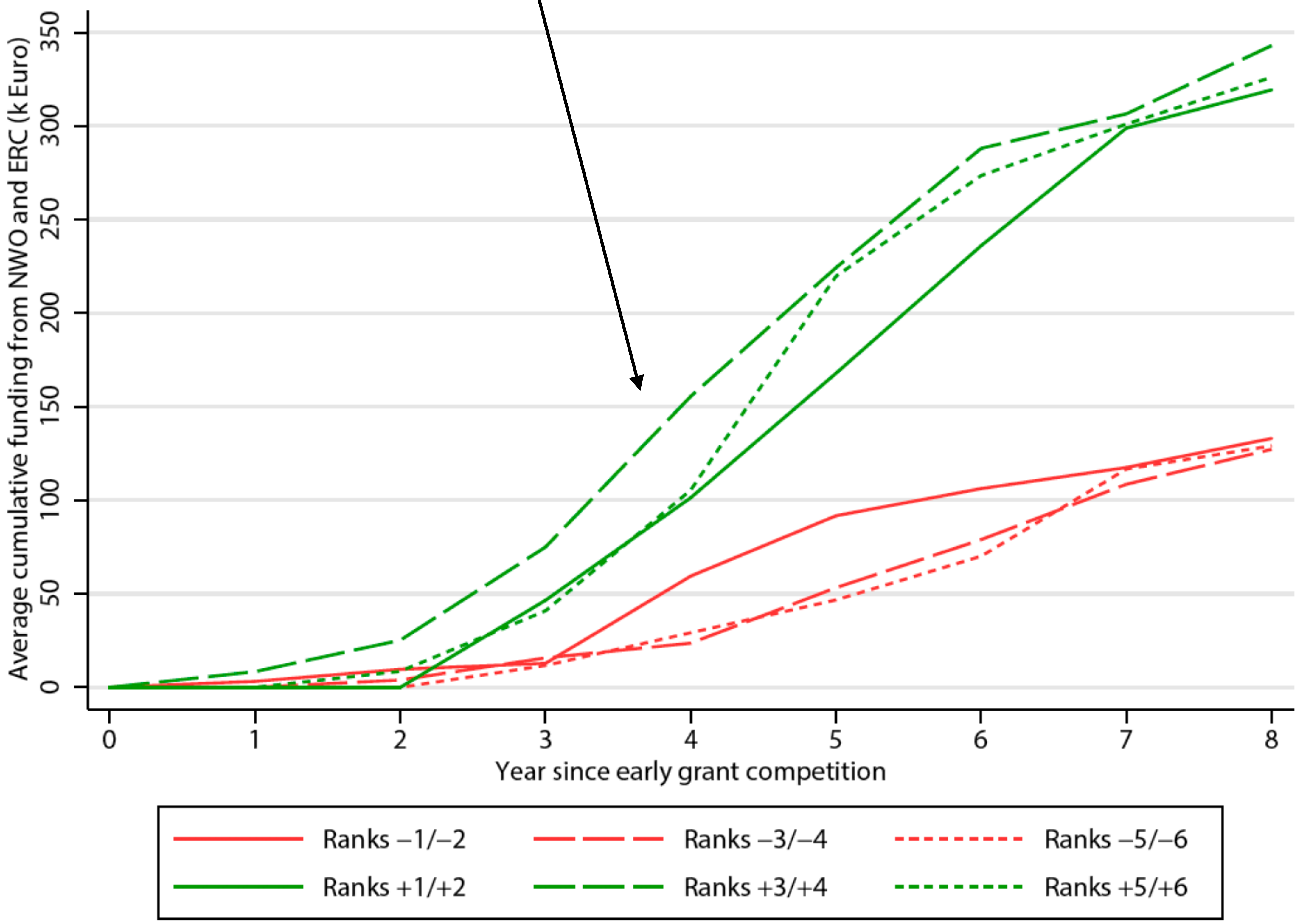
Context is often ignored in evaluation

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

Those who won early grants, won more over their lifetime



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

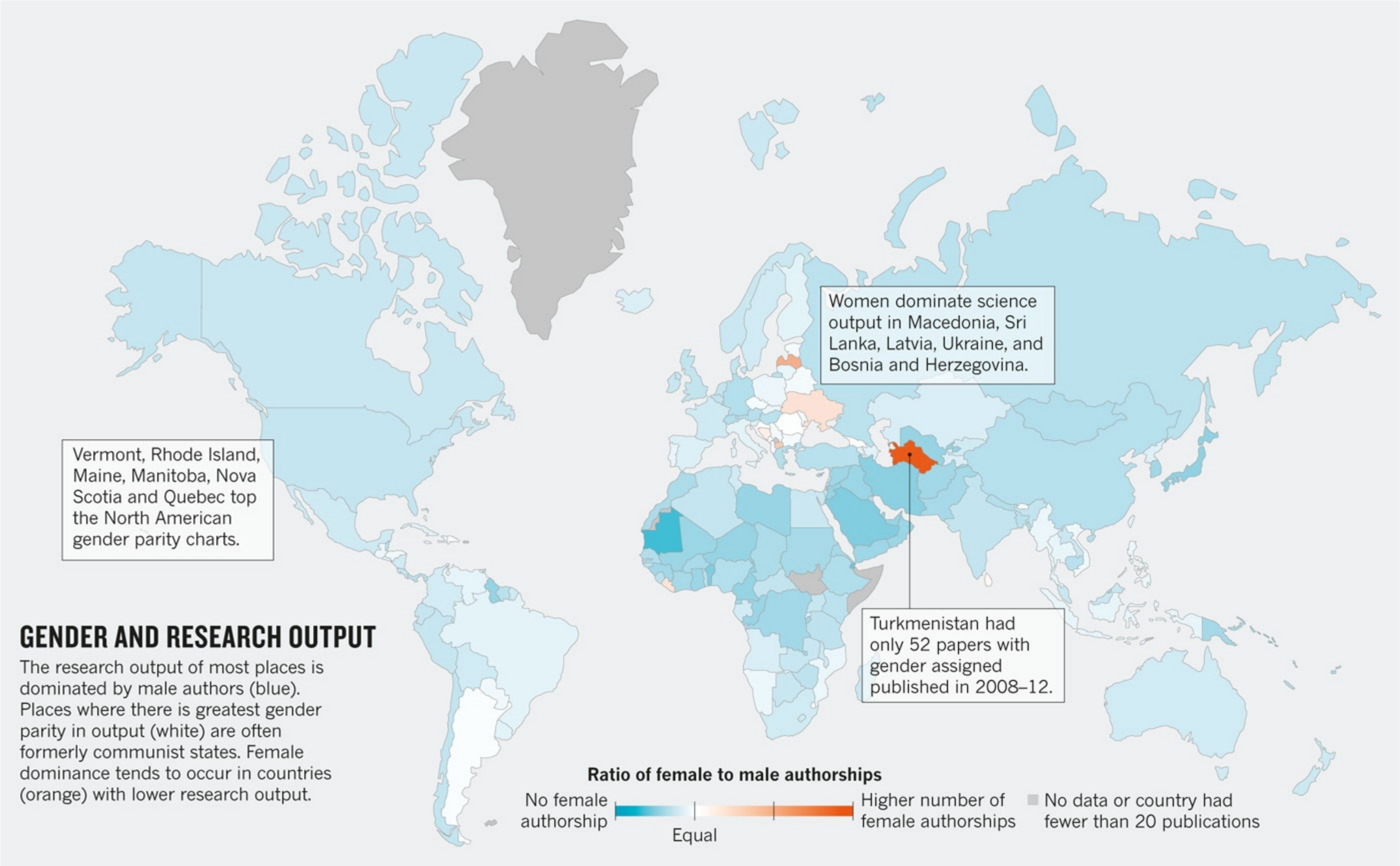
Merton, R. K. (1968). The Matthew Effect in Science. *Science*, *159*(3810), 56–63.

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.

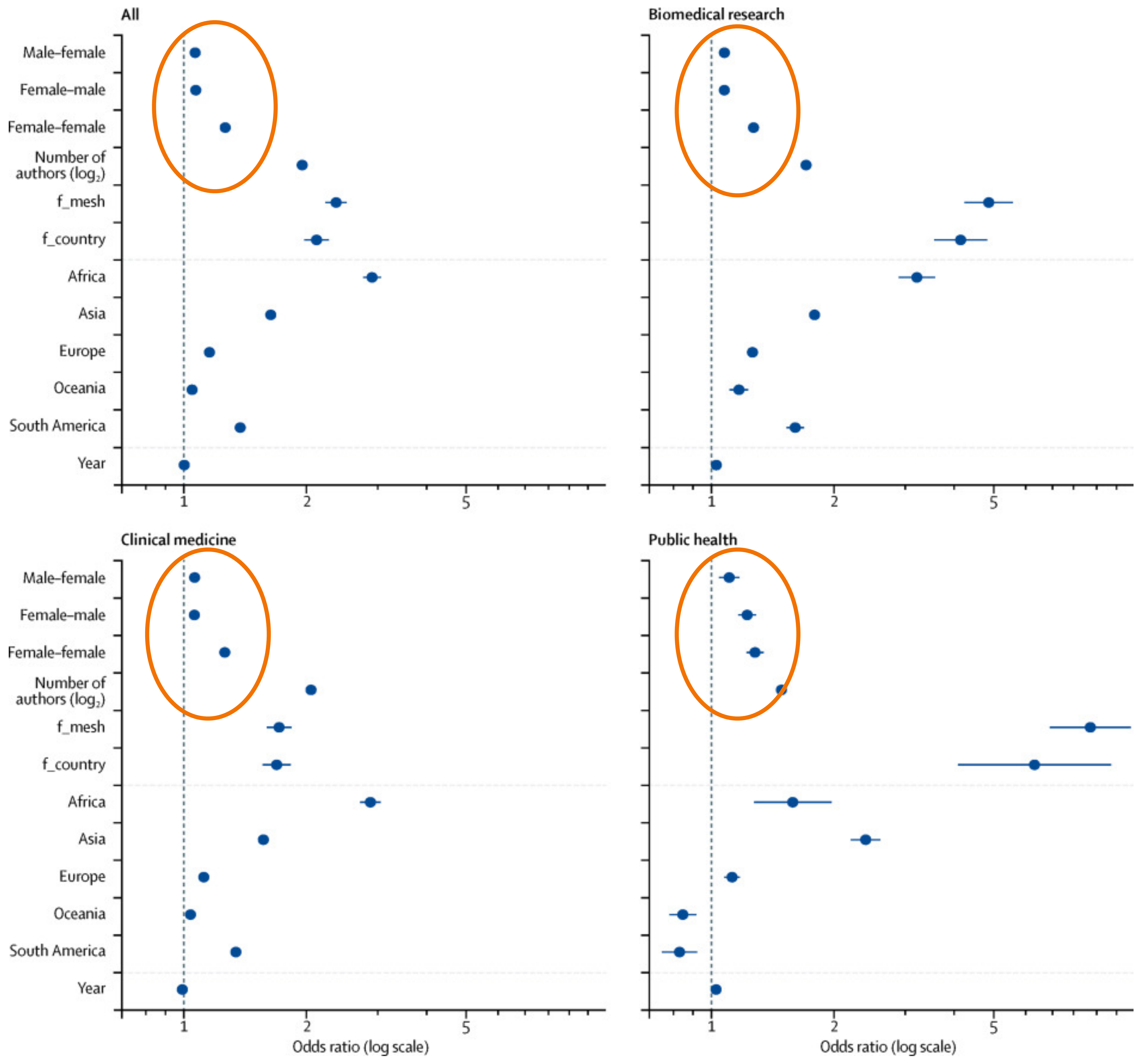
Women make up only around 30% of science worldwide



These disparities shape what we know about the world

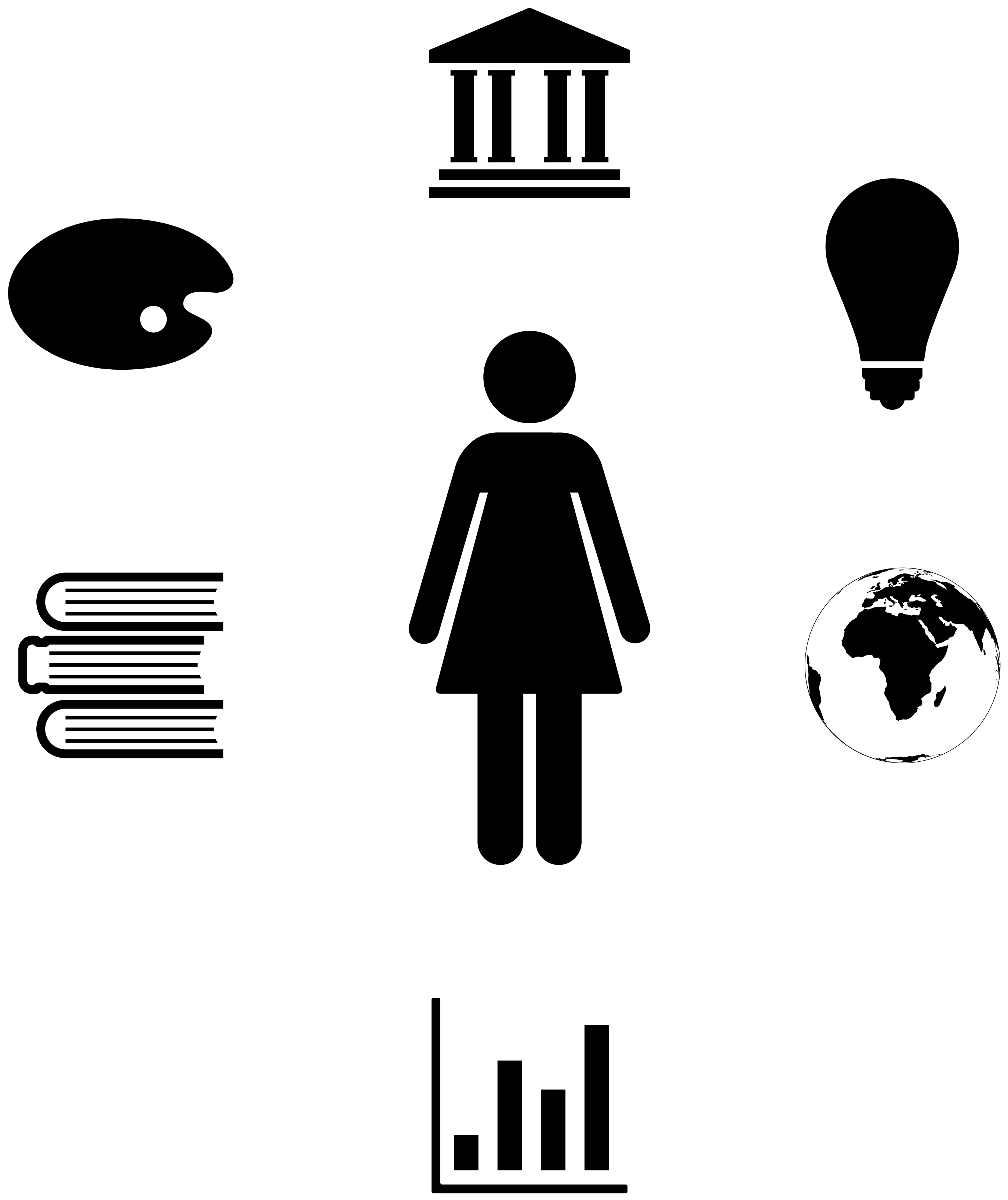
Sugimoto, C. R., Ahn, Y.-Y., Smith, E., Macaluso, B., & Larivière, V. (2019). Factors affecting sex-related reporting in medical research: A cross-disciplinary bibliometric analysis. *The Lancet*, *393*(10171), 550–559.

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

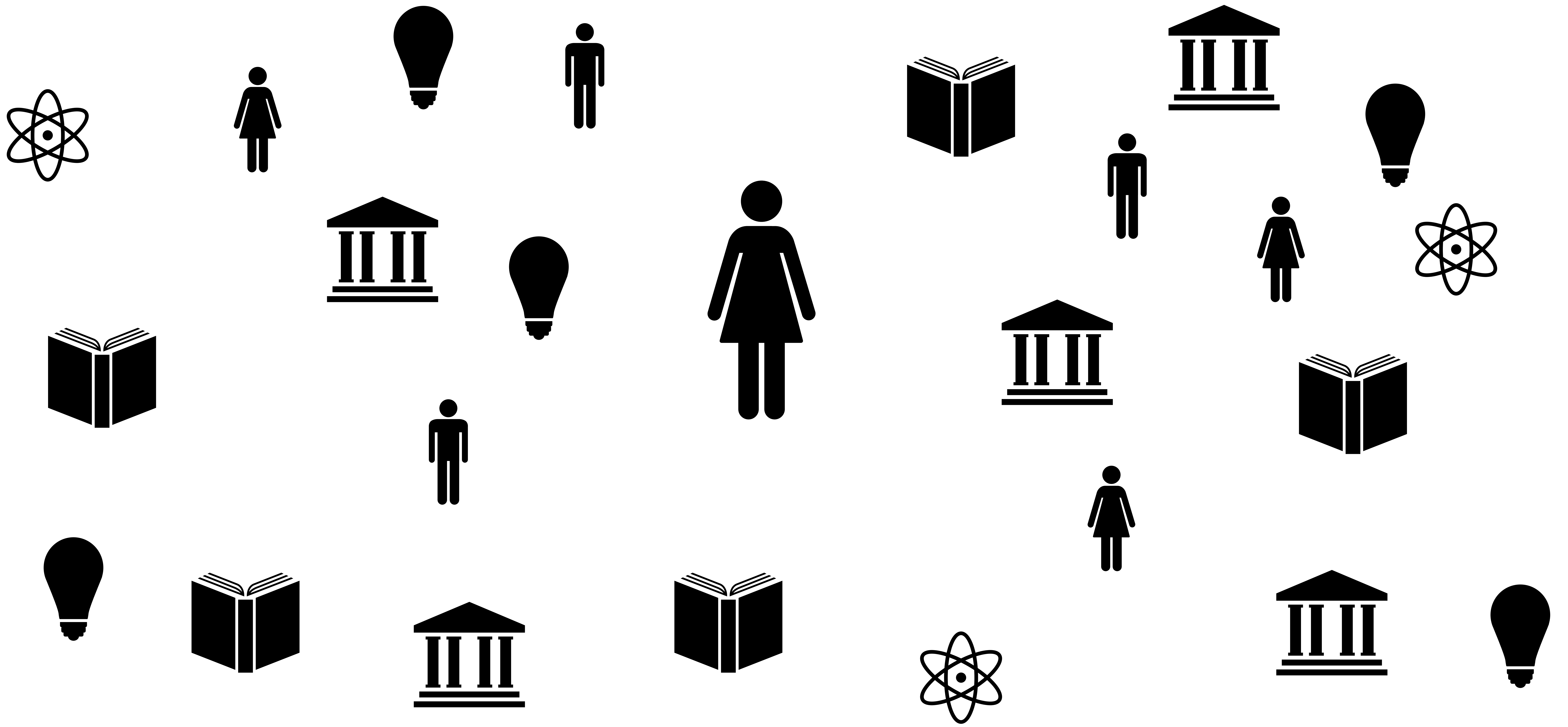


Researchers are not isolated

They have diverse characteristics that shape their careers



They are *embedded* in a wider ecosystem



Thank you!

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

Slides at: dakotamurray.me/2020-northeastern

Appendix

Appendix—eLife

eLife

- Life sciences journal
- Open access
- Consultative peer review

King, S. R. (2017). Peer Review: Consultative review is worth the wait. *ELife*, *6*, e32012..

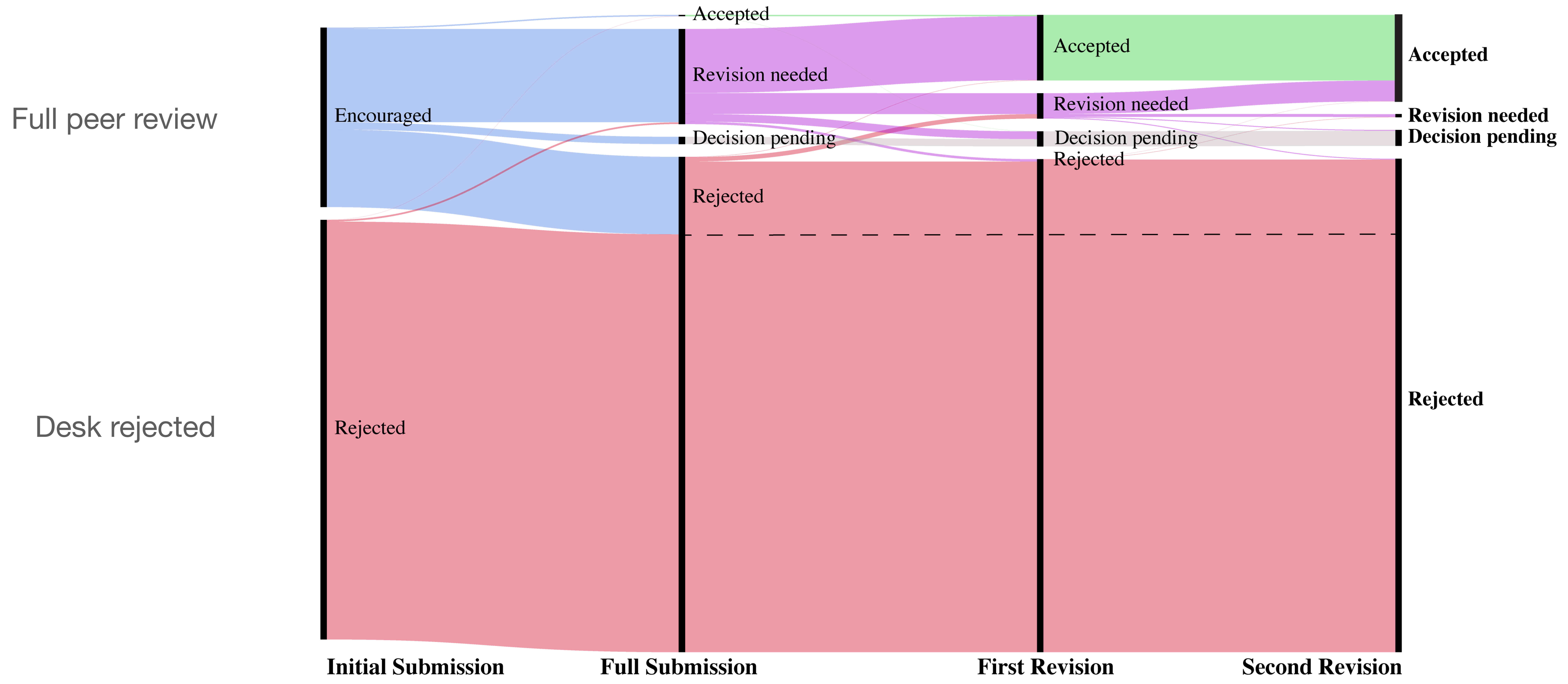
eLife

- Life sciences journal
- Open access
- Consultative peer review
- Does gender relate to outcomes?

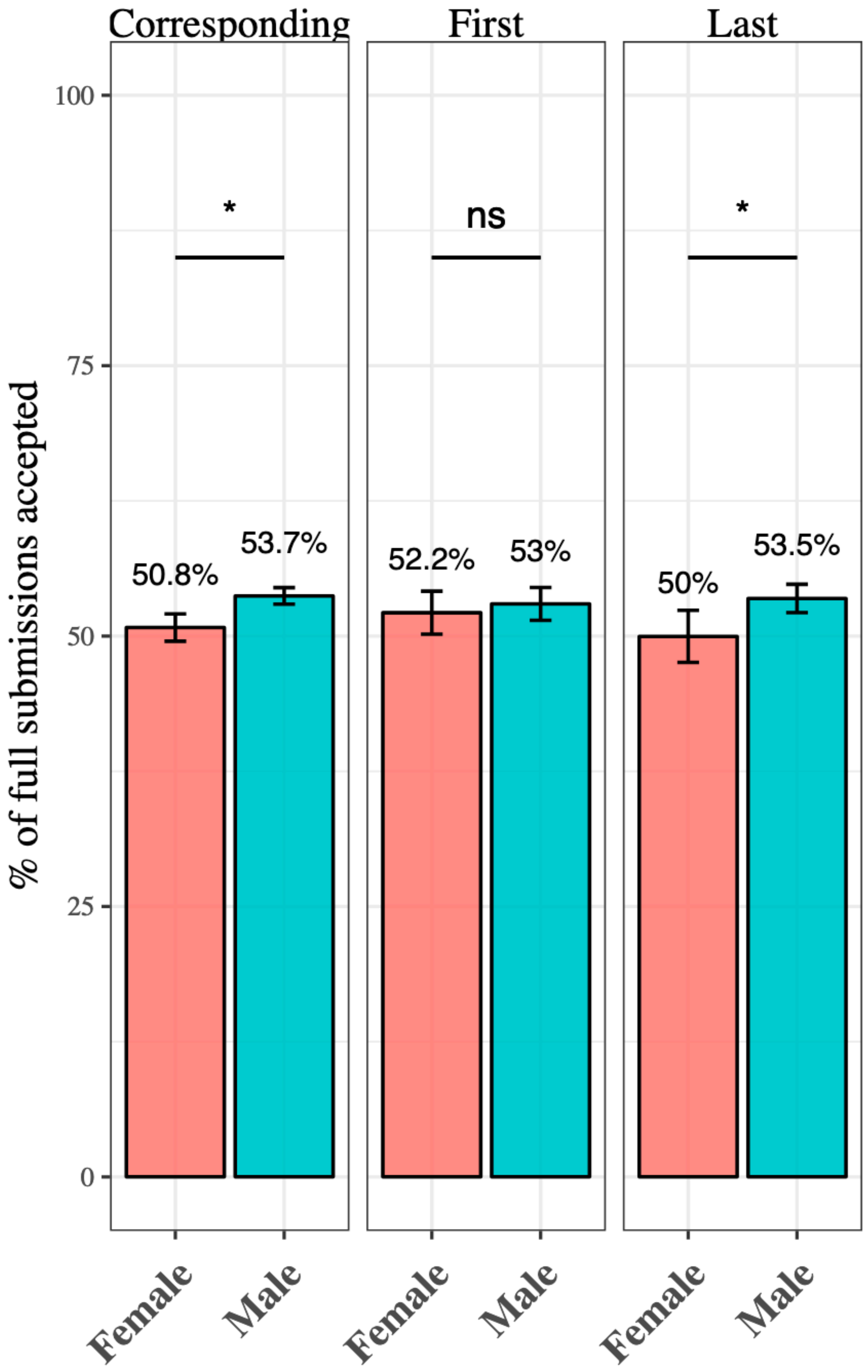
King, S. R. (2017). Peer Review: Consultative review is worth the wait. *ELife*, *6*, e32012..

