Toward a theory of scientific discovery and reproducibility

Bert Baumgaertner[1], Erkan Buzbas[1], Berna Devezer[1], Gustavo Nardin[2],

May 23, 2019, Science of Team Science Conference (Lansing)

[1]University of Idaho

[2]Brandenburg Institute of Technology

Staring summary

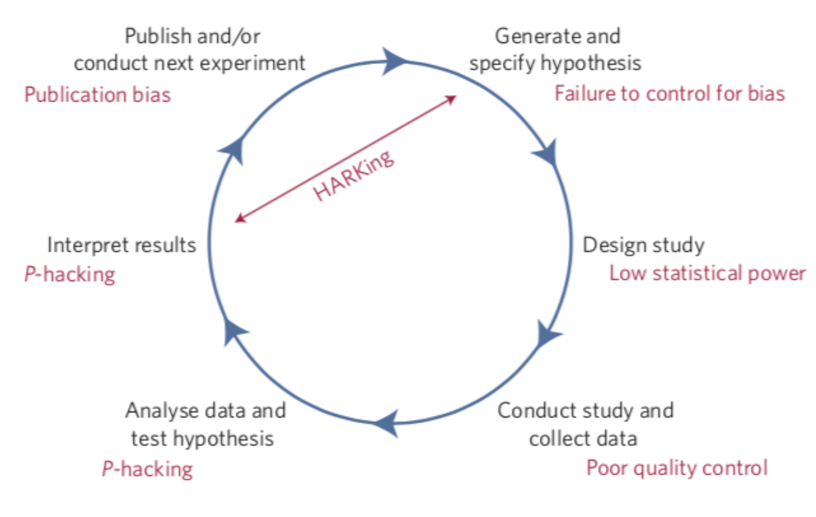
- (Cruising over a long history)
- Scientists have mostly relied on their intuition to understand the role of replications and reproducibility.
- No formal theory of reproducibility to explain how it is related to open science or truth.

Where are we now?

A “new” crisis. . .

- Diederik Stapel: 58 retractions since 2011; social psychology
- Yoshitaka Fujii: 183 retractions since 2012; anesthesiology
- Bayer studies (2011): out of 67 projects (oncology, womens health and cardiovascular medicine), only 14 could reproduce past work
- Amgen studies (2012): Team of 100 scientists; 47/53 cancer studies not replicated
- Psych reproducibility project (OSF, 2015): 61 out of 100 results failed to replicate
- Ongoing large-scale replication projects by OSF: Social psych, cancer biology, developmental psych, primate research, neuroscience,. . .

Putative causes



Recent theoretical research on reproducibility

- Focus on null hypothesis significance testing
- Effect of incentive structures and questionable research practices
- How to improve statistical and open science practices[†]

[†] Ioannidis 2005, McElreath and Smaldino 2015, Smaldino and McElreath 2016, Higginson and Munafò 2016, Nissen et al. 2016

Our questions

- How does reproducibility work at the baseline and how is it linked to the (assumed) truth and other properties of scientific discovery?
- Some fields progress by building, comparing, selecting, and re-building models. How can we represent scientific progress within a model-centric framework?

Theses

1. Even if erroneous methodological practices were absent, reproducible results are still not guaranteed.
2. Reproducibility and scientific progress are not always perfectly aligned.

Our approach to modeling scientific process

A model-centric view of scientific discovery

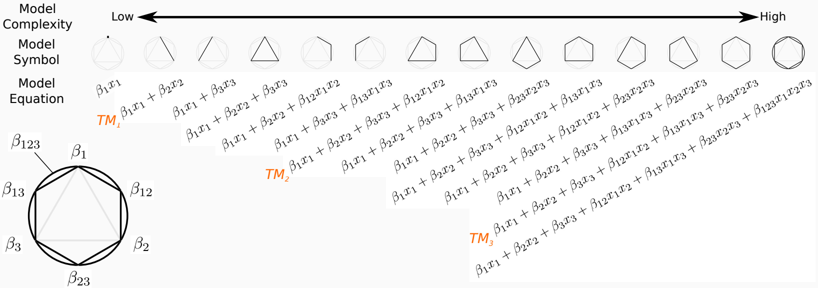
Science progresses by building, comparing, selecting, and re-building models.

We build a model-centric framework to study the properties of scientific discovery.

Scientific discovery is a temporal process.

We begin by imagining a model universe.

A universe of linear models



A stochastic process of scientific discovery

- Infinite population of scientists conducting a sequence of idealized experiments $\xi^{(t)} := (K^{(t)}, M^{(t)}, S, D^{(t)})$, indexed by time $t = 1, 2, \cdots$.
- $M^{(t)} \in \{M_1, M_2, \cdots, M_L\}$ known to all scientists.
- There are A distinct types of scientists, each with a well-defined research strategy for proposing a model in their experiment.
- These strategies depend on a global model $M_G^{(t)}$, which represents the consensus of the scientist population at time t, and the type of scientist.