Data

23,876 total submissions, of which 6,509 were reviewed and a decision made

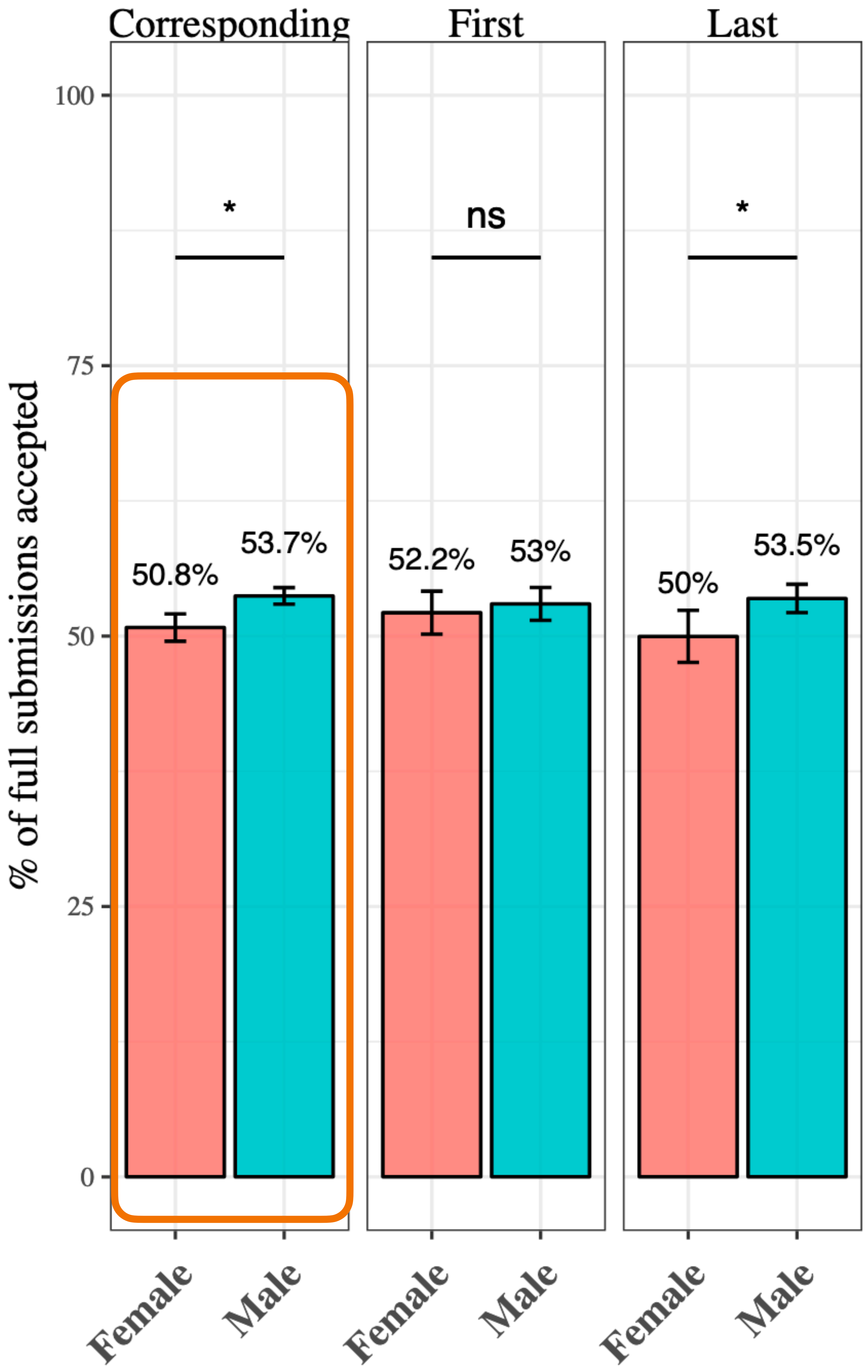


Women’s papers accepted less



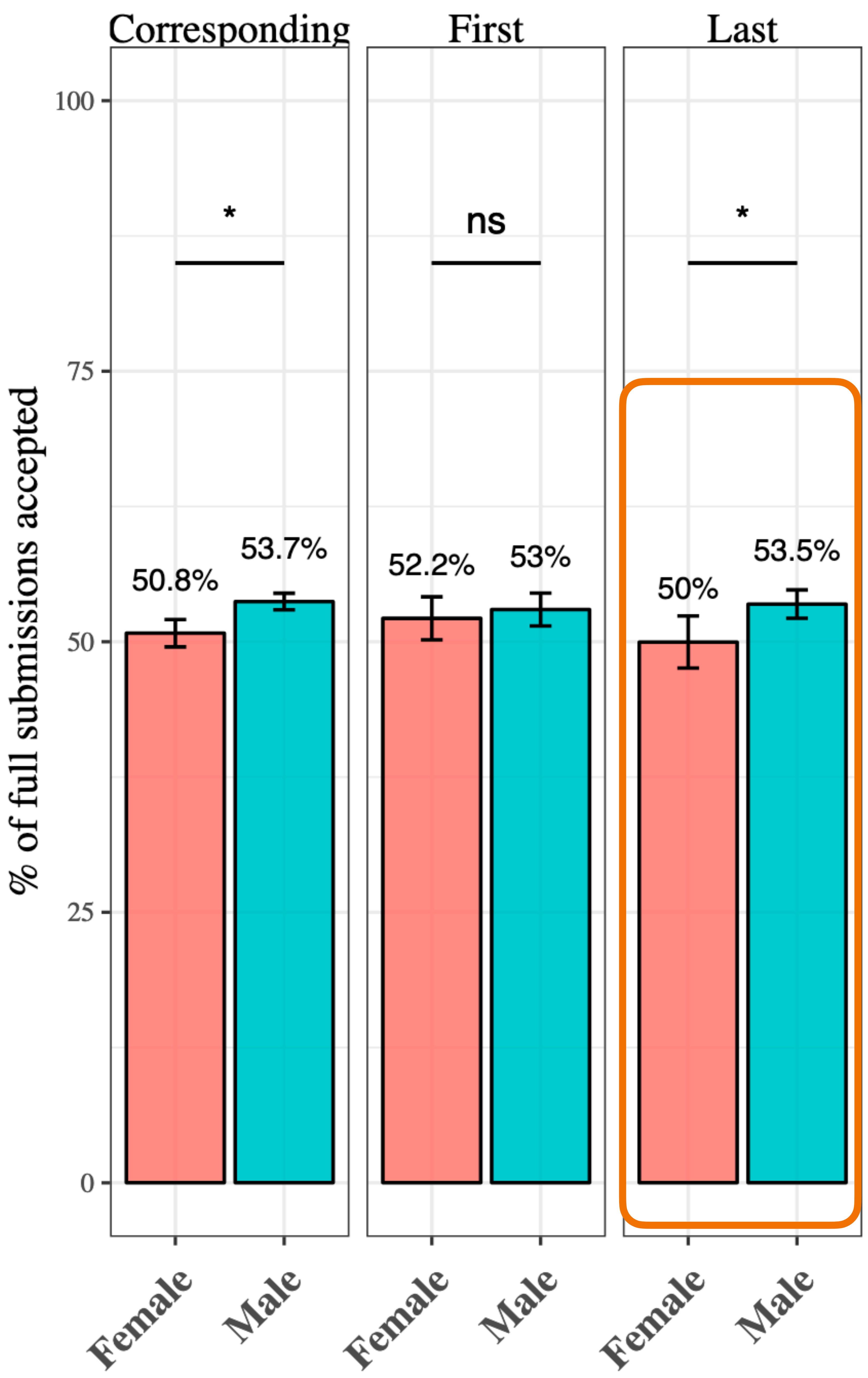
Women’s papers accepted less

- 2.9% gender disparity for corresponding authors



Women’s papers accepted less

- 2.9% gender disparity for corresponding authors
- 3.5% disparity for last authors



Lesson:

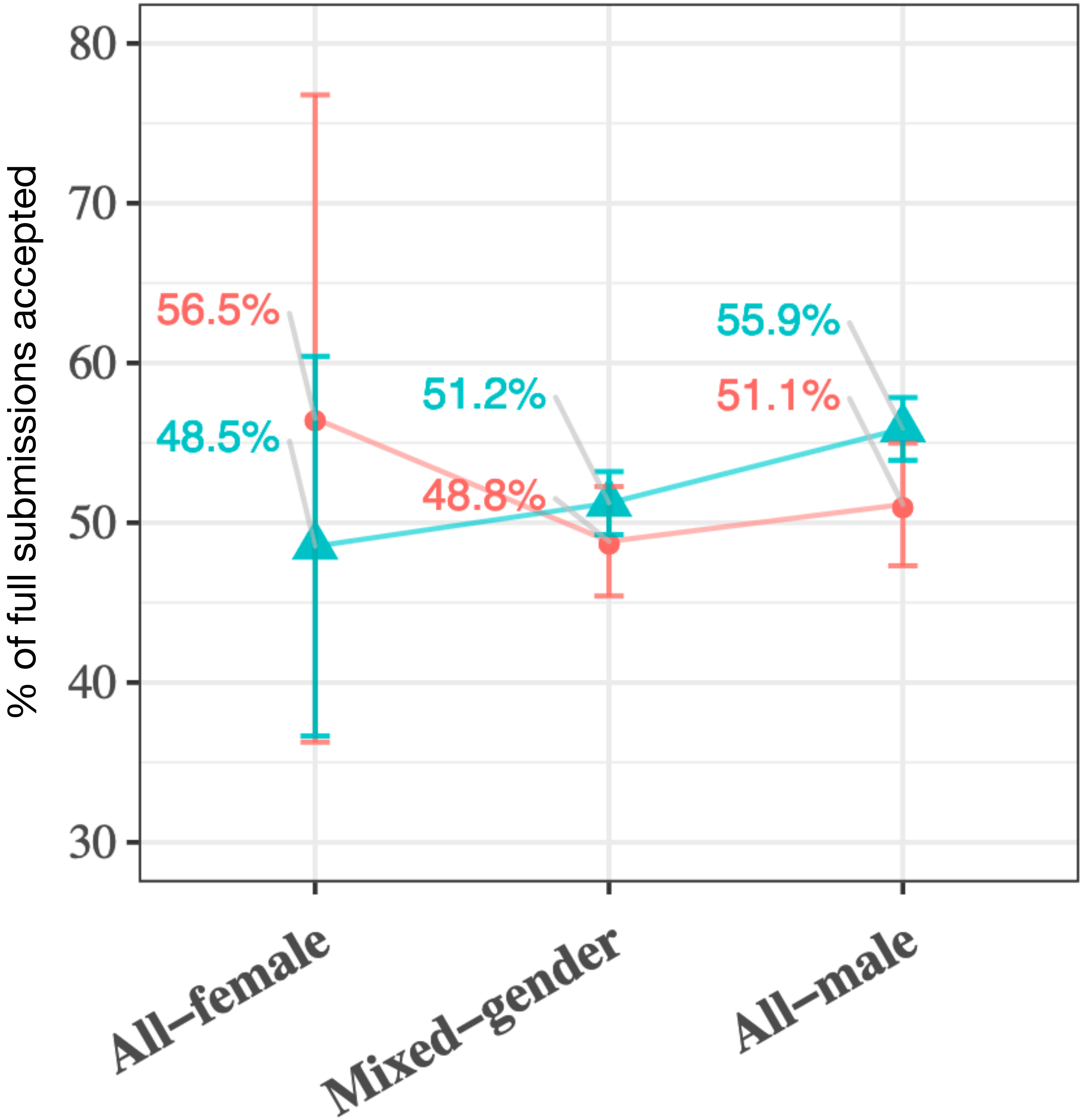
Peer review is a possible source of gender disparities

Due to Matthew Effects, even small biases can lead to big disparities

Day, T. E. (2015). The big consequences of small biases: A simulation of peer review. *Research Policy*, *44*(6), 1266–1270.

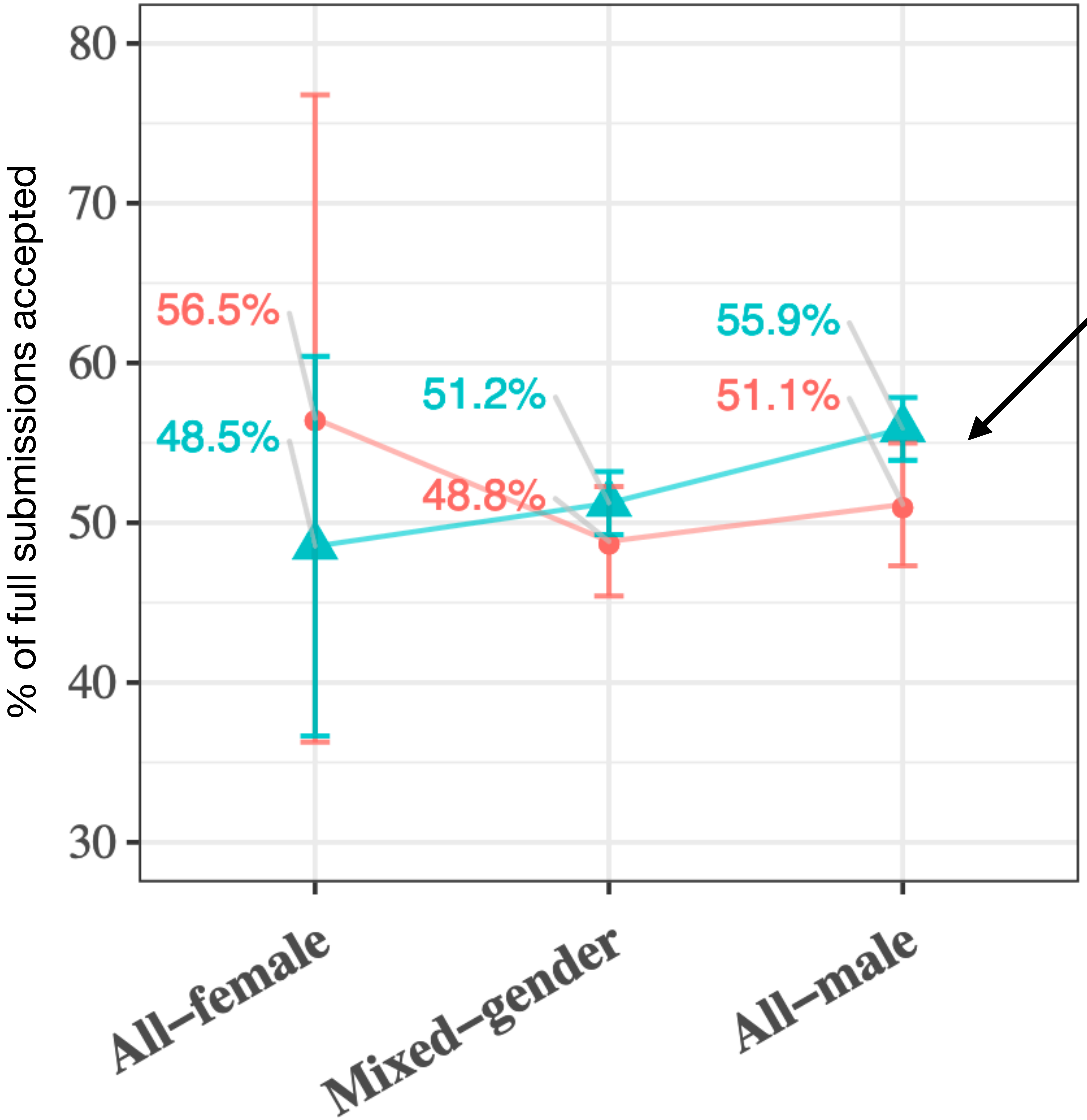
Reviewers matters

Disparities differ by the composition of the reviewer group



Reviewers matters

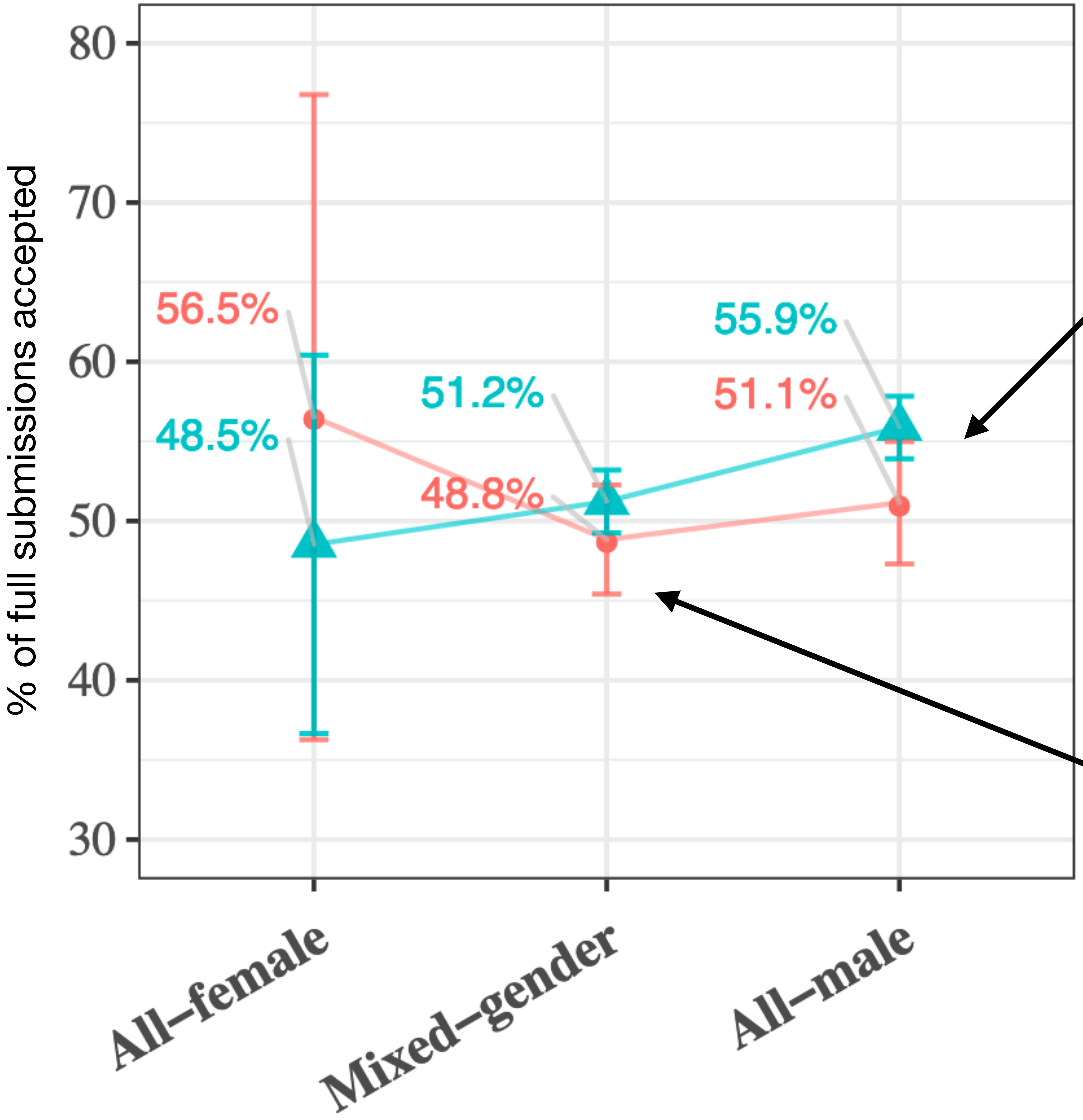
Disparities differ by the composition of the reviewer group



4.8% disparity for all-male teams

Reviewers matters

Disparities differ by the composition of the reviewer group



4.8% disparity for all-male teams

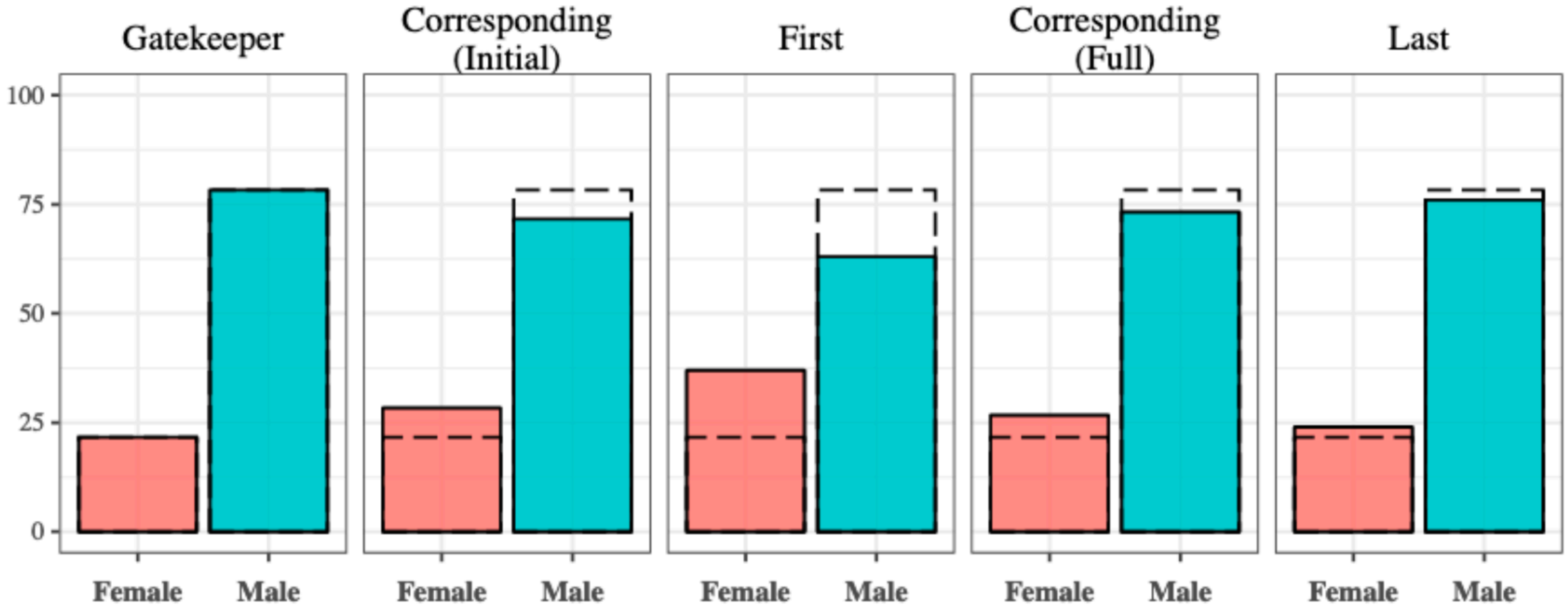
2.4% disparity for mixed-gender teams

Lesson:

Gatekeepers matter!

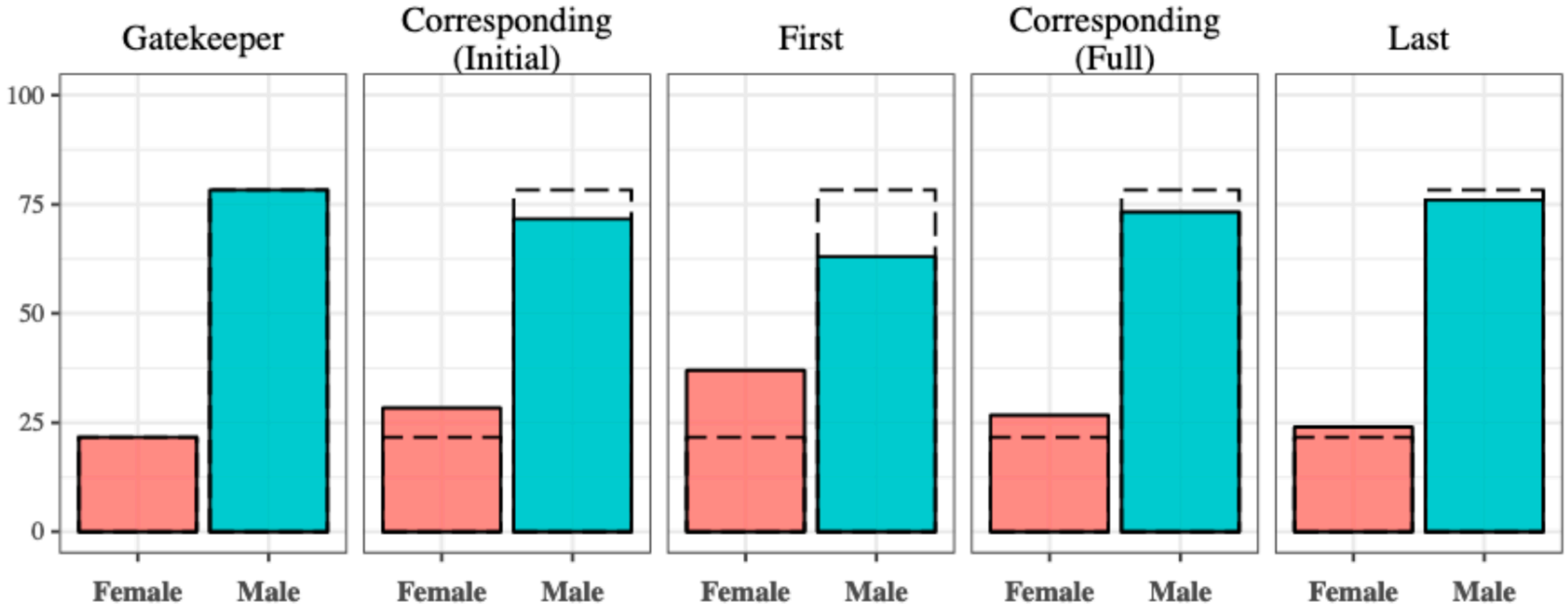
An easy solution

Make reviewers represent the authorship



An easy solution

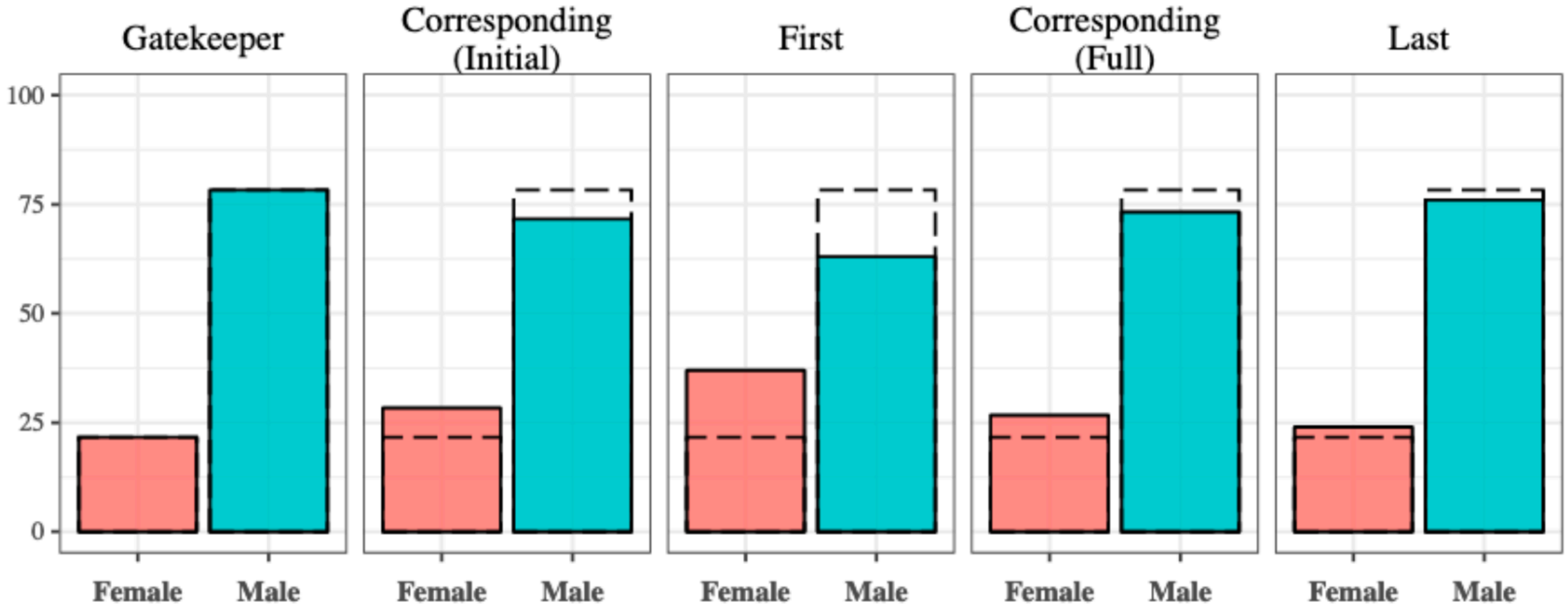
Make reviewers represent the authorship



This is something that eLife does well!

An easy solution

Make reviewers represent the authorship



This is something that eLife does well!

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

Lesson:

Open data on scholarly activity is essential to identify and correct issues

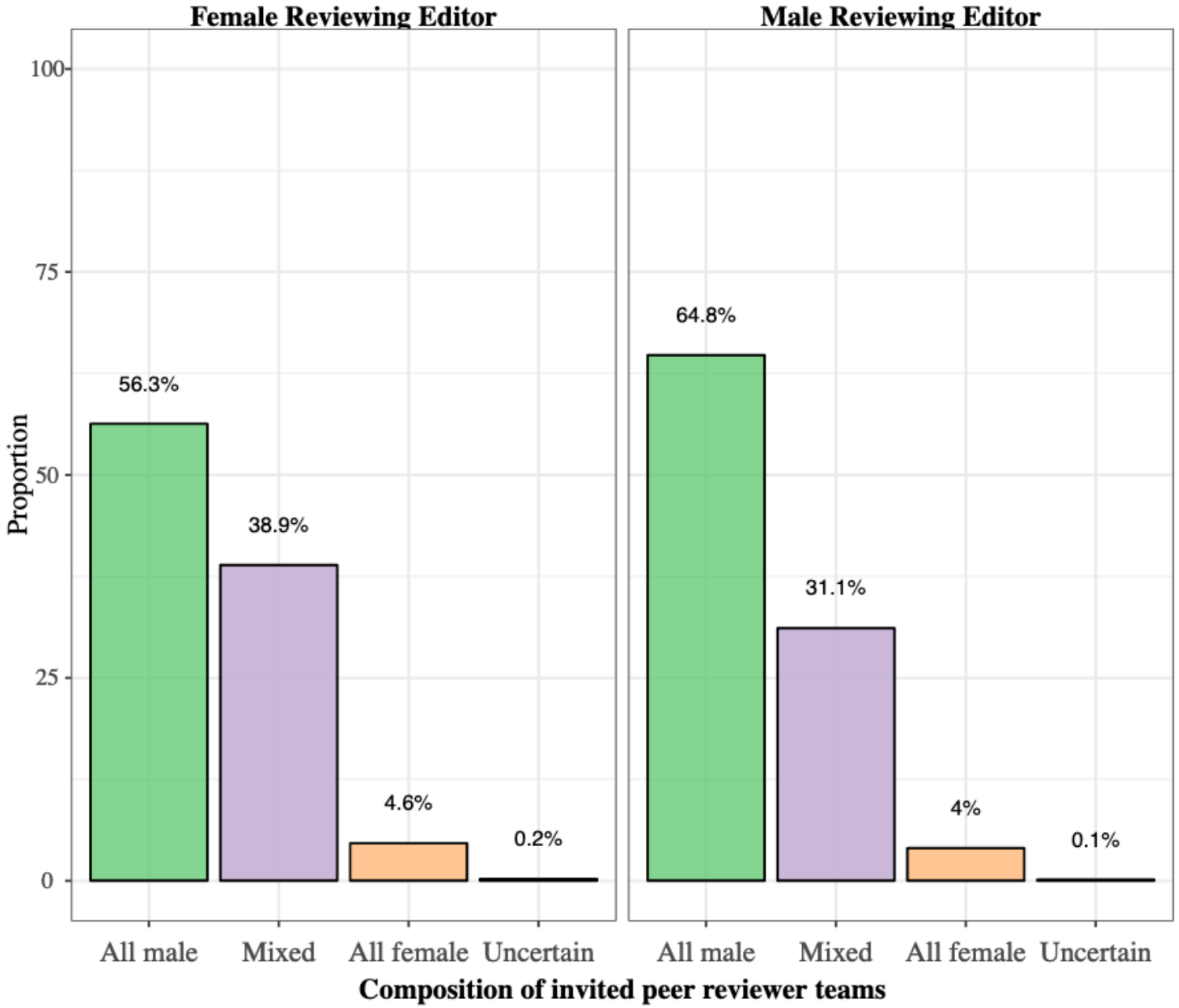
Read more, including results relating to nationality!

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.

https://www.biorxiv.org/content/10.1101/400515v3.full

Men and women editors make different teams

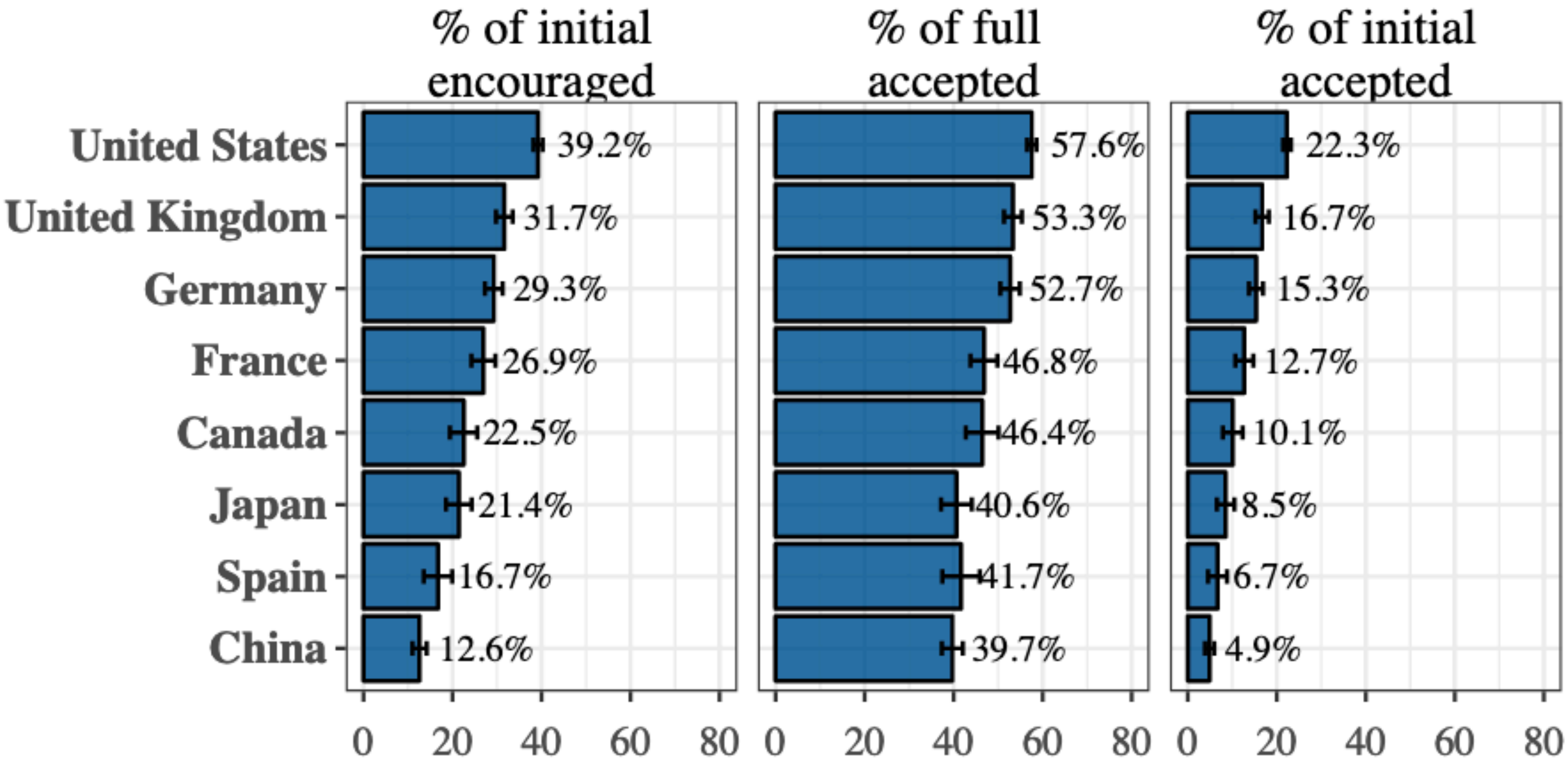
Women editors are more likely to recruit mixed-gender reviewer teams



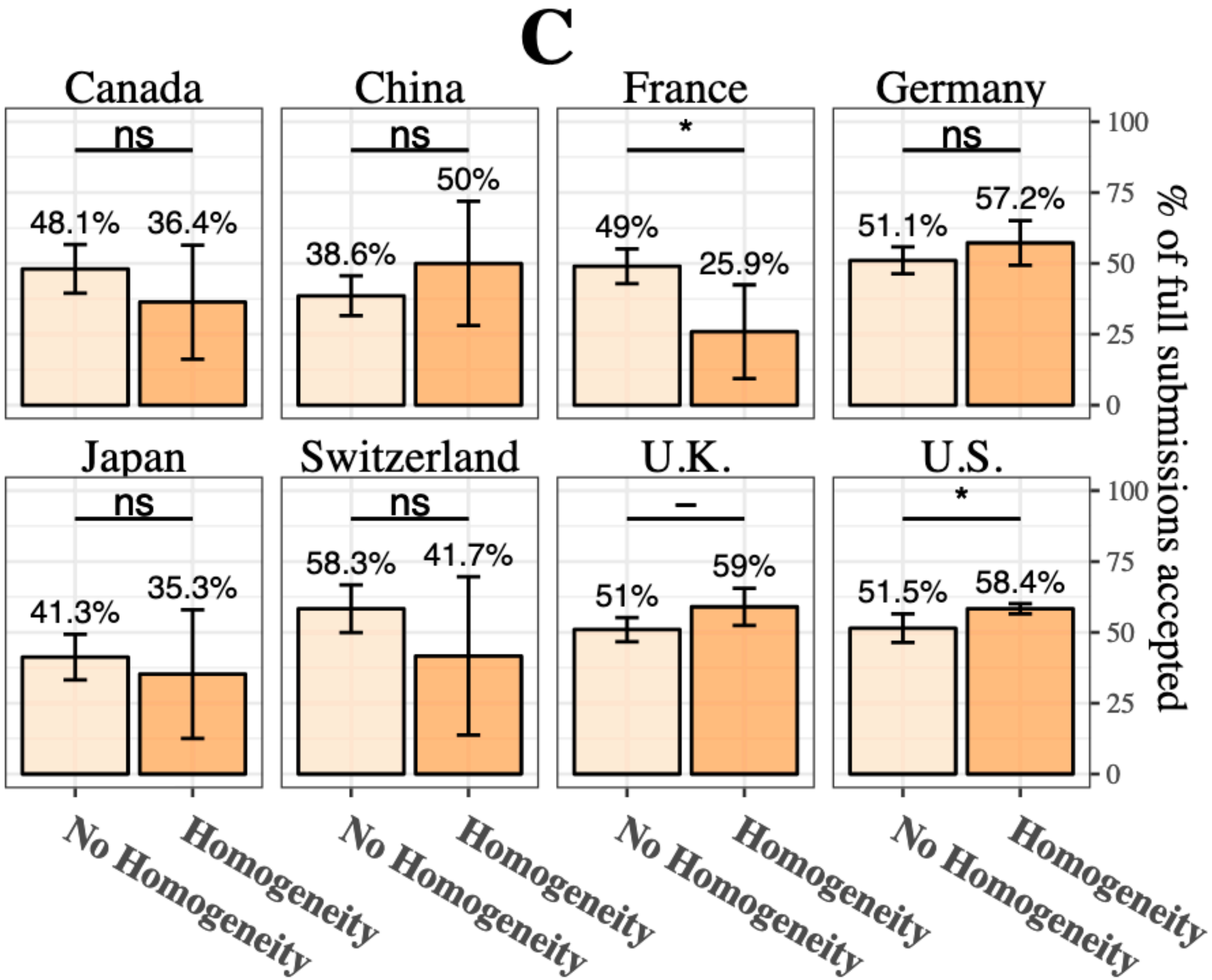
Also see:

Helmer, M., Schottdorf, M., Neef, A., & Battaglia, D. (2017). Research: Gender bias in scholarly peer review. *ELife*, *6*, e21718.

eLife—Nationality



eLife — Nationality

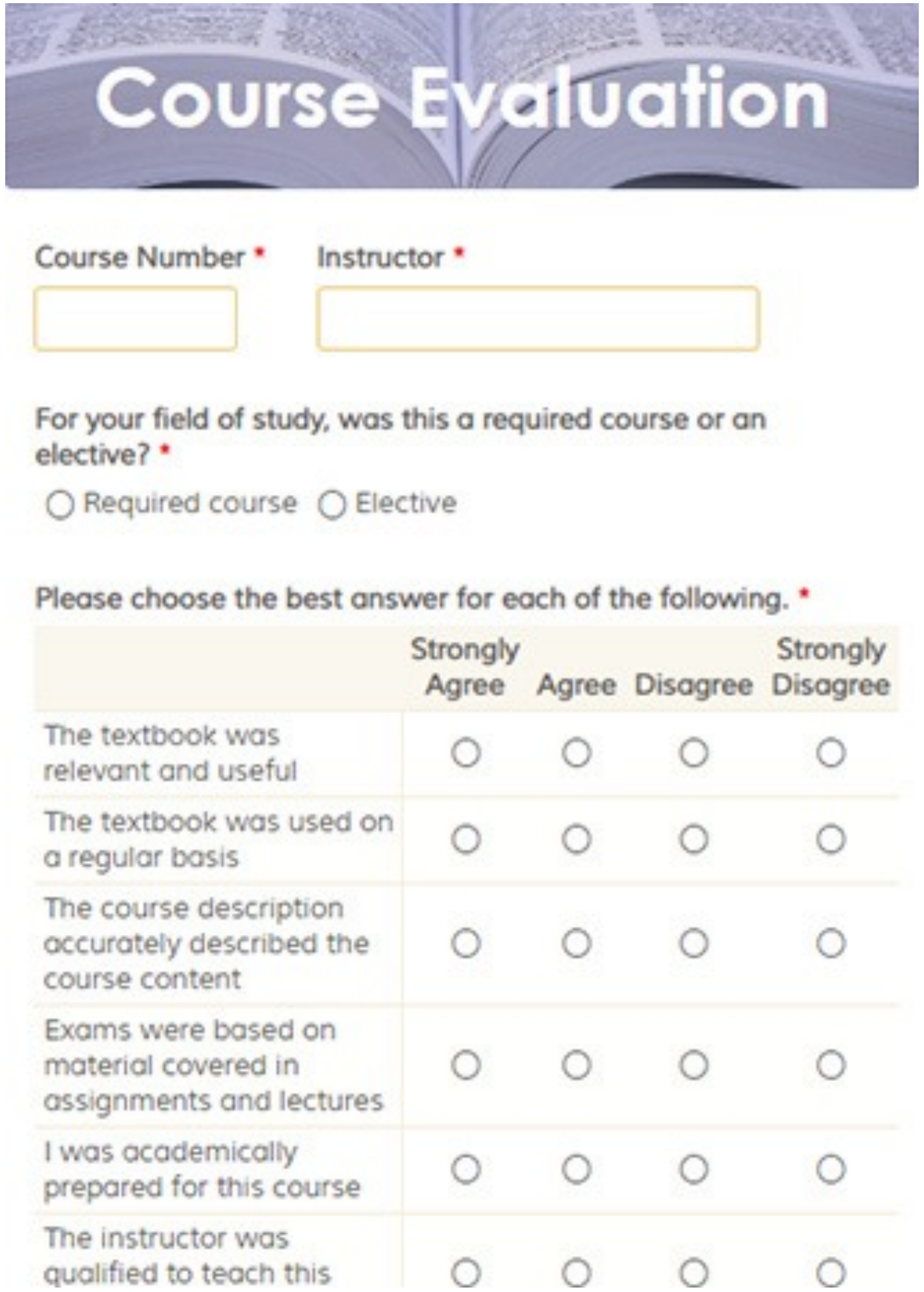


Student-Teacher evaluations

- Evaluate university faculty for hiring and promotion
- Gender, racial, and ethnic bias
- Invalid measures of teaching
- Data not open, little research at scale

Chávez, K., & Mitchell, K. M. W. (2019). Exploring Bias in Student Evaluations: Gender, Race, and Ethnicity. *PS: Political Science & Politics*, 1–5.

Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn?: A Meta-Analysis and Review of the Literature. *Journal of Marketing Education*, *31*(1), 16–30.



Course Evaluation

Course Number * Instructor *

For your field of study, was this a required course or an elective? *

○ Required course ○ Elective

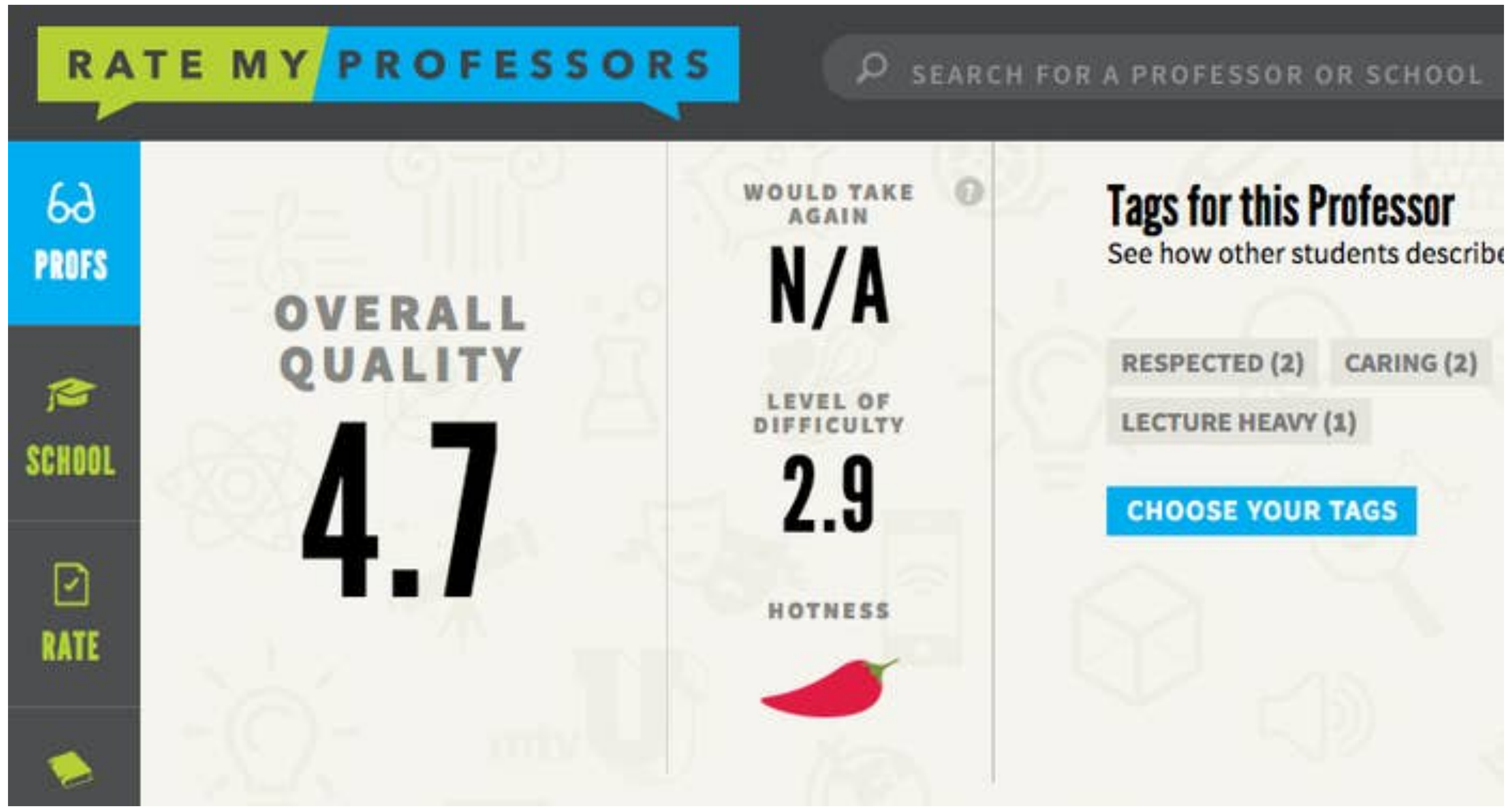
Please choose the best answer for each of the following. *

	Strongly Agree	Agree	Disagree	Strongly Disagree
The textbook was relevant and useful	○	○	○	○
The textbook was used on a regular basis	○	○	○	○
The course description accurately described the course content	○	○	○	○
Exams were based on material covered in assignments and lectures	○	○	○	○
I was academically prepared for this course	○	○	○	○
The instructor was qualified to teach this	○	○	○	○

Appendix—RateMyProfessors

RateMyProfessors.com

- Public ratings of instructors in the U.S., U.K., and Canada
- Shown to be an effective proxy for student-teacher evaluations



Clayson, D. E. (2014). What does ratemyprofessors.com actually rate? *Assessment & Evaluation in Higher Education*, *39*(6), 678–698.

Otto, J., Jr, D. A. S., & Ross, D. N. (2008). Does ratemyprofessor.com really rate my professor? *Assessment & Evaluation in Higher Education*, *33*(4), 355–368.

RateMyProfessors.com

How do we identify TT faculty?

Scraped over 1.7 million profiles and 19 million ratings for instructors at U.S. institutions of higher education

RateMyProfessors.com

How do we identify TT faculty?

RATE MY PROFESSORS

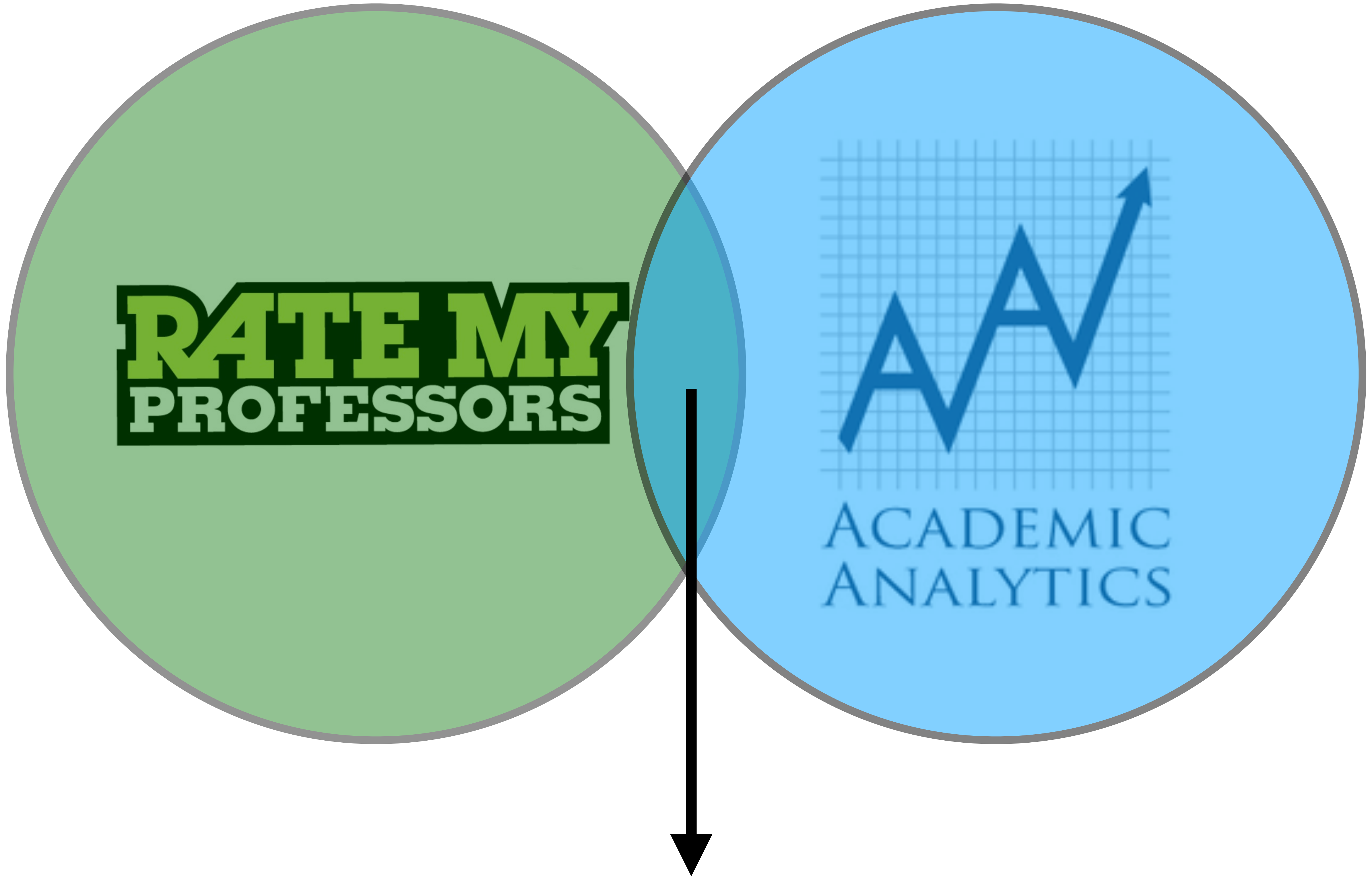
Scraped over 1.7 million profiles and 19 million ratings for instructors at U.S. institutions of higher education

ACADEMIC ANALYTICS

Metadata on 165,000 Tenure and Tenure Track faculty U.S. listed in *Academic Analytics*

Match datasets

- Based on names, department,
- 18,946 faculty
- Include relevant metadata from both datasets
- Gender, race inferred



ID	Rating	Gender	Race	...
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…

Read more,

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.

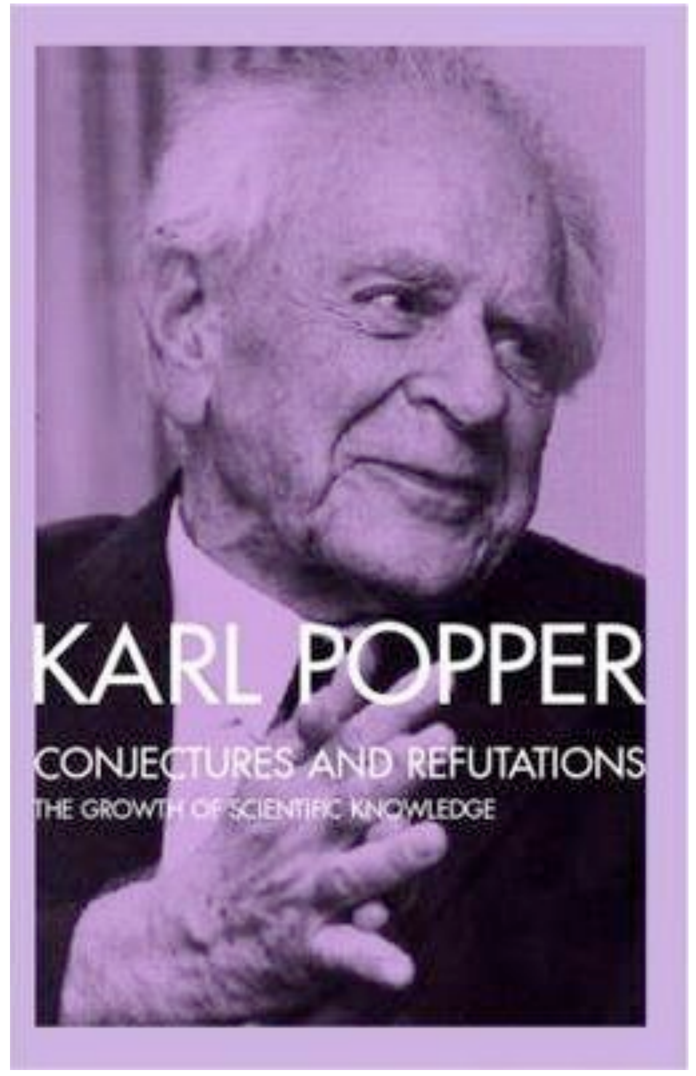
10.1371/journal.pone.0233515

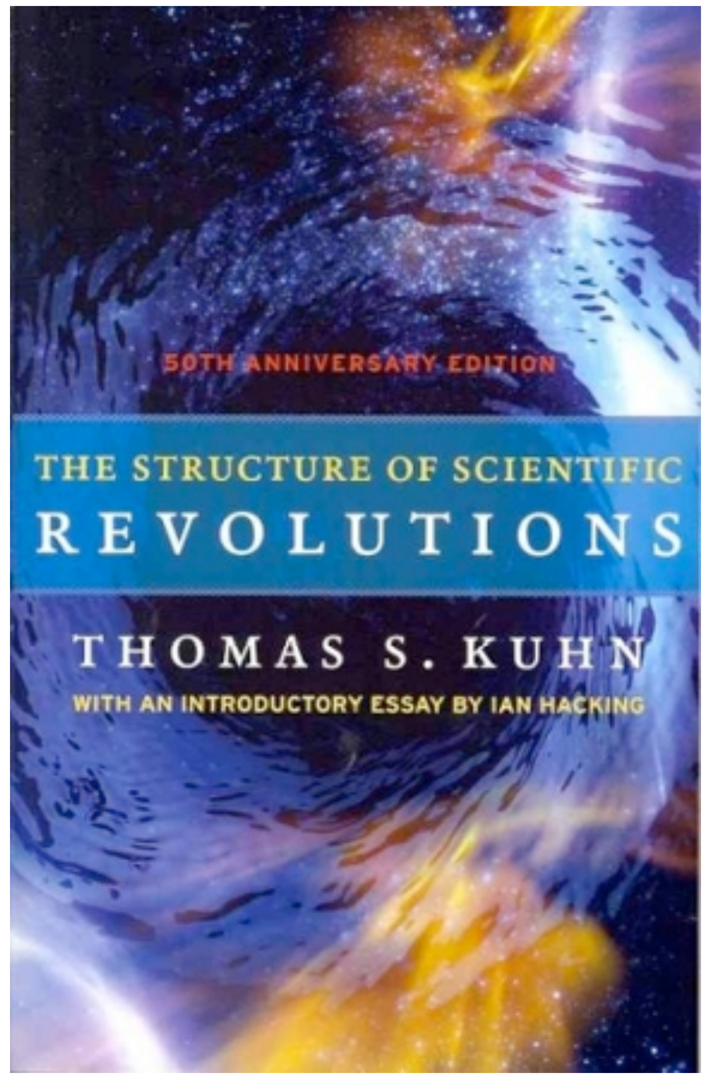
Appendix—disagreement

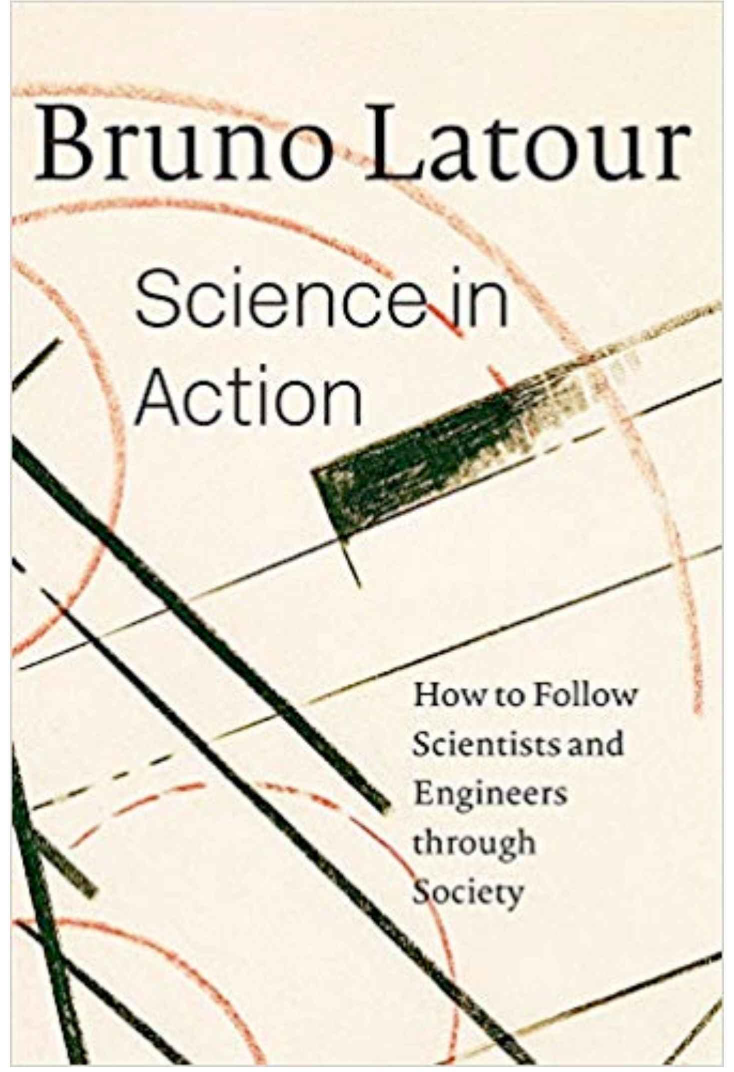
Disagreement in science

One of the ways we evaluate each other

- Dialectic debates
- Popper's falsification
- Kuhn's anomalies and revolutions
- Latour's controversies







Past work?

“Negative citations”

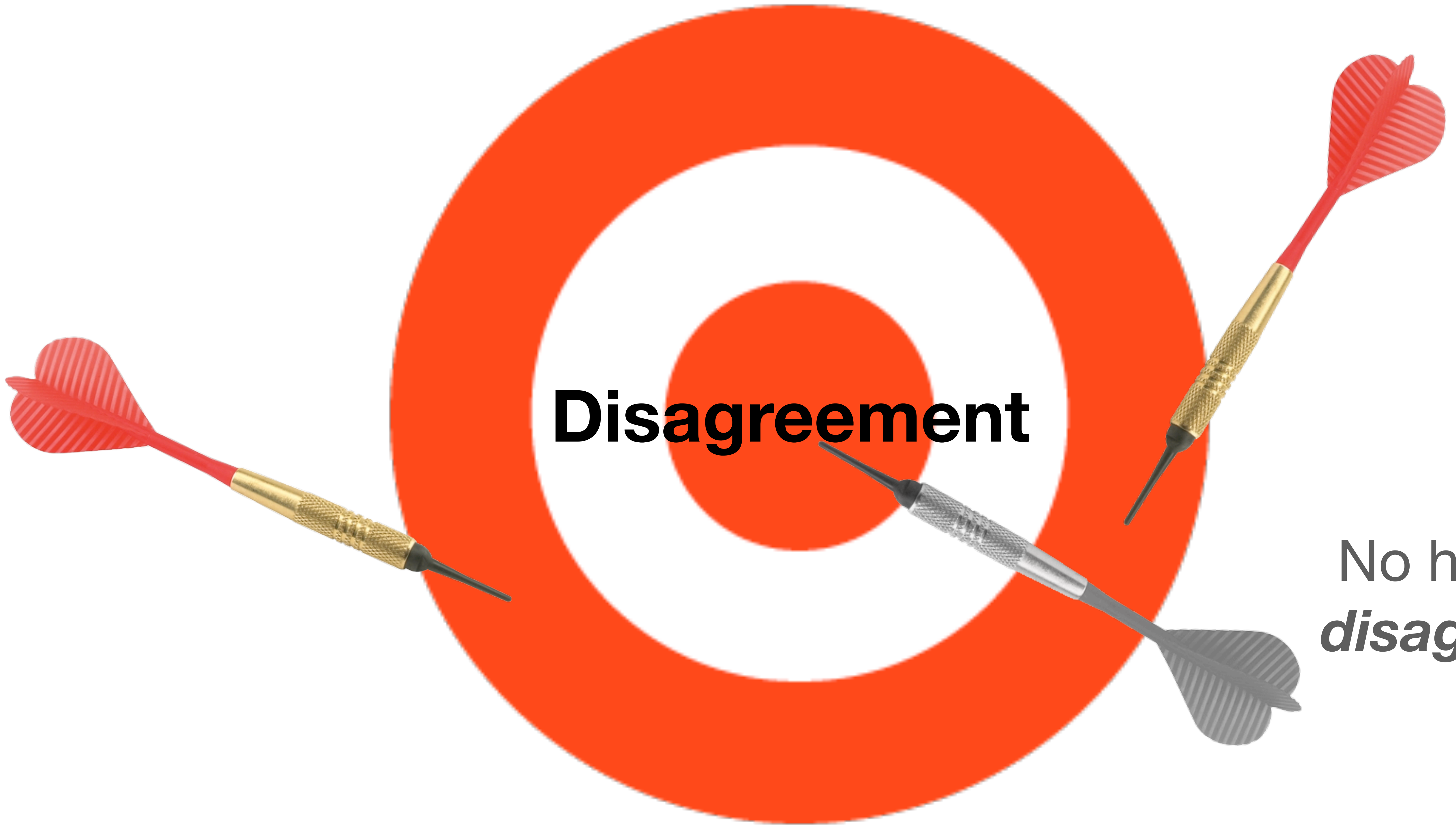
“Uncertainty”

Chen, C., Song, M., & Heo, G. E. (2018). A scalable and adaptive method for finding semantically equivalent cue words of uncertainty. JOI

Catalini, C., Lacetera, N., & Oettl, A. (2015). The incidence and role of negative citations in science. PNAS

Past work?

“Uncertainty”



“Negative citations”

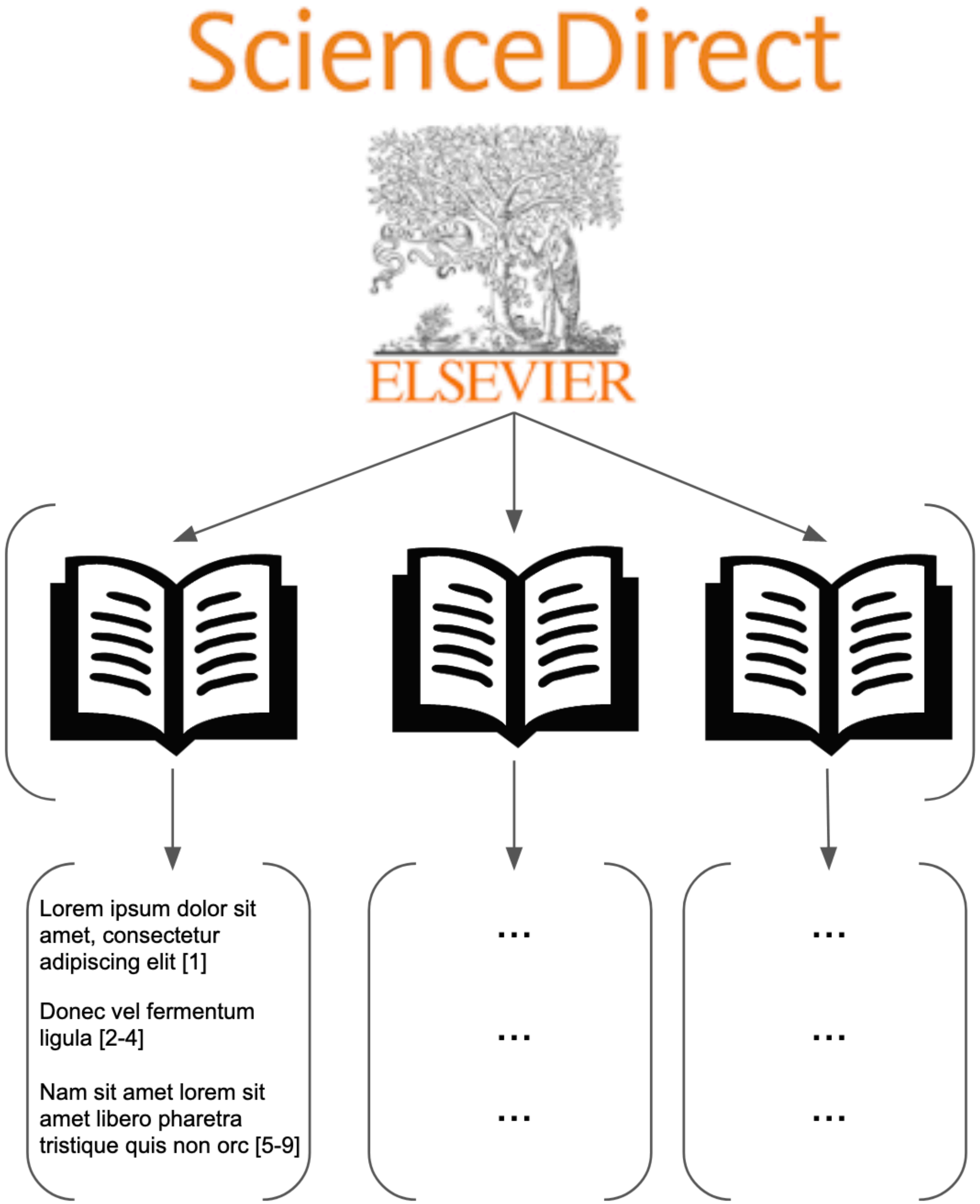
No holistic analysis of ***disagreement…***

Chen, C., Song, M., & Heo, G. E. (2018). A scalable and adaptive method for finding semantically equivalent cue words of uncertainty. JOI

Catalini, C., Lacetera, N., & Oettl, A. (2015). The incidence and role of negative citations in science. PNAS

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*					

"...recruiting participants was **challenging**..."

"However, recent studies have **disagreed** with this approach"

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**

a **Percent Agreement**

	standalone	+studies	+ideas	+methods	+results	AVERAGE
challenge*	92.0	86.0	86.0	98.0	92.0	**90.8**
conflict*	96.0	94.0	90.0	98.0	86.0	**92.8**
contradict*	88.0	88.0	82.0	94.0	86.0	**87.6**
contrary	70.0	82.0	66.0	78.0	90.0	**77.2**
contrast*	86.0	68.0	92.0	80.0	82.0	**81.6**
controvers*	90.0	92.0	84.0	96.0	94.0	**91.2**
debat*	90.0	88.0	86.0	90.0	88.0	**88.4**
differ*	100.0	82.0	90.0	74.0	90.0	**87.2**
disagree*	78.0	92.0	82.0	78.0	92.0	**84.4**
disprov*	70.0	92.0	90.0	86.0	86.0	**84.8**
no consensus	96.0	98.0	95.0	98.0	92.0	**95.8**
questionable	64.0	48.0	82.0	82.0	48.0	**64.8**
refut*	92.0	82.0	84.0	78.0	84.0	**84.0**
AVERAGE	**85.5**	**84.0**	**85.3**	**86.9**	**85.4**	

% Agree
100
50
0

Validity

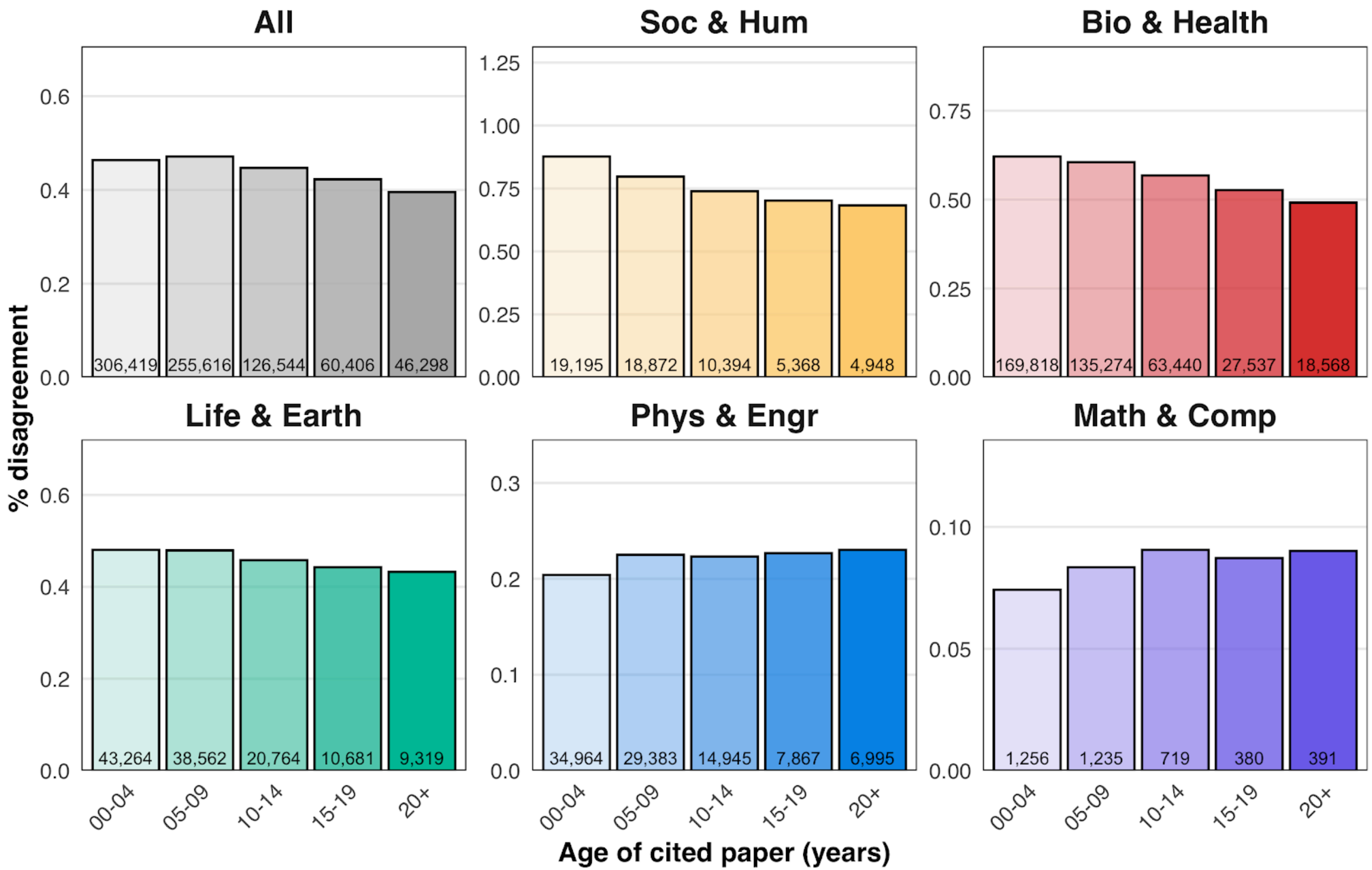
Coder 1: ✓ Valid **Coder 2: ✓ Valid**	Coder 1: ✓ Valid Coder 2: X Invalid
Coder 1: X Invalid Coder 2: ✓ Valid	Coder 1: X Invalid Coder 2: X Invalid

b **Percent Validity**

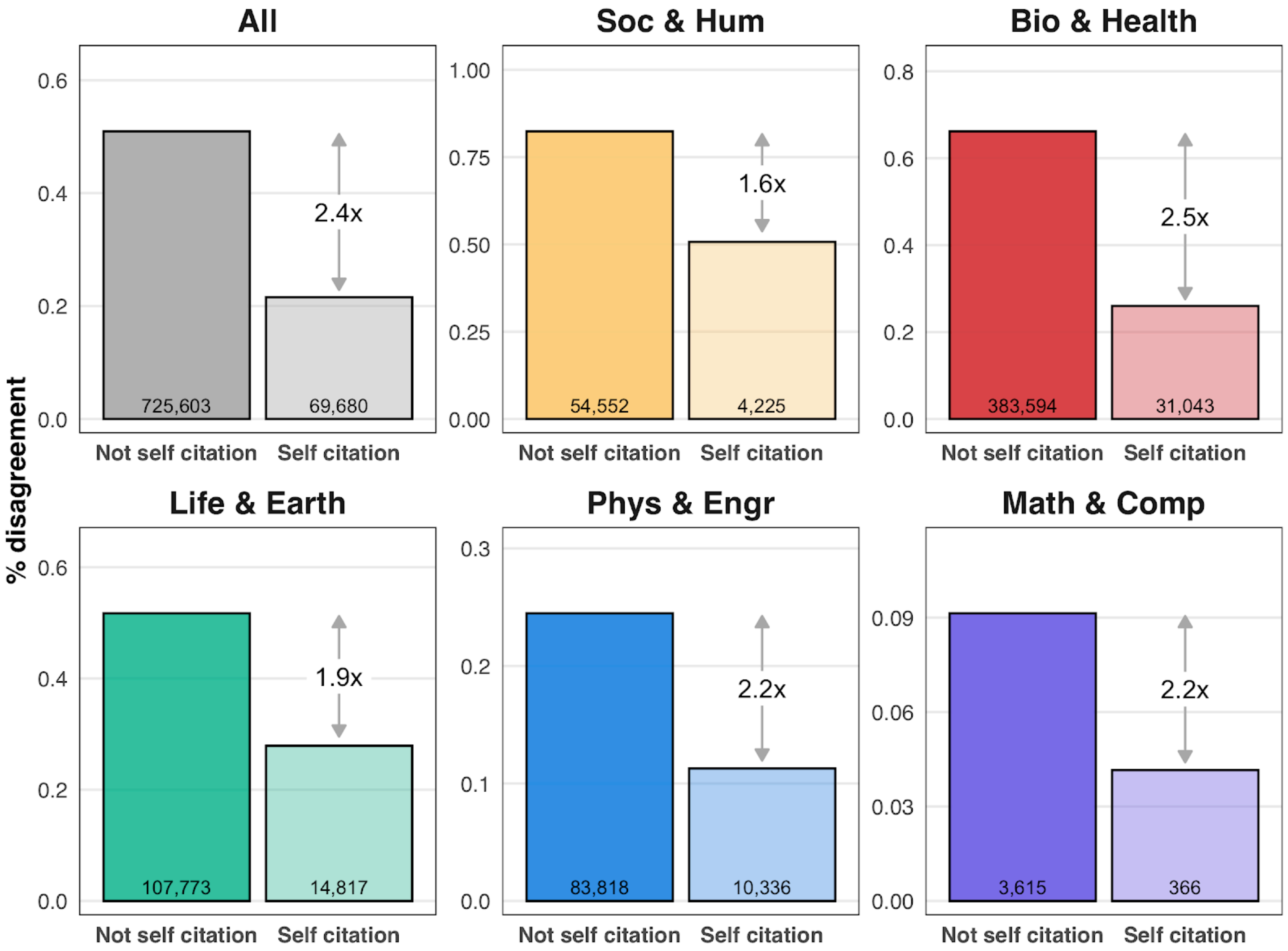
	standalone	+studies	+ideas	+methods	+results	AVERAGE
challenge*	4.0	16.0	60.0	14.0	30.0	**24.8**
conflict*	40.0	82.0	12.0	16.0	78.0	**45.6**
contradict*	70.0	88.0	64.0	64.0	80.0	**73.2**
contrary	38.0	76.0	62.0	78.0	86.0	**68.0**
contrast*	20.0	54.0	80.0	20.0	62.0	**47.2**
controvers*	82.0	92.0	84.0	94.0	92.0	**88.8**
debat*	80.0	86.0	78.0	84.0	84.0	**82.4**
differ*	0.0	8.0	4.0	18.0	20.0	**10.0**
disagree*	50.0	86.0	70.0	72.0	90.0	**73.6**
disprov*	58.0	84.0	78.0	74.0	66.0	**72.0**
no consensus	94.0	98.0	92.5	98.0	92.0	**94.9**
questionable	28.0	38.0	72.0	80.0	10.0	**45.6**
refut*	68.0	72.0	72.0	54.0	70.0	**67.2**
AVERAGE	**48.6**	**67.7**	**63.7**	**58.9**	**66.2**	

% Valid
100
50
0

Less disagreement for older papers

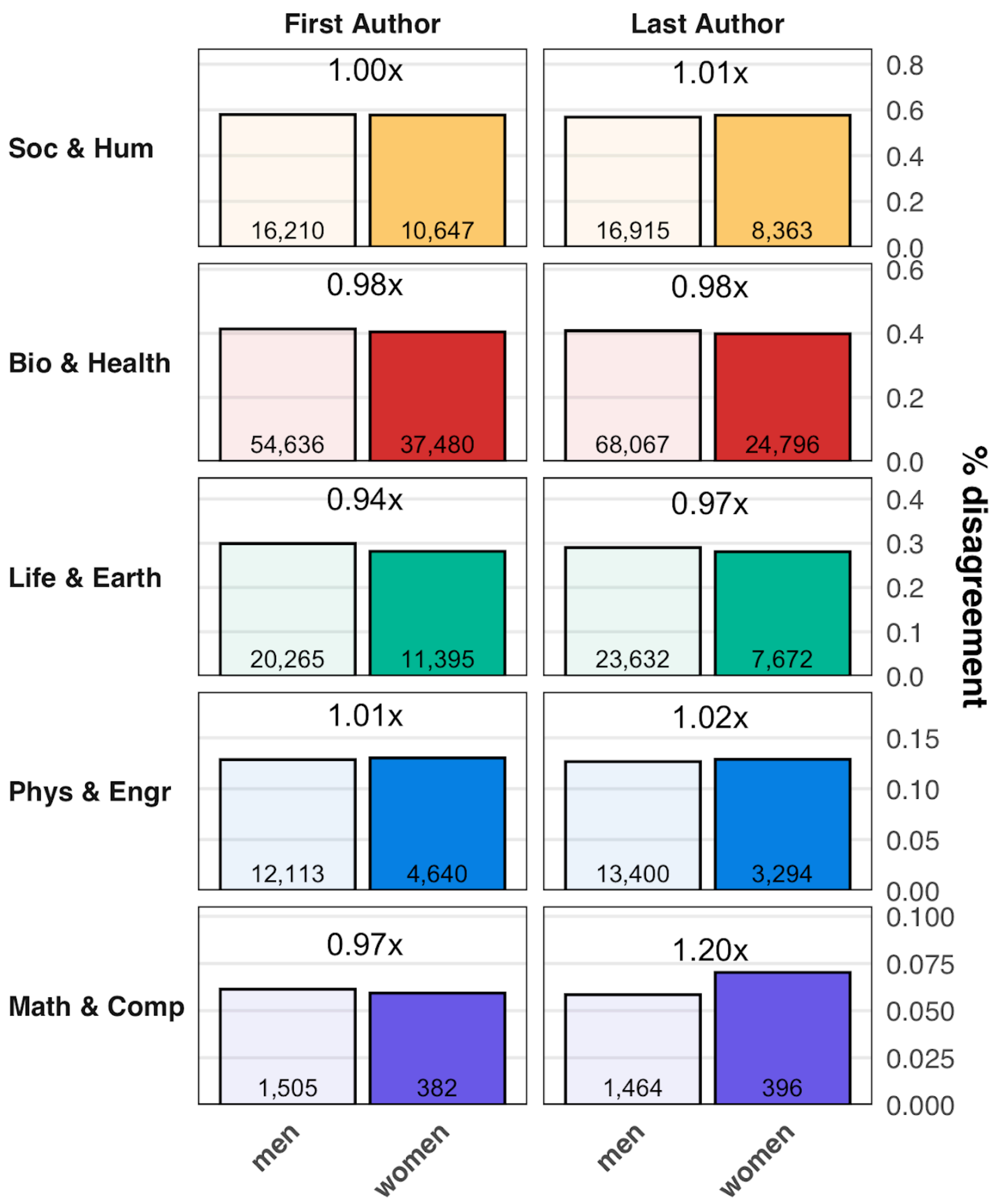


Disagreement and self-citation



dakotamurray.me/talks/2020-northeastern/

Disagreement and gender



Appendix—Mobility

Mobility is central to science

Institutionalized in evaluation

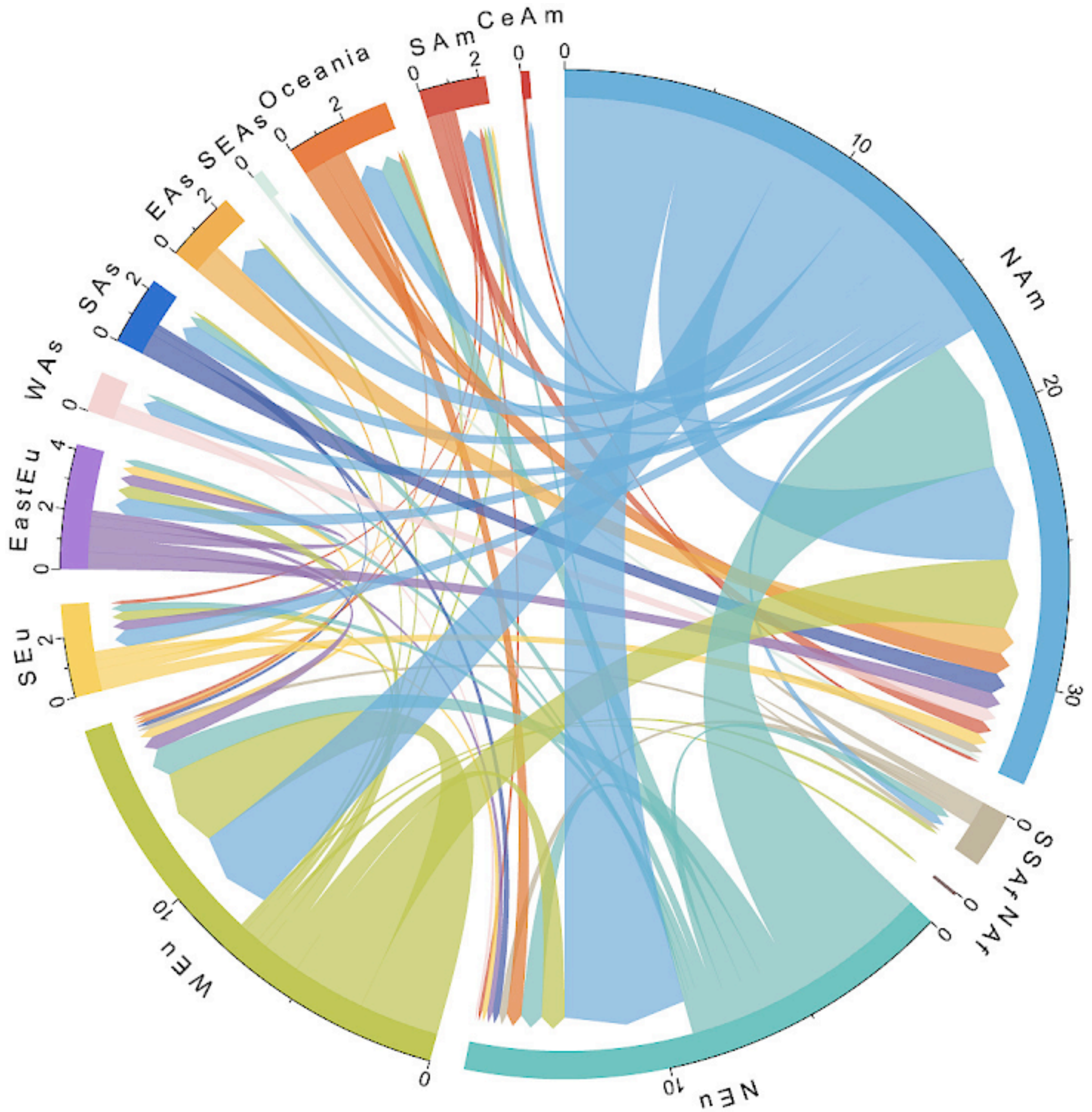
Article 19: Requirement for International Visits
When applying for promotion to full professor or equivalent rank, the applicants who were born after January 1, 1970, must complete at least a 6-month international visit.

- Hangzhou Dianzi University

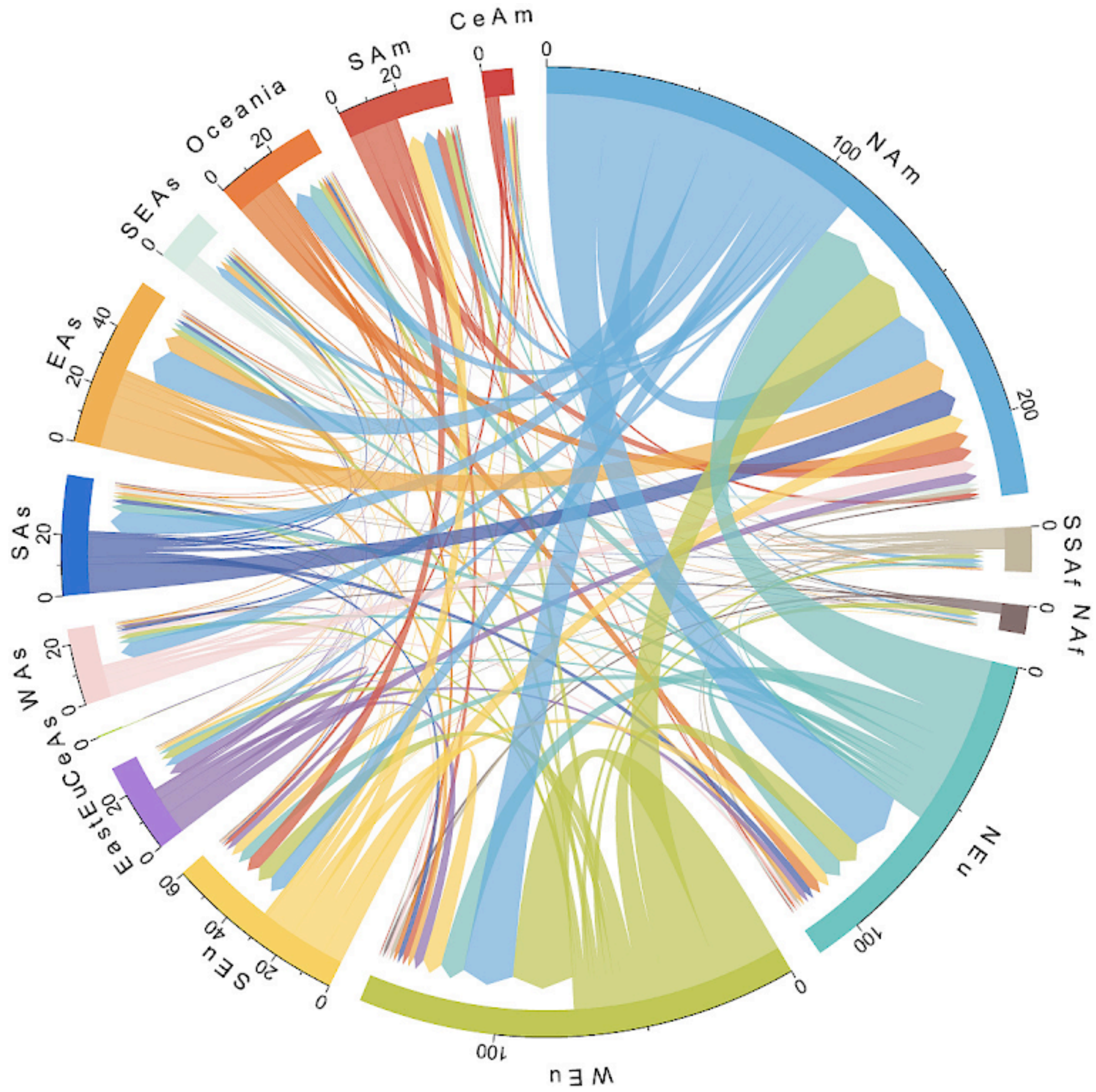
Thanks to Fei She

Science is becoming increasingly global

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



1970

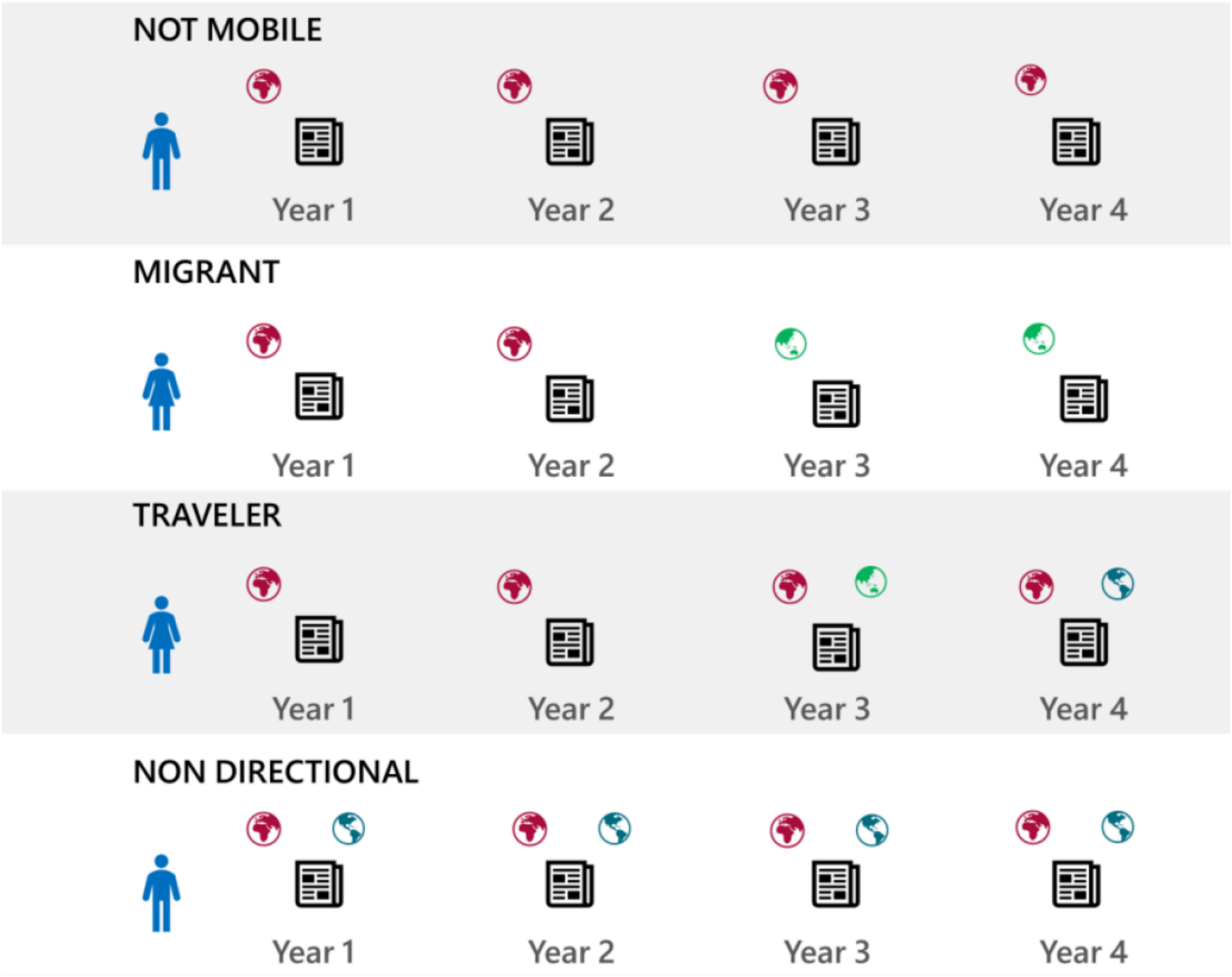


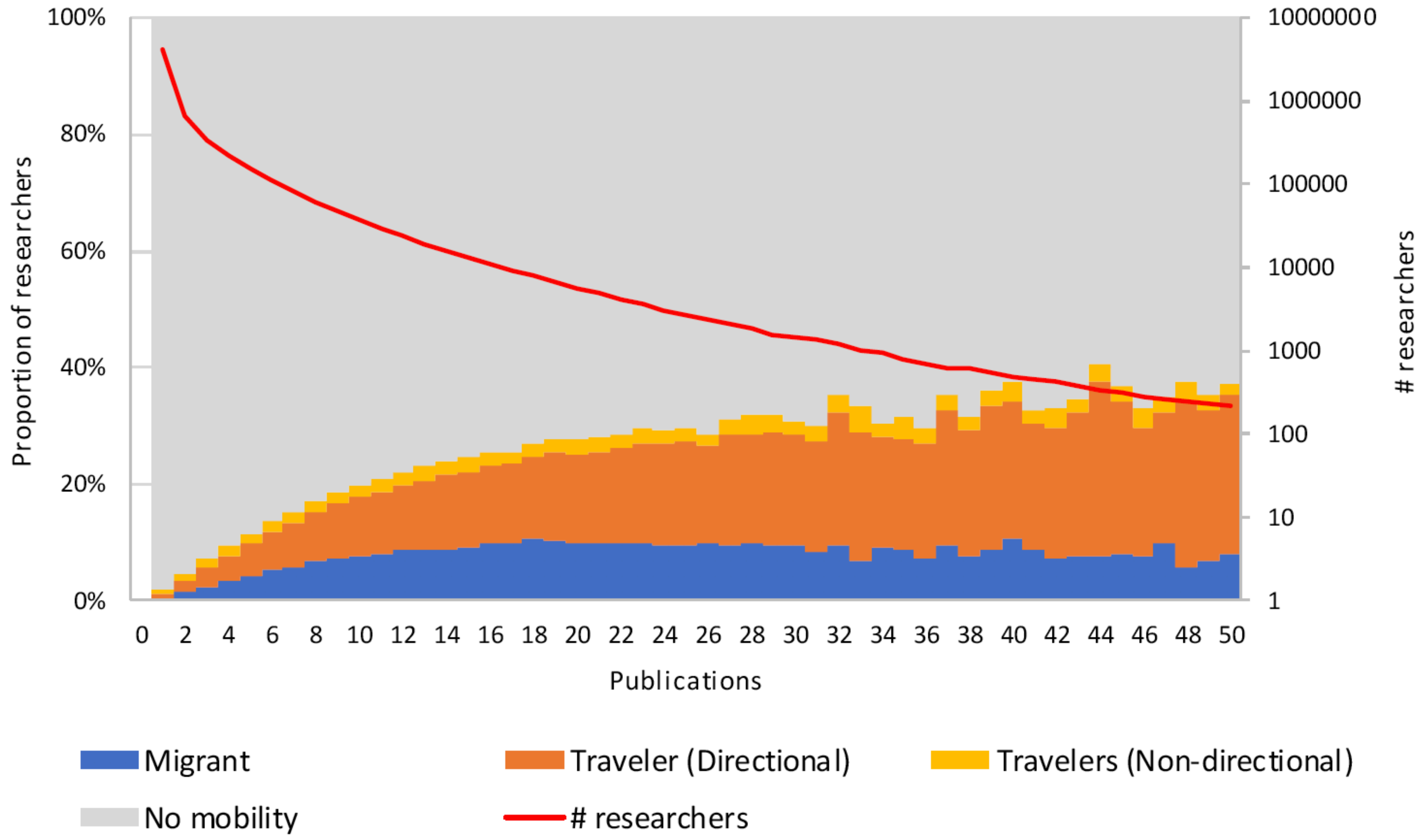
2014

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*.

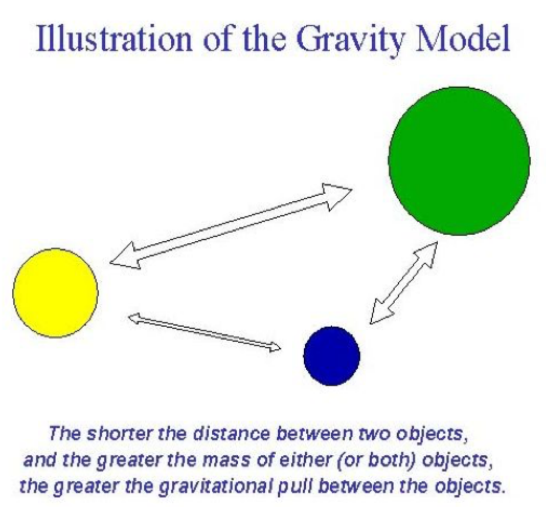
Mobility is complicated





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*.

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

Target Context

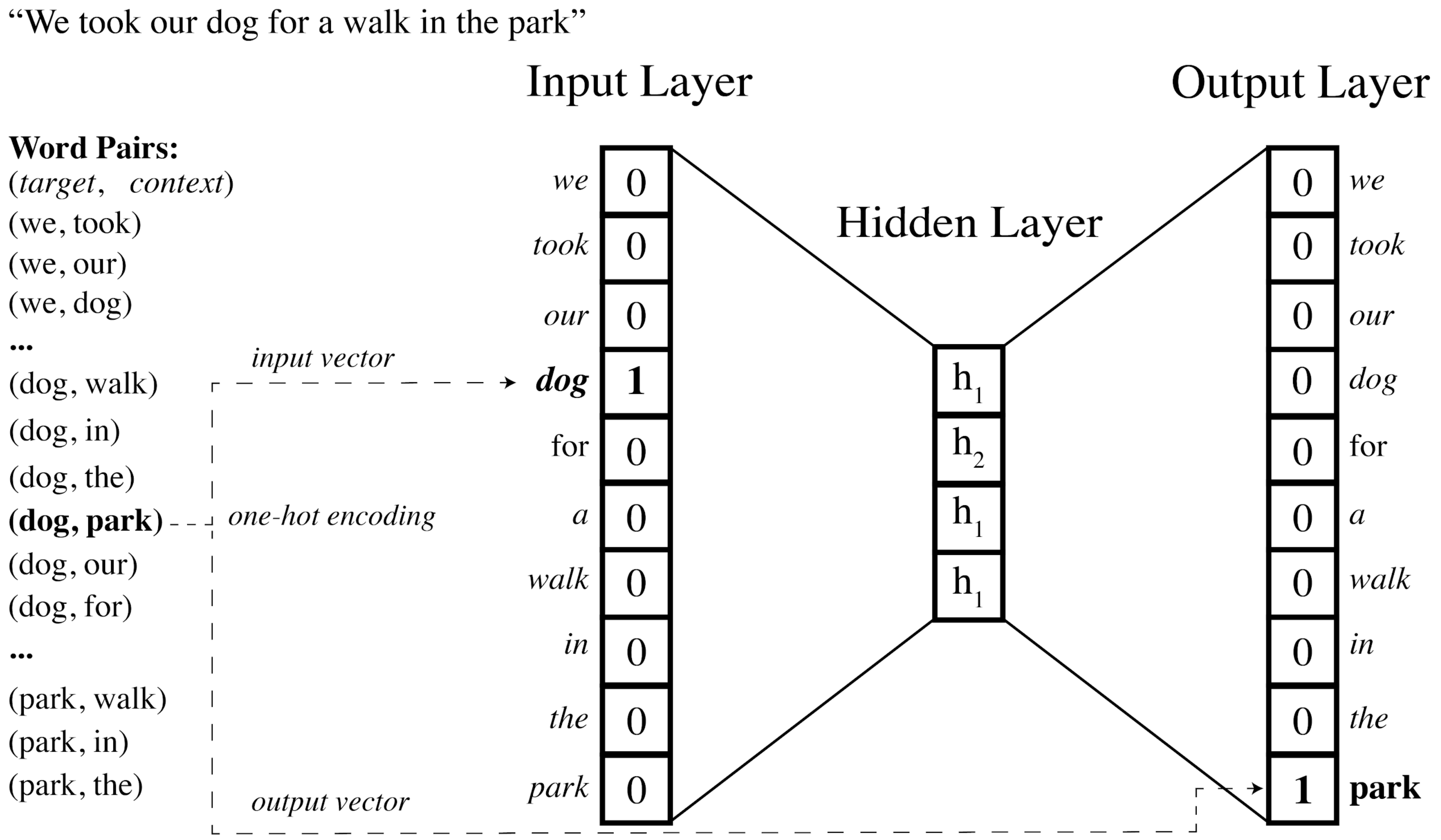
Let's assume that we can calculate the conditional probability with a function (e.g. dot product) of word **vectors** and **learn those vectors with a neural network**.

Learning word embeddings

Train a neural network to predict context words given a target

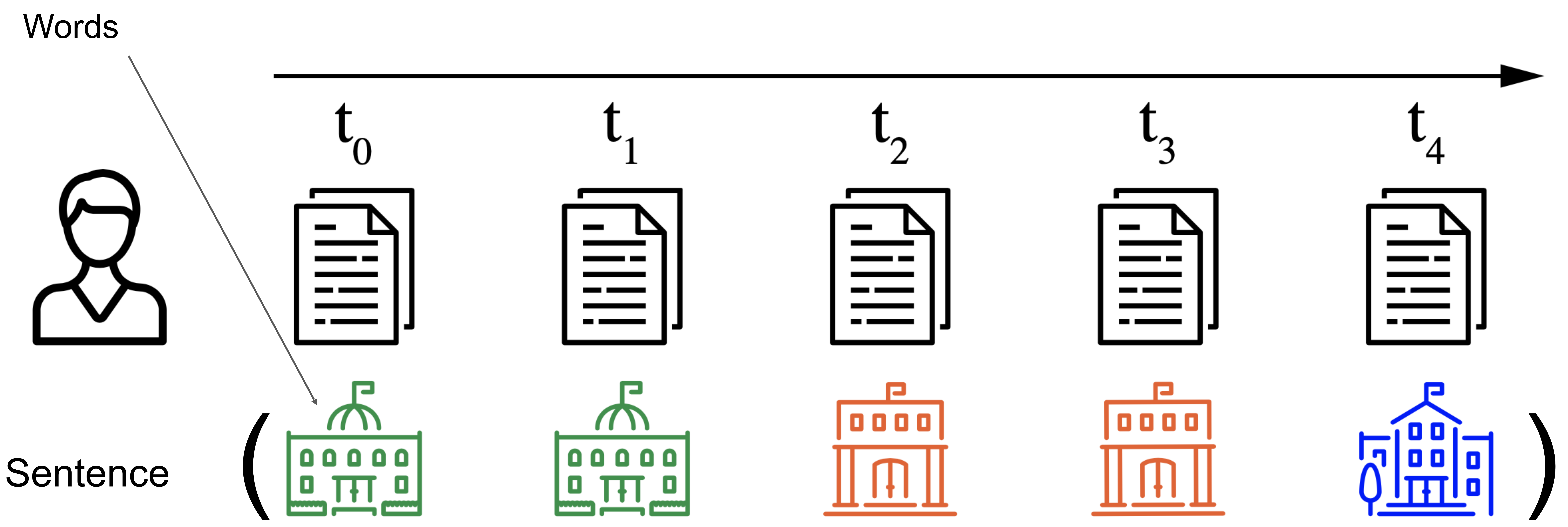
The hidden layer maps targets to concepts!

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



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

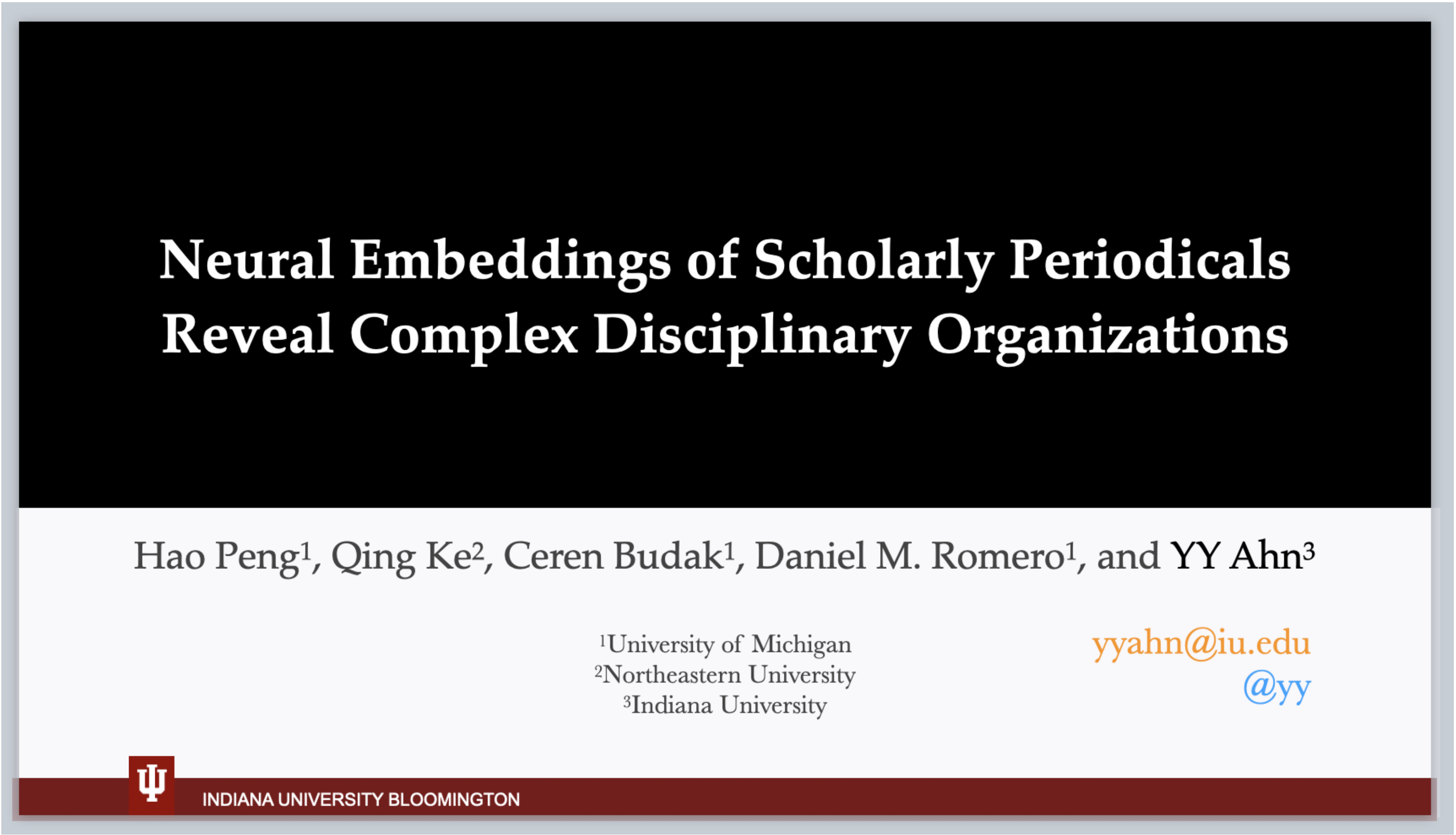
Any "sentences" — a sequence of elements from a finite vocabulary — work! We can use **trajectories of scientists** as **sentences** and **organizations** as **words**.



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**.

(

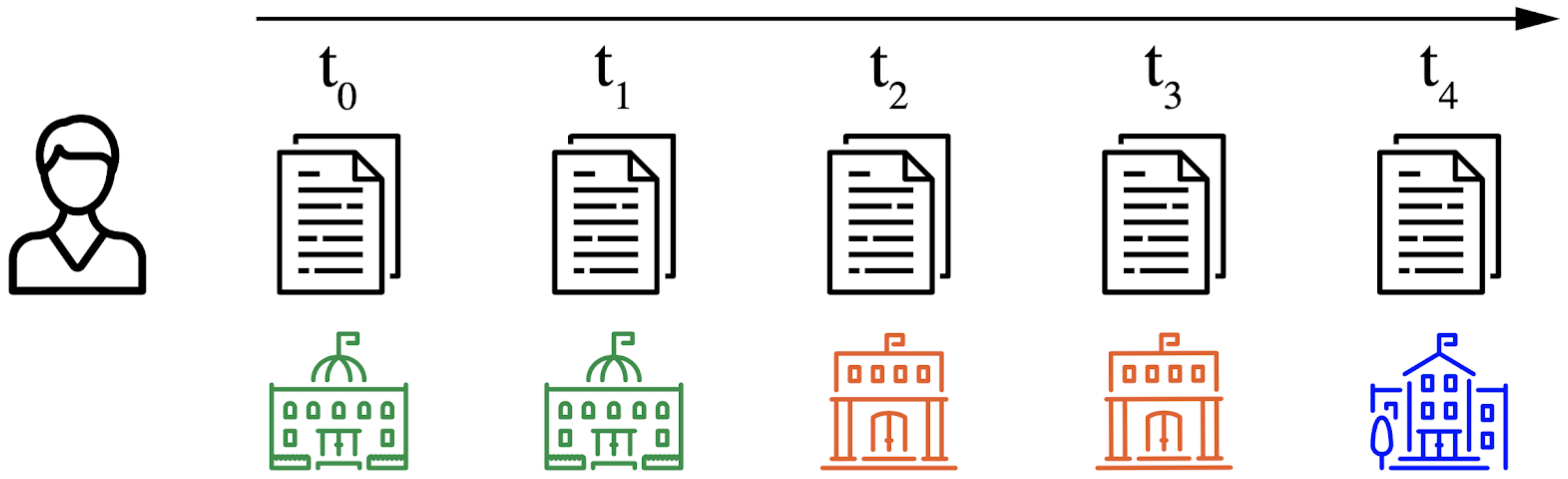


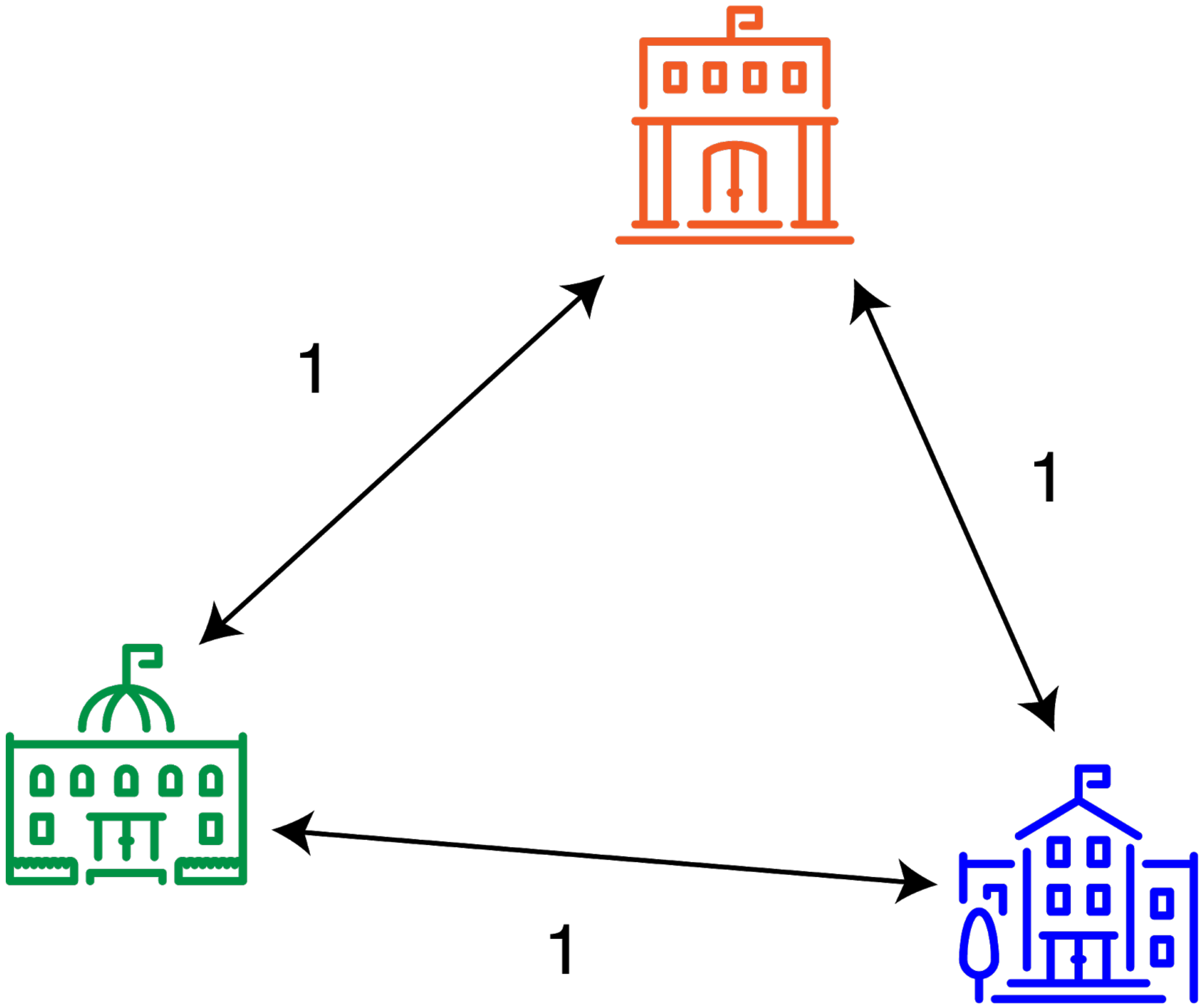
IC2S2 "embedding" session

)

Each a trajectory of organizations

Derive “flux” from scientists career trajectories





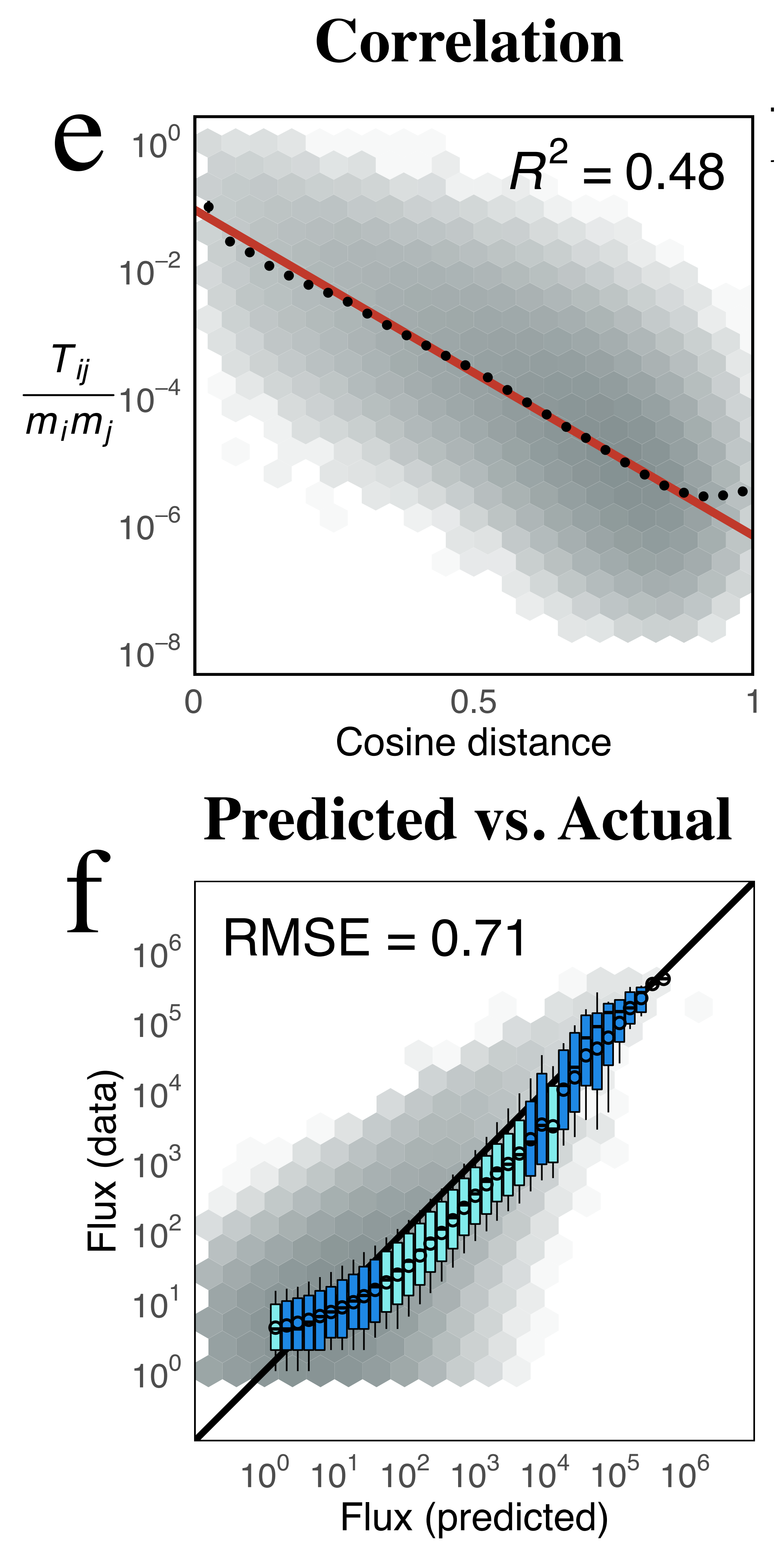
Aside: why does it work?

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!

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

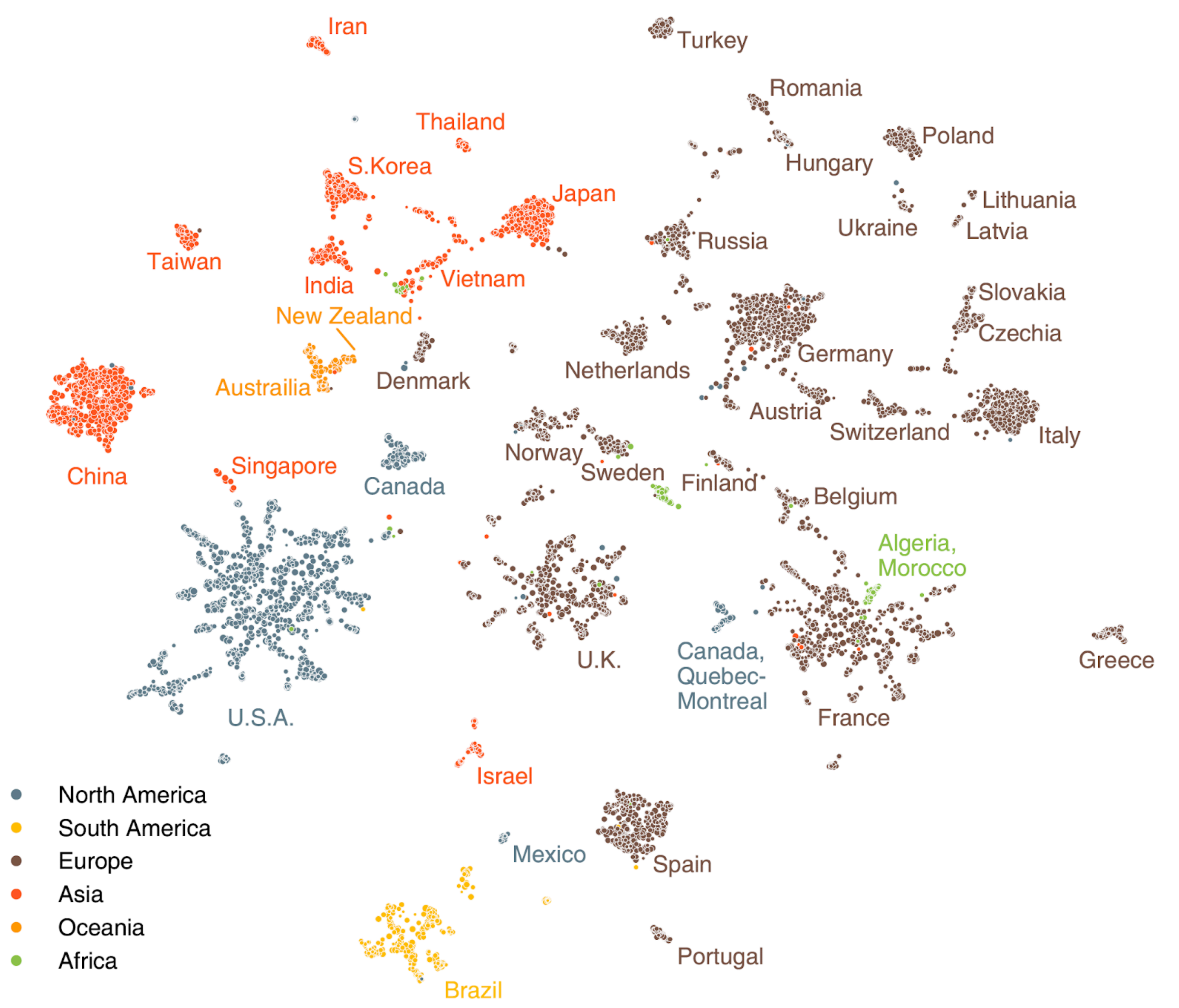
$$\frac{T_{ij}}{m_i m_j} = f(r_{ij})$$

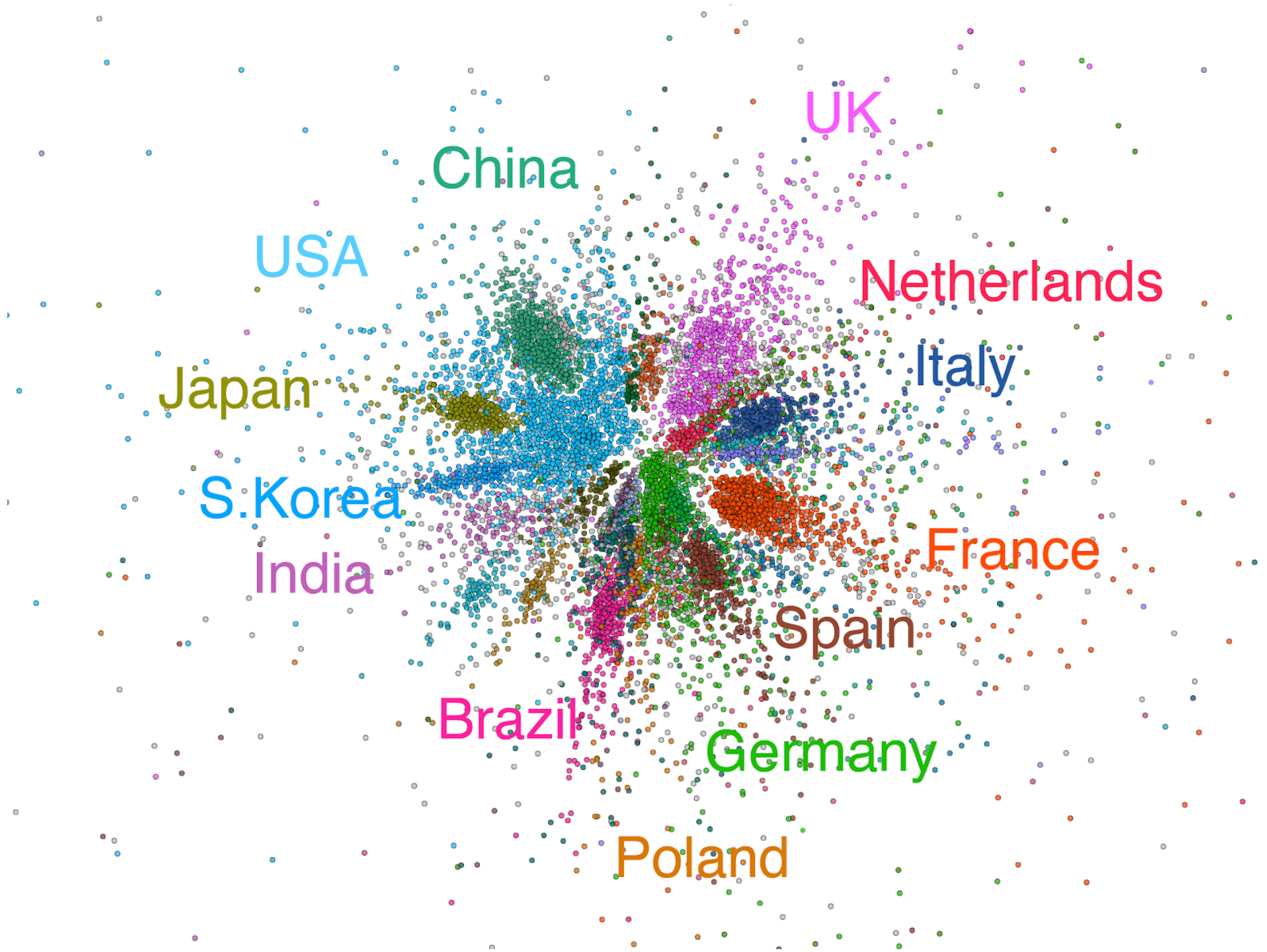


Murray, D., Yoon, J., Kojaku, S., Costas, R., Jung, W.-S., Milojević, S., & Ahn, Y.-Y. (2020). Unsupervised embedding of trajectories captures the latent structure of mobility. *(Under review, Nature)*

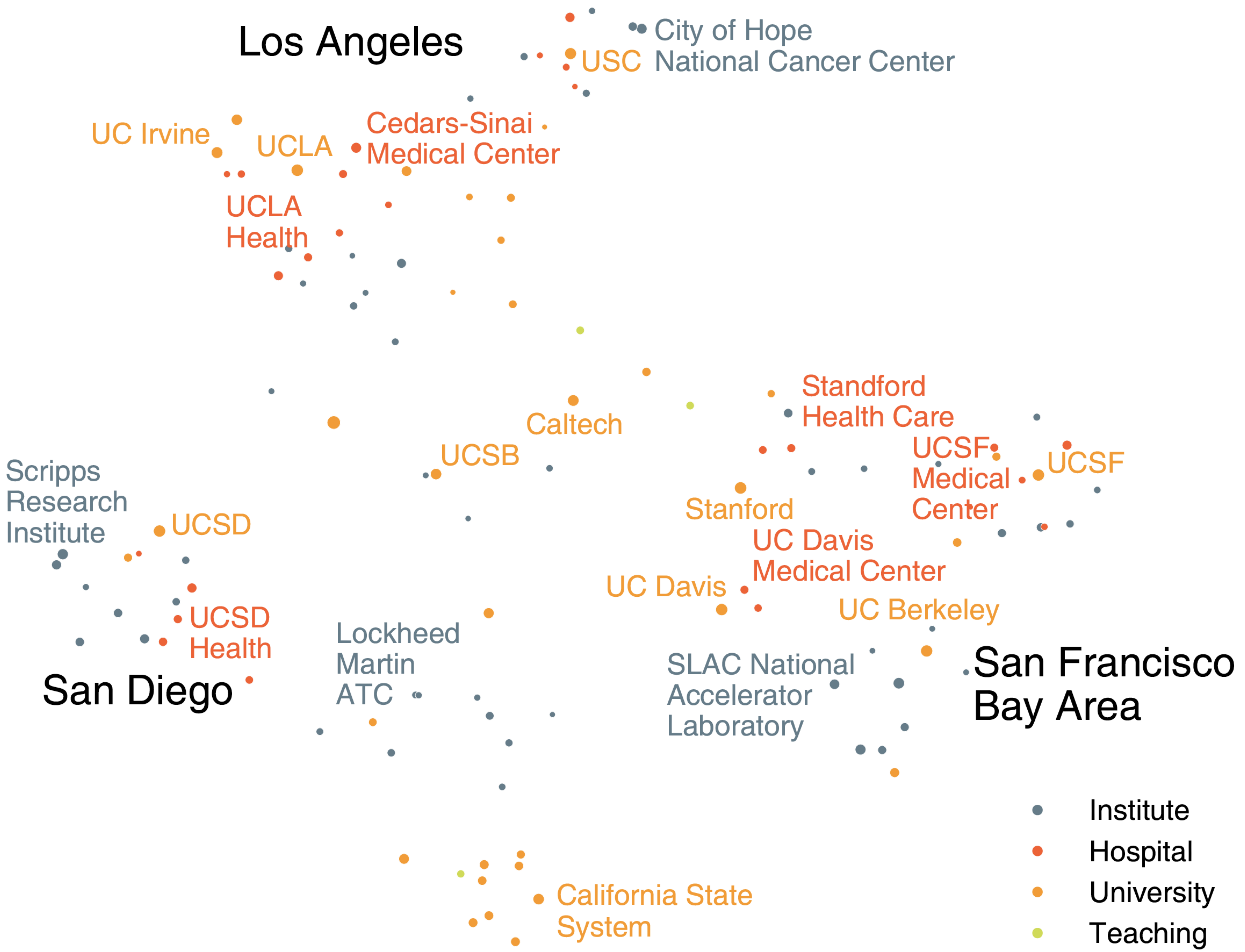
Why not networks?

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

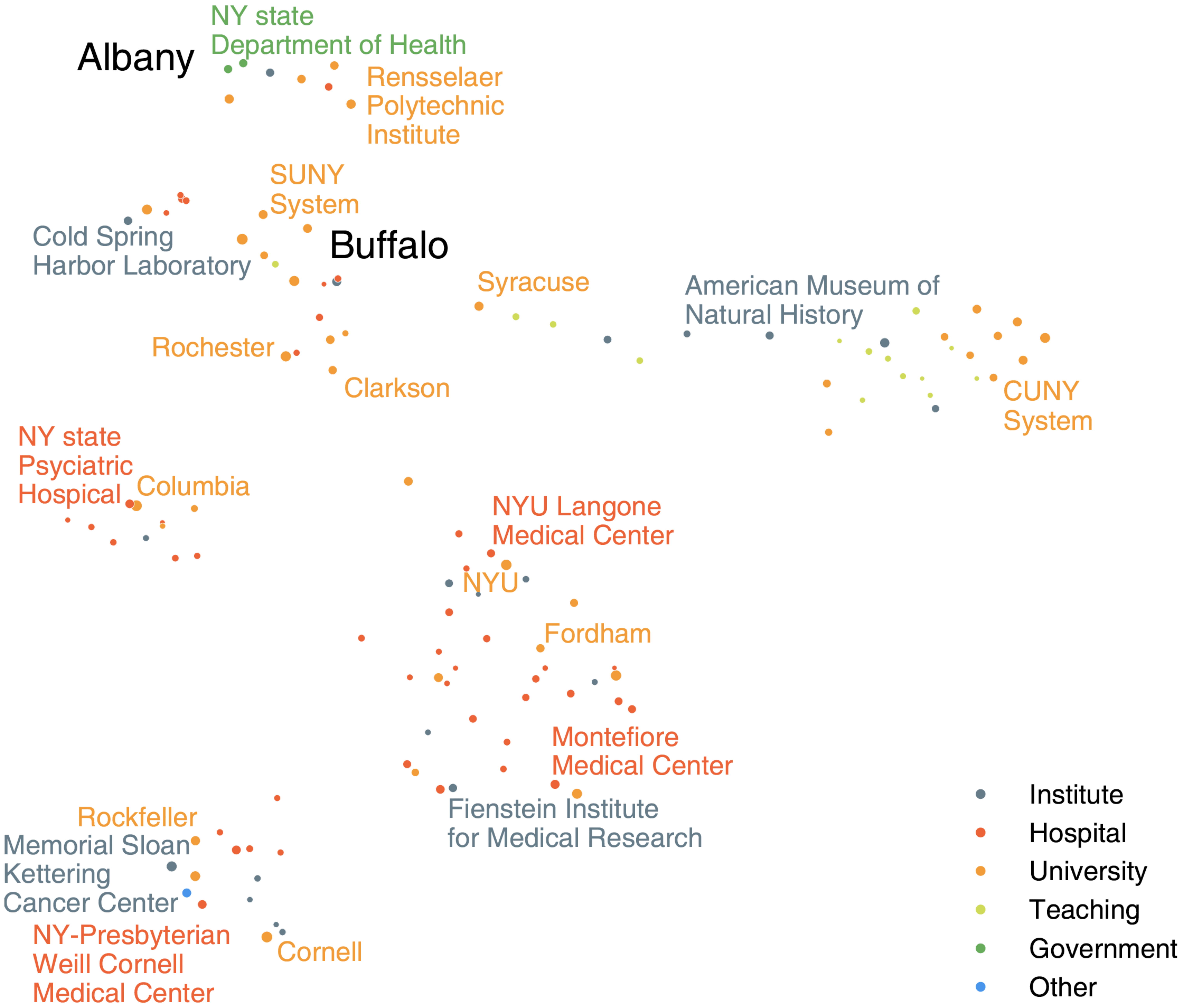




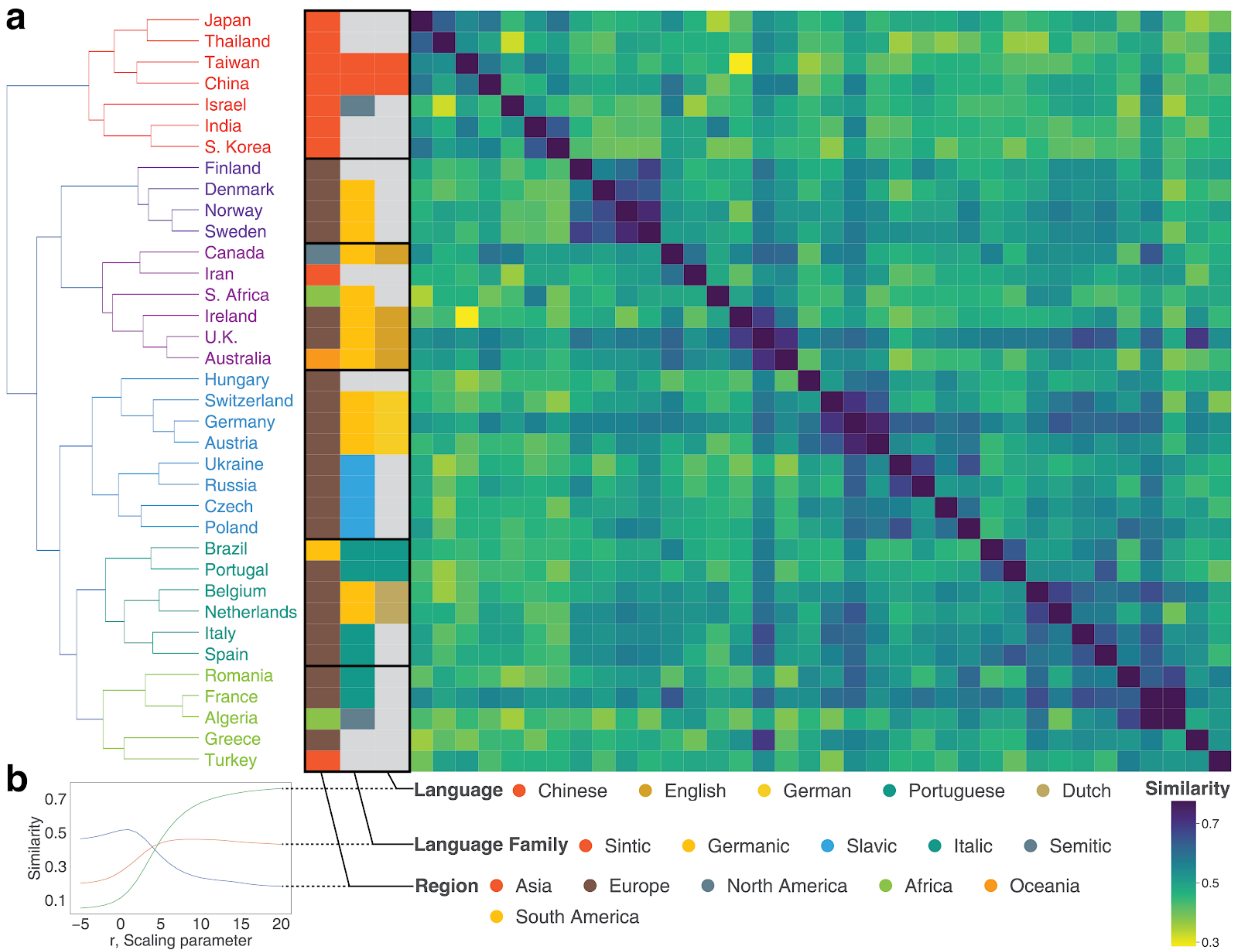
California Structure



New York Structure



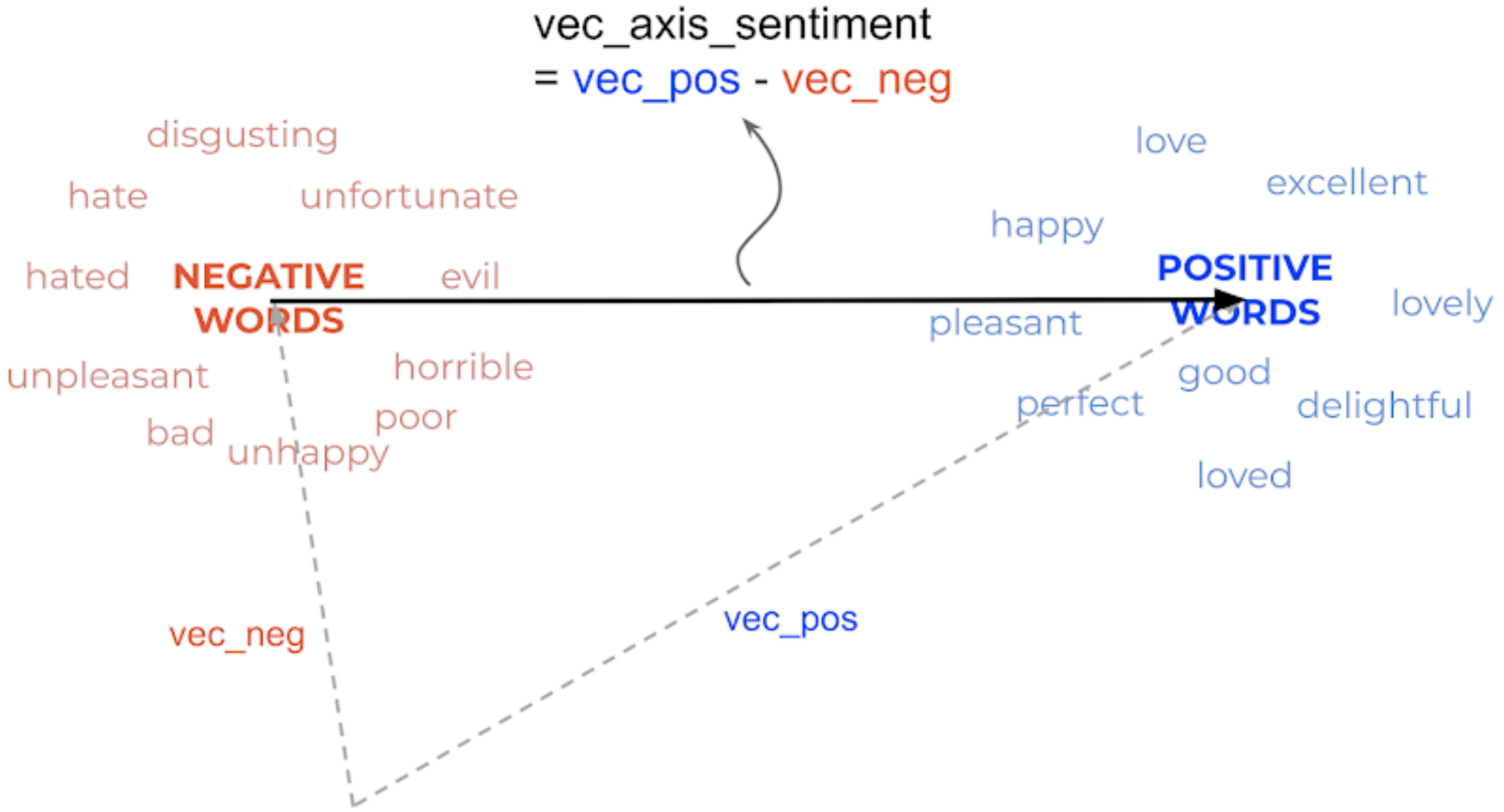
Geography, then language, structure the vector space



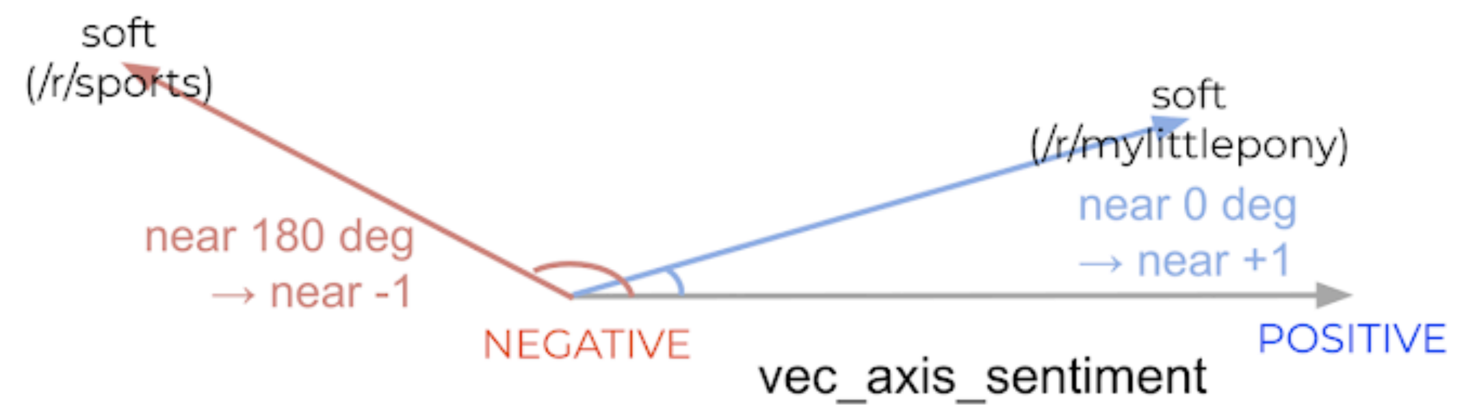
But also prestige…

Leverage the semantic properties of the embedding

Step 1: Define a semantic axis

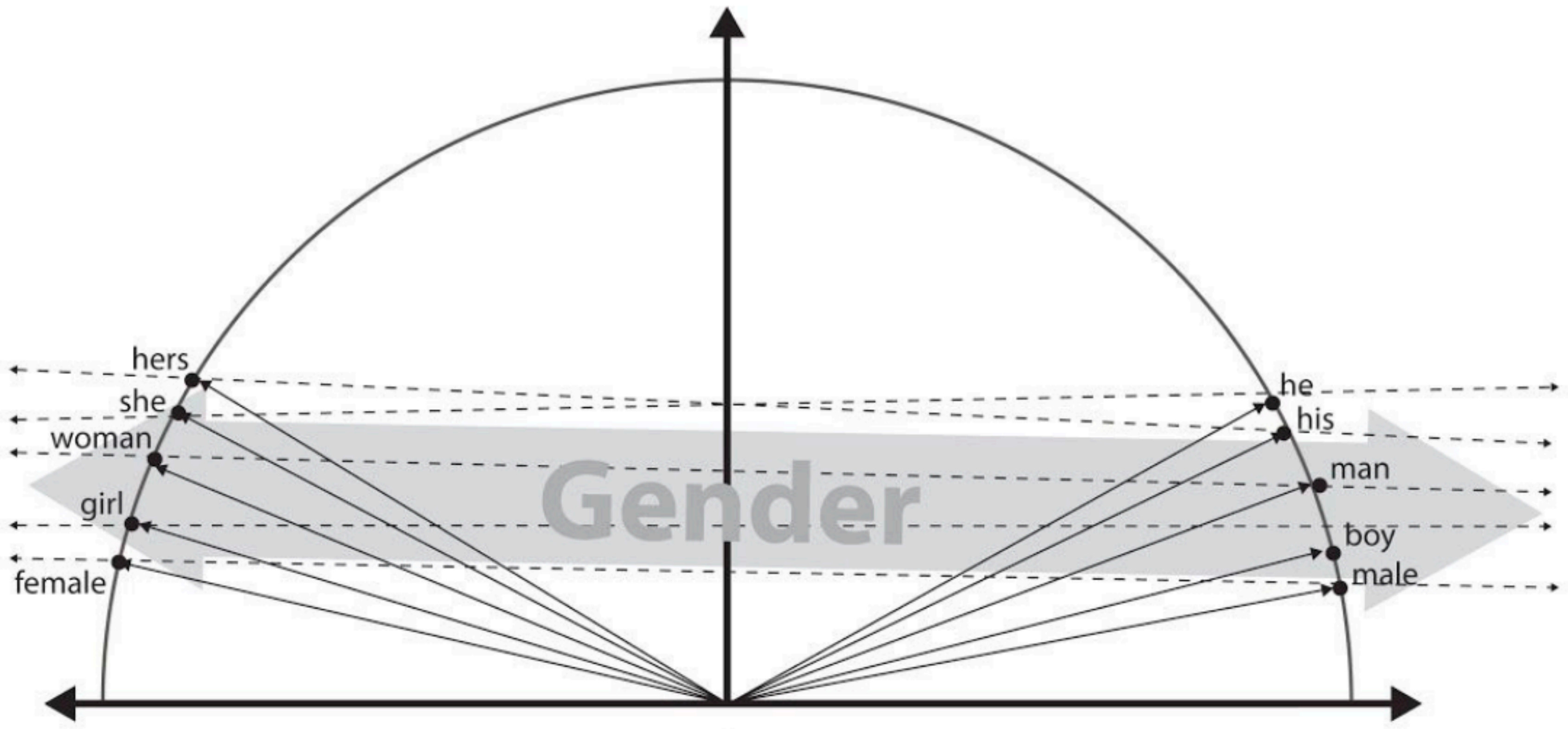


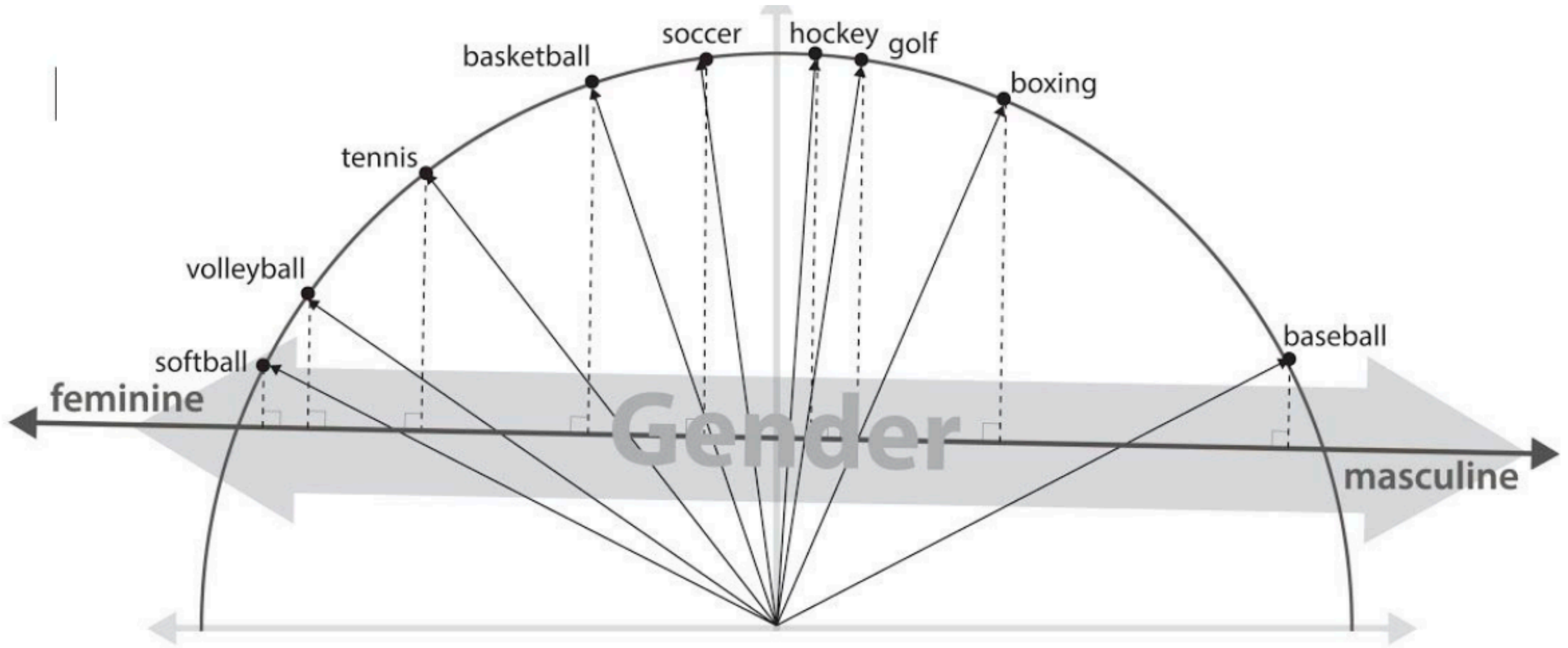
Step 2. Locate words on the semantic axis



Compute a cosine similarity between a word vector and axis vector

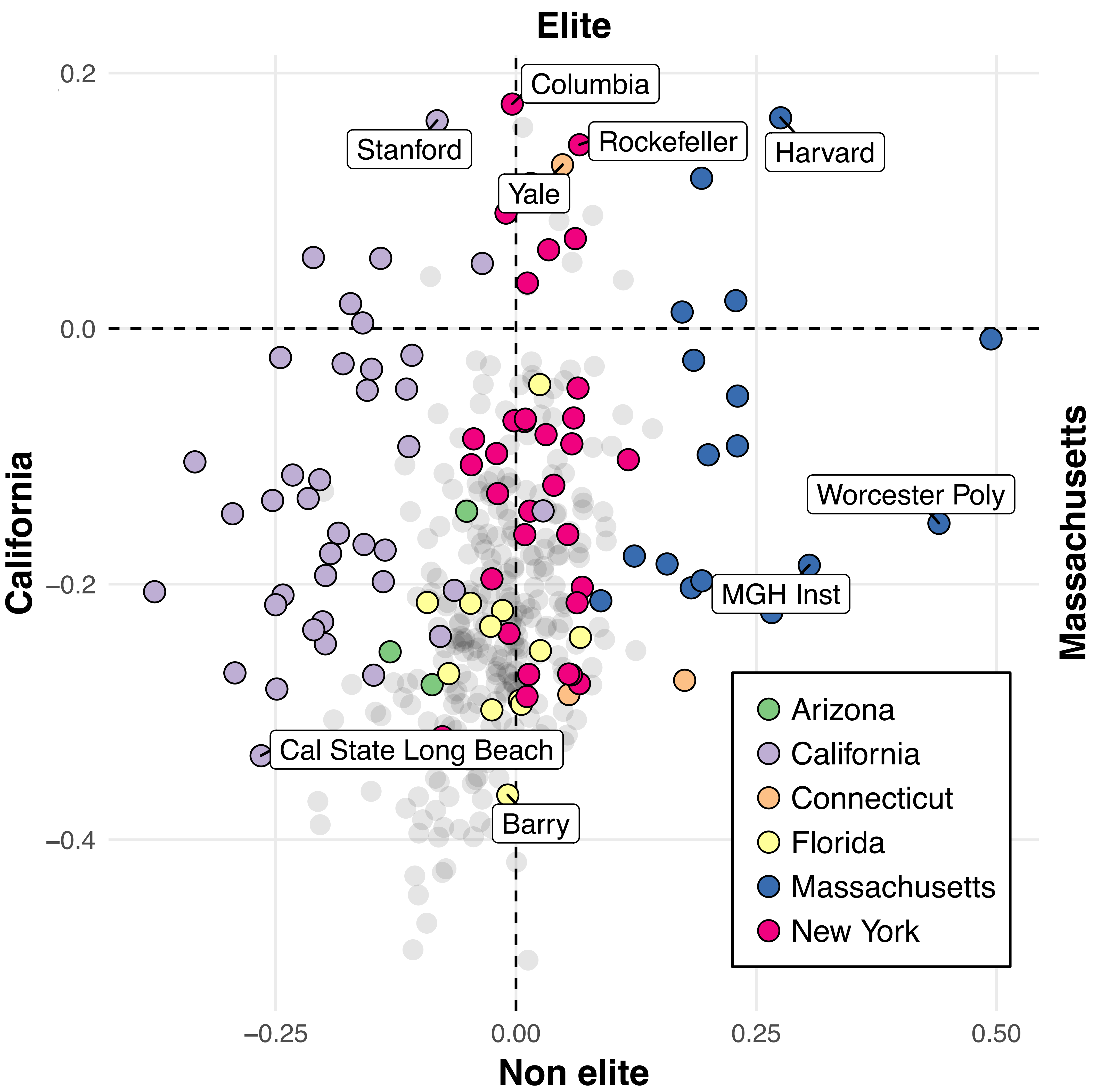
An, J., Kwak, H., & Ahn, Y.-Y. (2018). SemAxis: A Lightweight Framework to Characterize Domain-Specific Word Semantics Beyond Sentiment. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2450–2461.



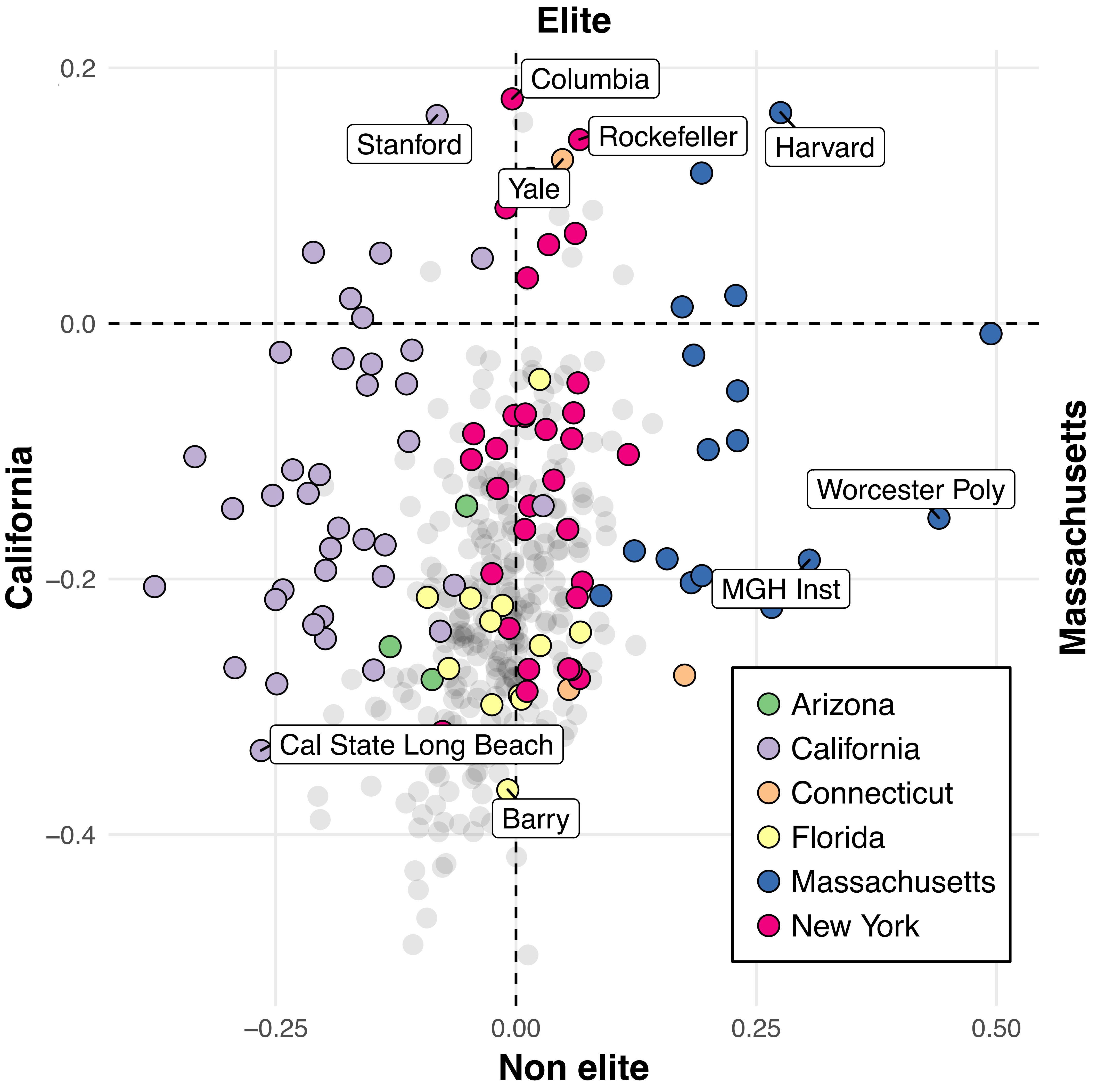


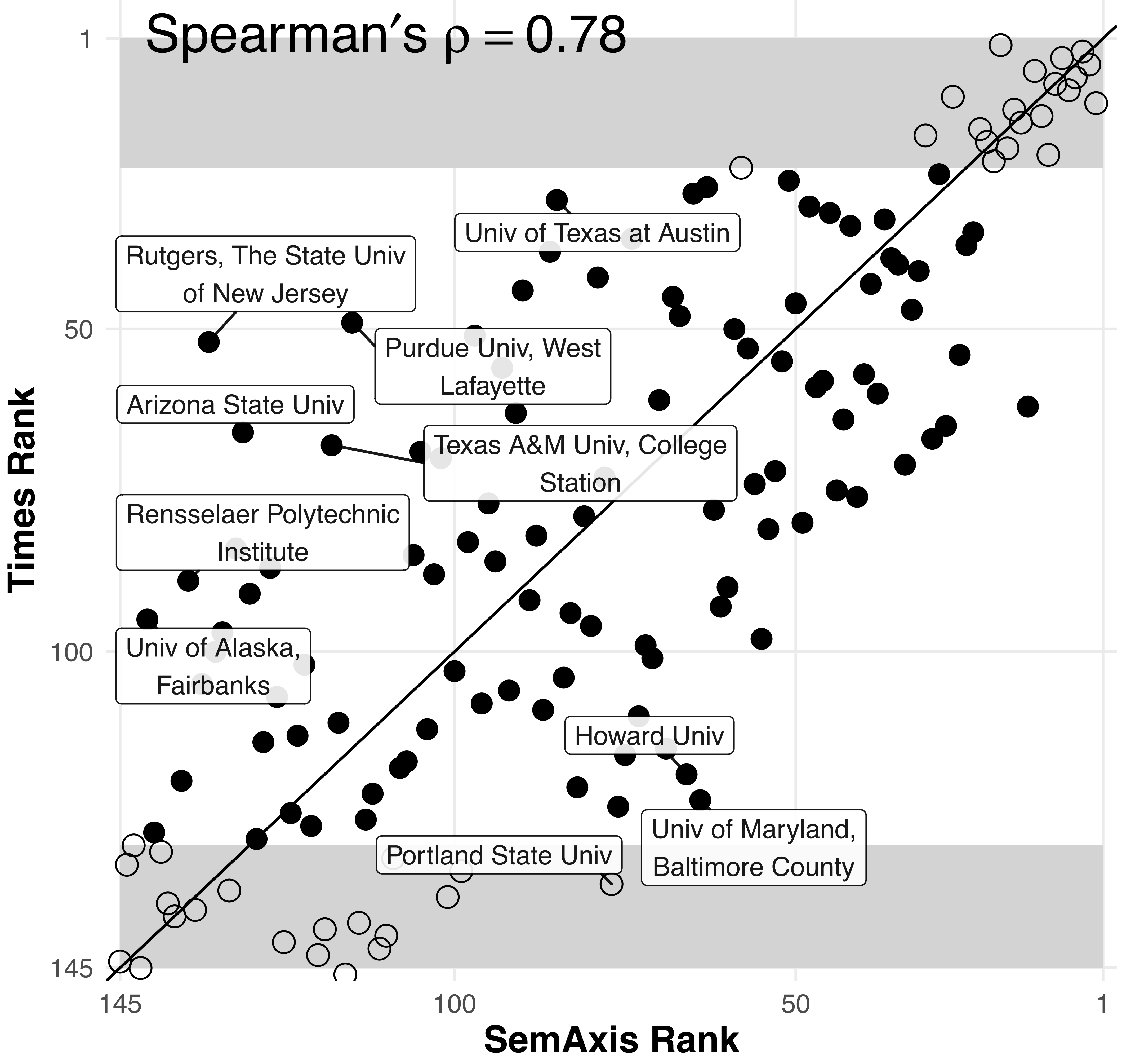
Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*, *84*(5), 905–949.

SemAxis using Geography and Prestige



SemAxis reconstructs university prestige

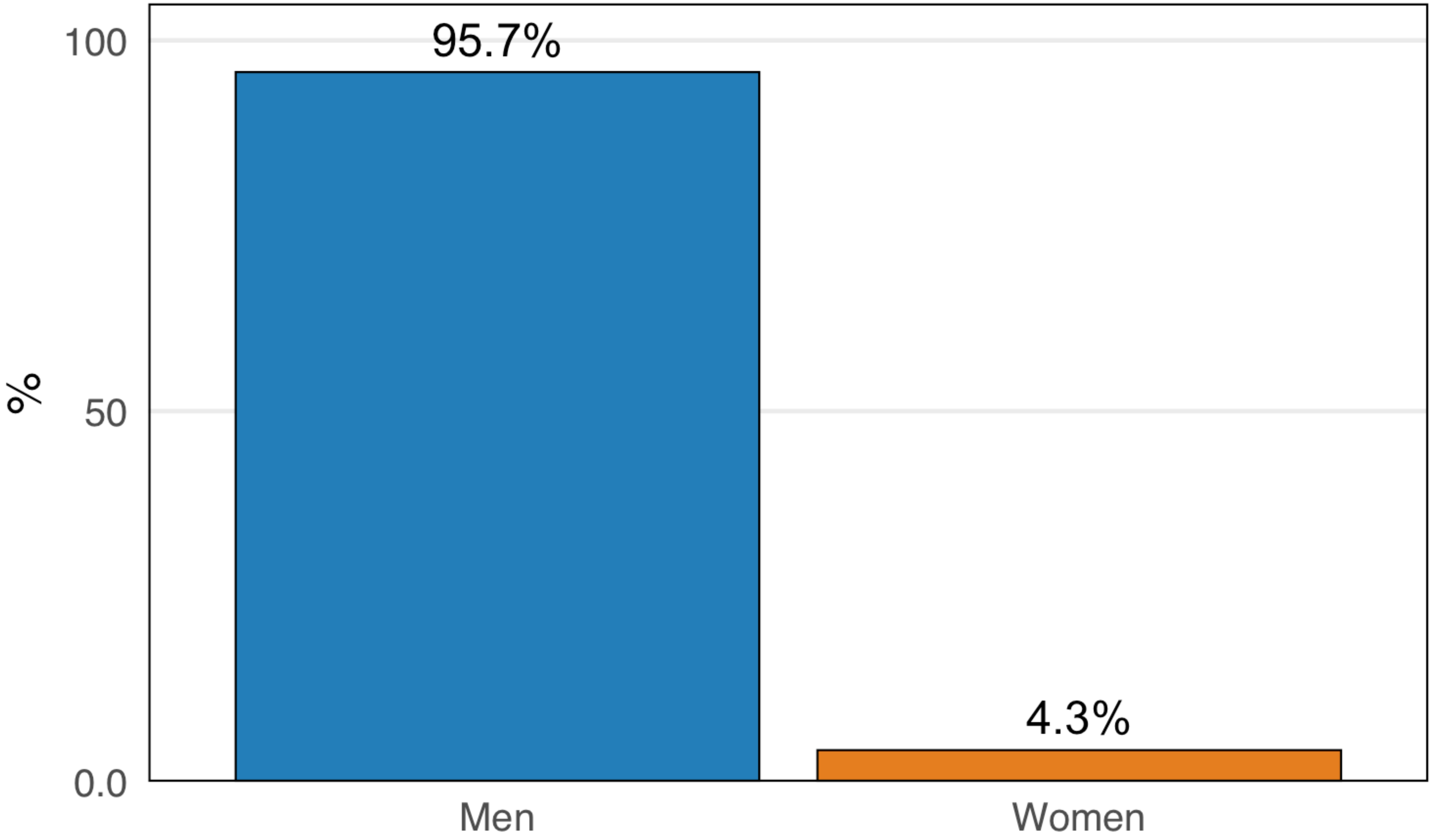




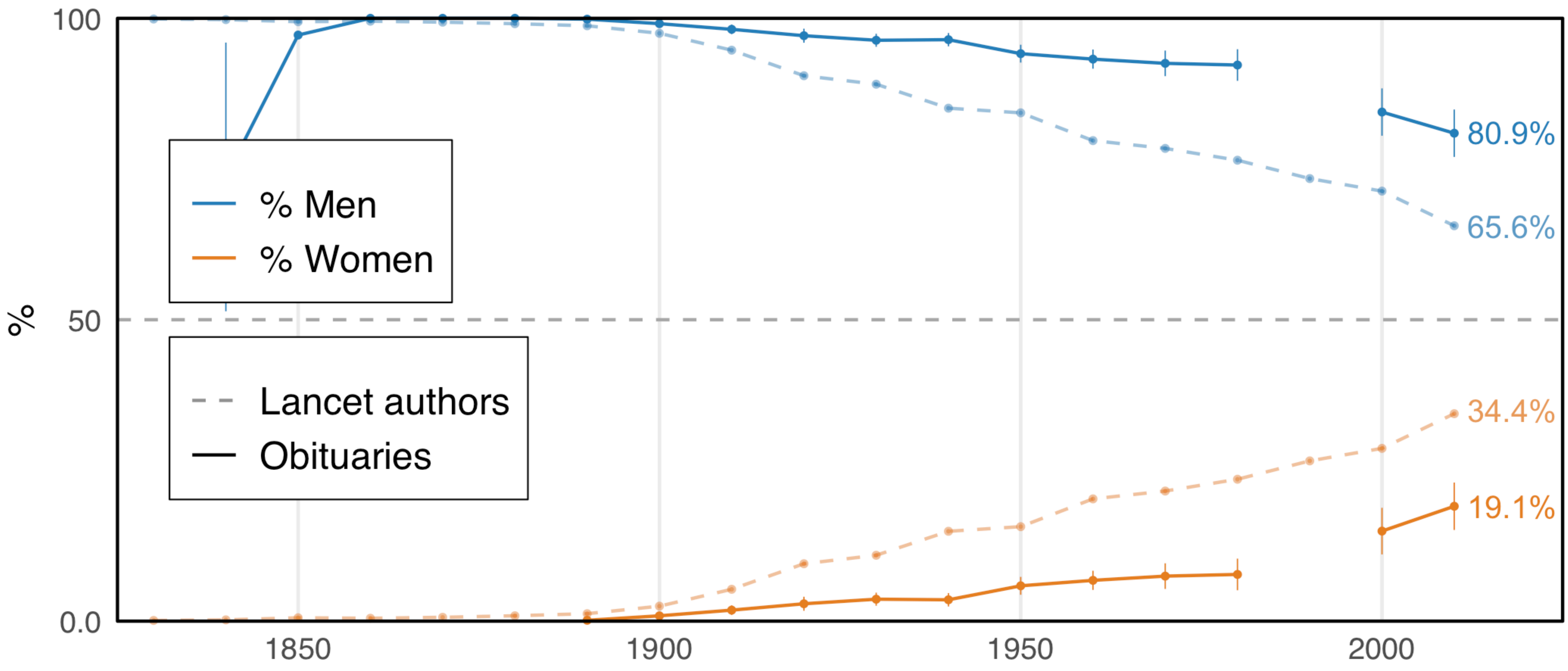
Appendix—Obituaries

Number of obituaries over time

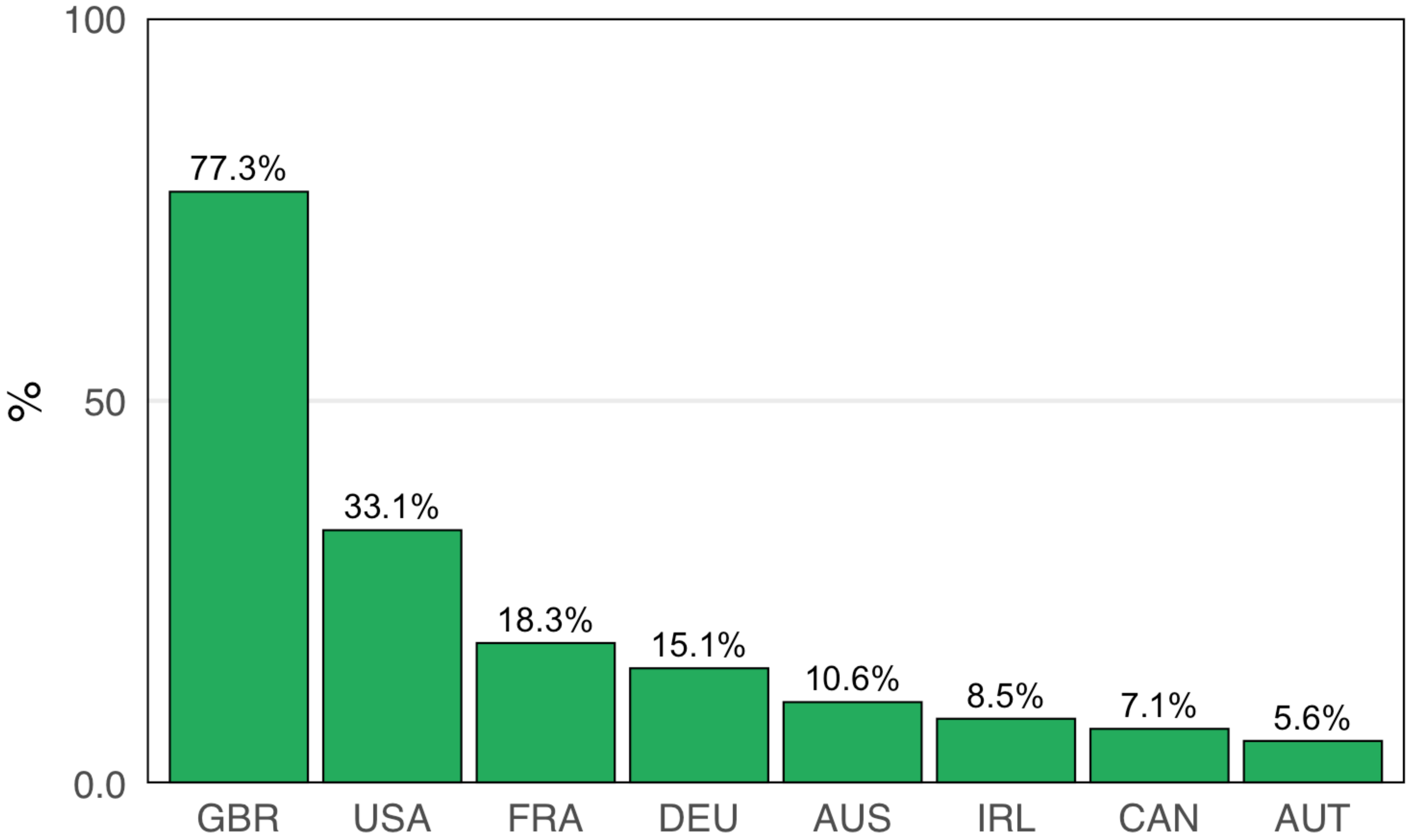
Largely dominated by men



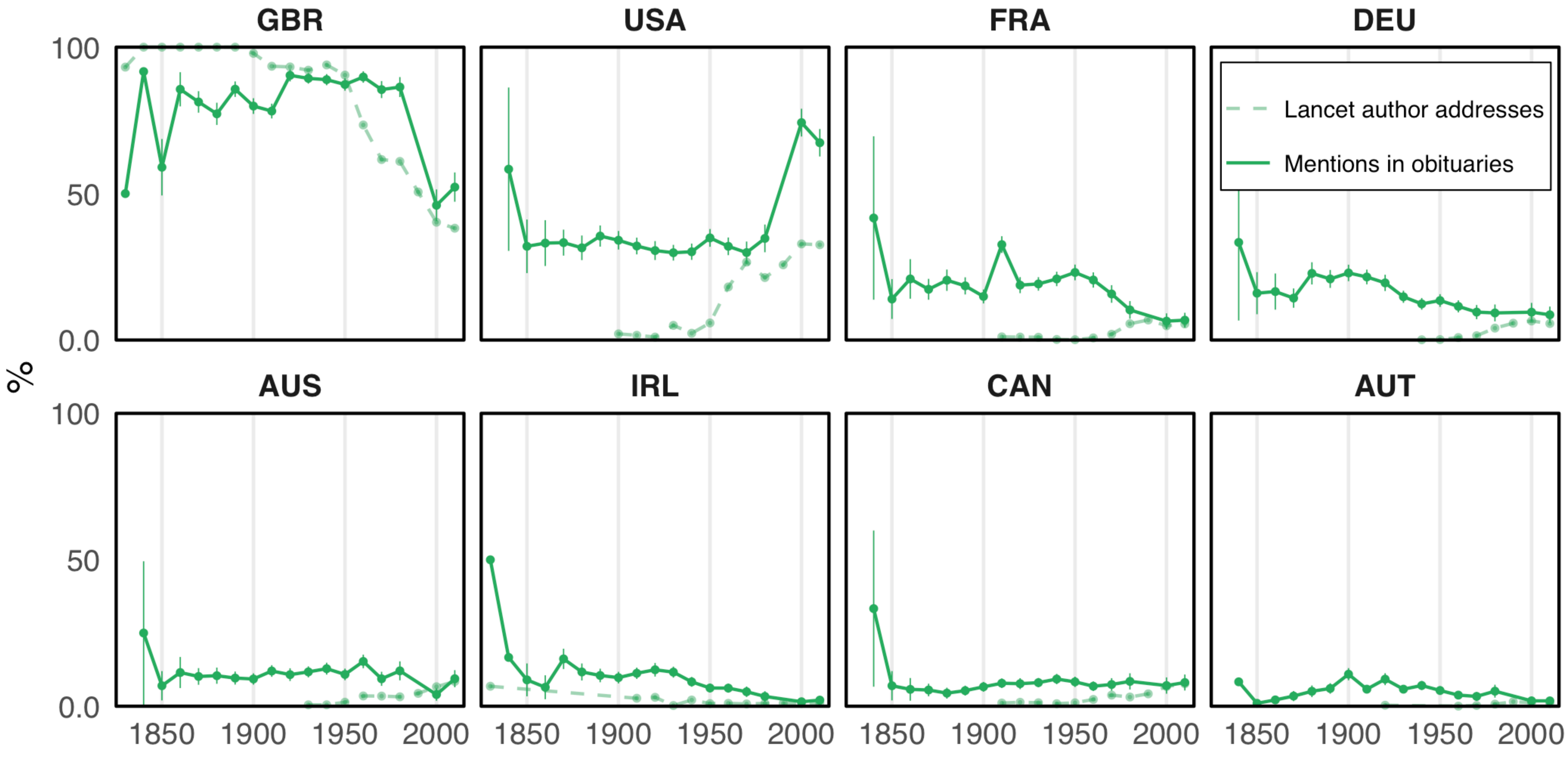
Gender representation has improved, but still behind



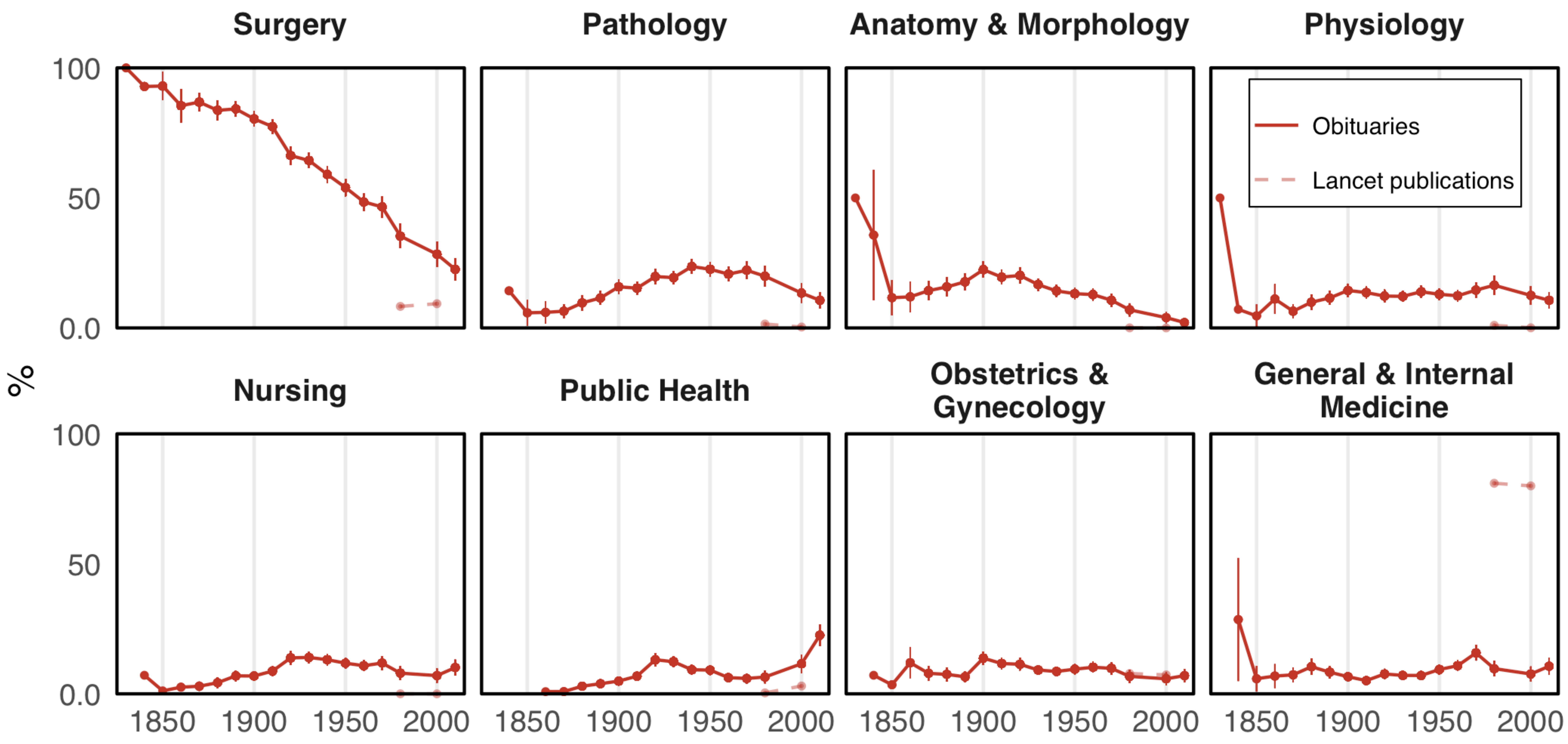
Mostly people from the U.K.



More U.S. authors in the 2000s

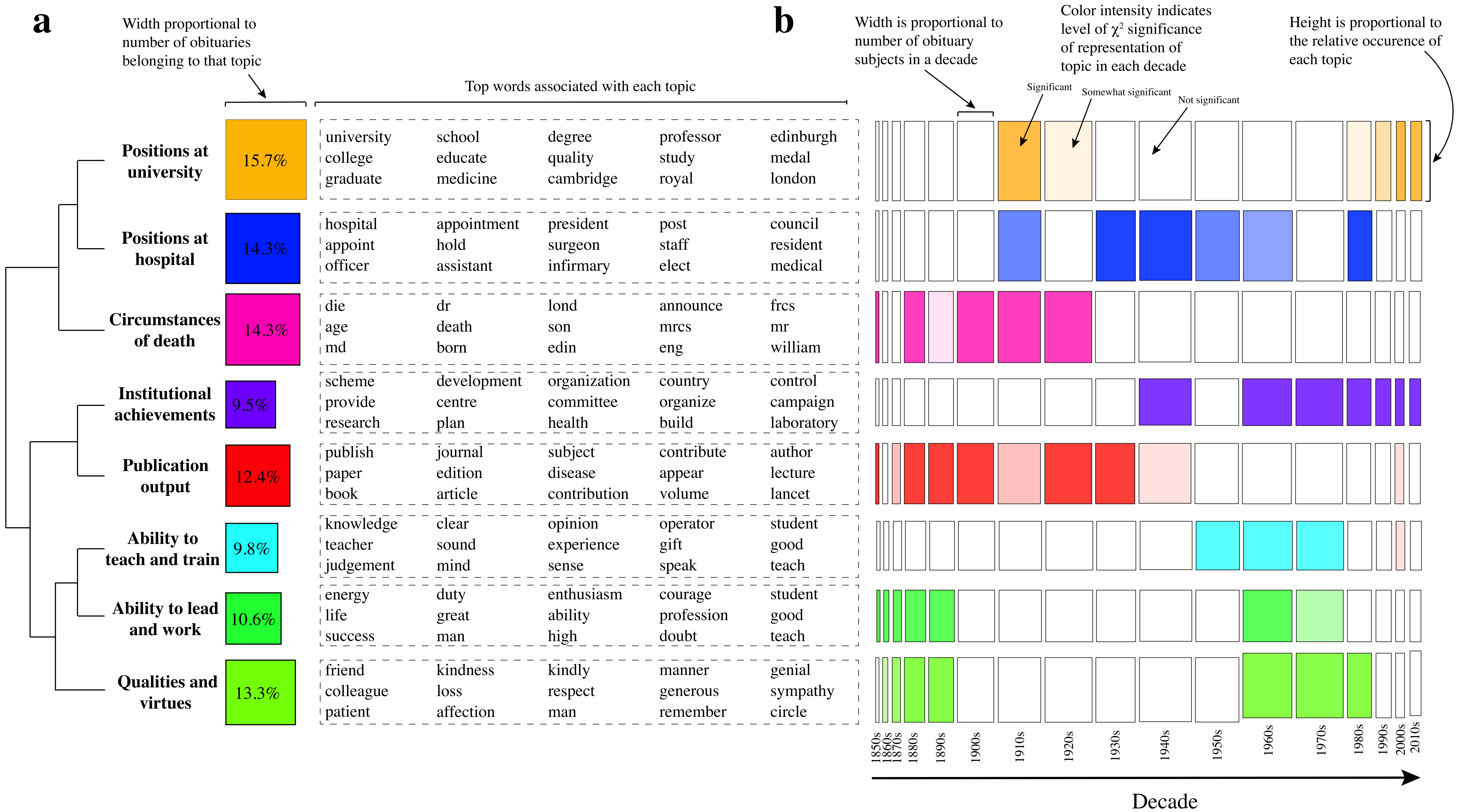


Specialties mentioned in text evolve over time



dakotamurray.me/talks/2020-northeastern/

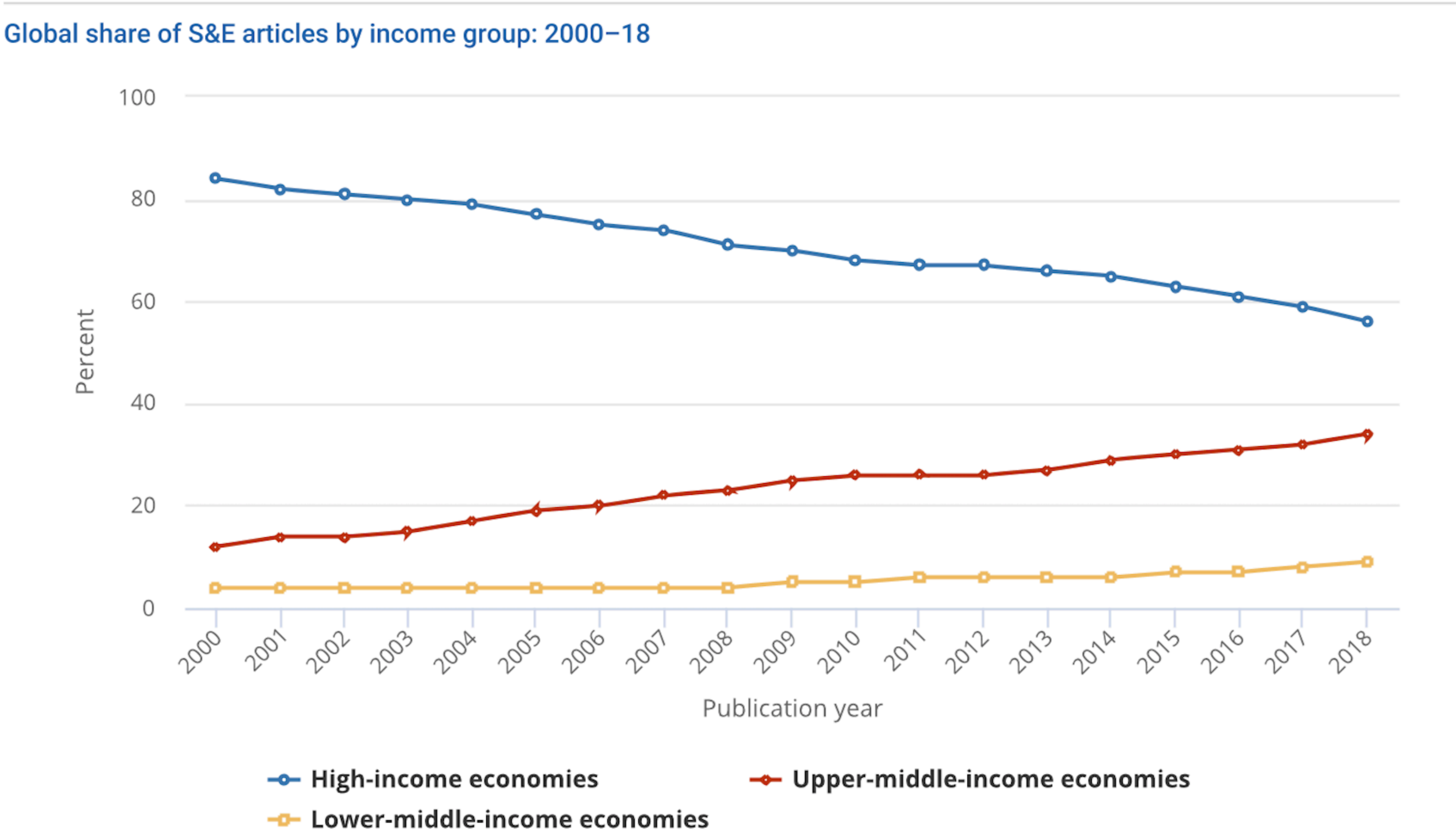
How we speak of the dead



Appendix—misc

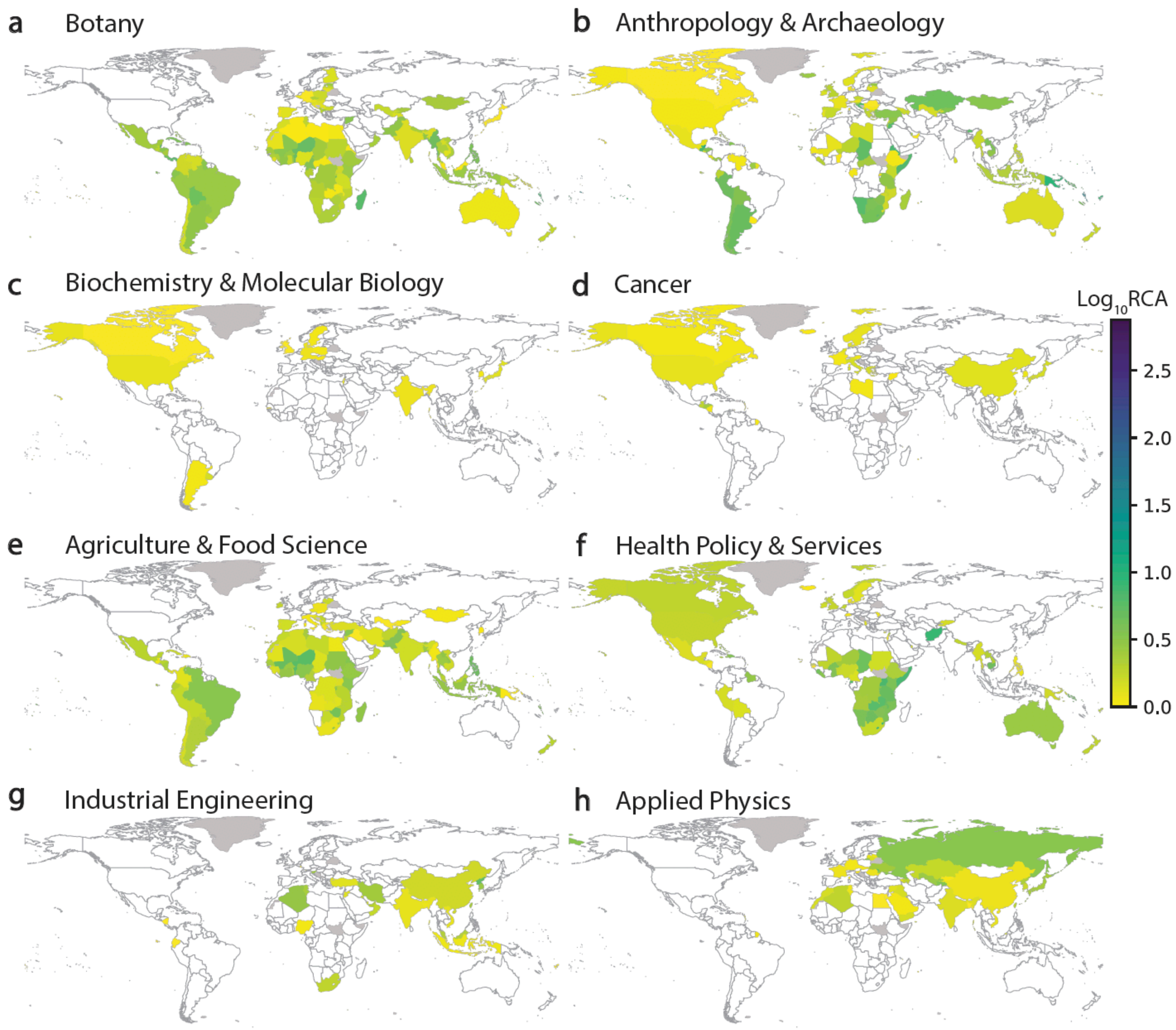
Similarly, certain countries dominate scientific output

NSF Statistics, Publication Output, by Region, Country, or Economy, https://ncses.nsf.gov/pubs/nsb20206



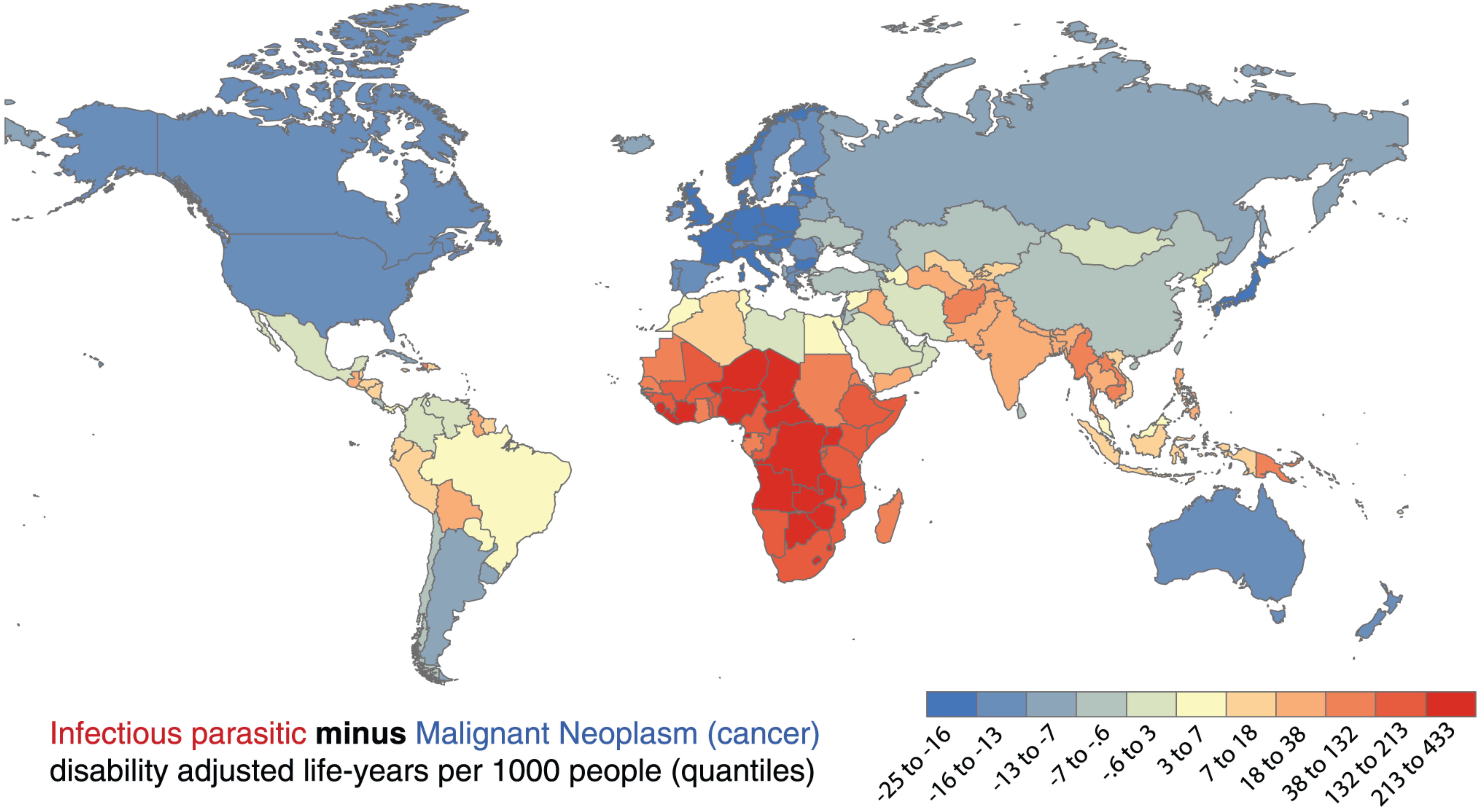
Countries do different kinds of research

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.



Infectious parasitic **minus** Malignant Neoplasm (cancer)
disability adjusted life-years per 1000 people (quantiles)