Scientist types and their research strategies

Rey, the replicator, replicates the preceding experiment $\xi^{(t-1)}$

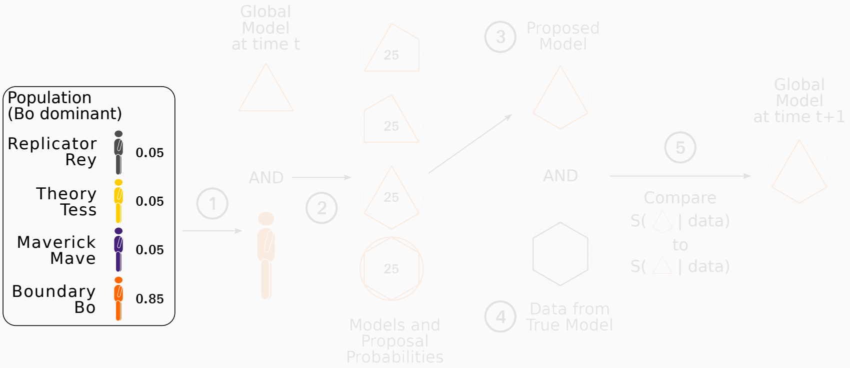
Tess, the theory tester, proposes a new model that is one step away from a global model $M_G^{(t)}$

Mave, the maverick, proposes a model uniformly randomly

Bo, the boundary tester, adds an interaction term to a global model $M_G^{(t)}$

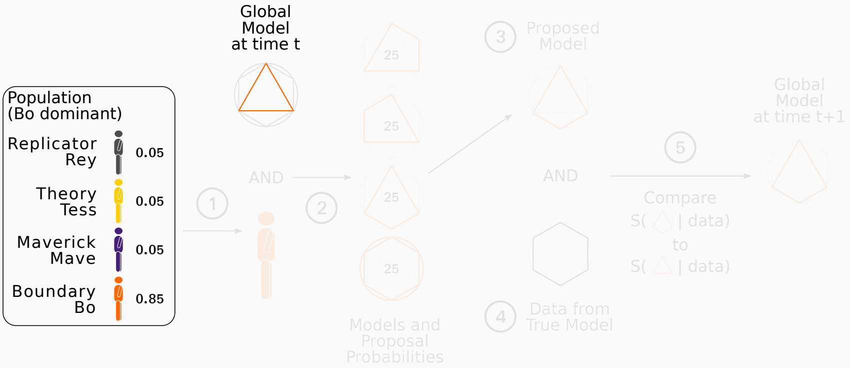
Example step from the process

using the idea of idealized experiment (K, M, S, D) as basis of scientific inquiry



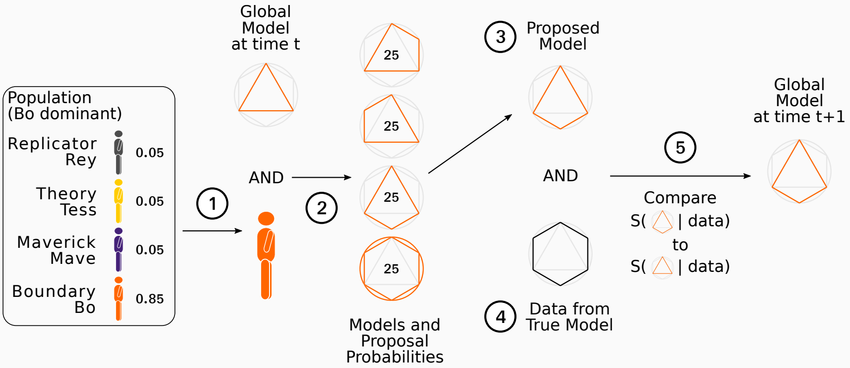
Example step from the process

using the idea of idealized experiment (K, M, S, D) as basis of scientific inquiry



Example step from the process

using the idea of idealized experiment (K, M, S, D) as basis of scientific inquiry



We track following properties of scientific discovery...

...in addition to reproducibility:

- Early discovery of truth
- Persistence on truth once it is discovered
- Time spent on truth in the long-term

We track following properties of scientific discovery...

...in addition to reproducibility:

Early discovery of truth

Persistence on truth once it is discovered

Time spent on truth in the long-term

in a stochastic process, when truth is a state, these can be studied by

We track following properties of scientific discovery...

...in addition to reproducibility:

Early discovery of truth –**mean first passage time**

Persistence on truth once it is discovered –**stickiness**

Time spent on truth in the long-term –**stationary distribution**

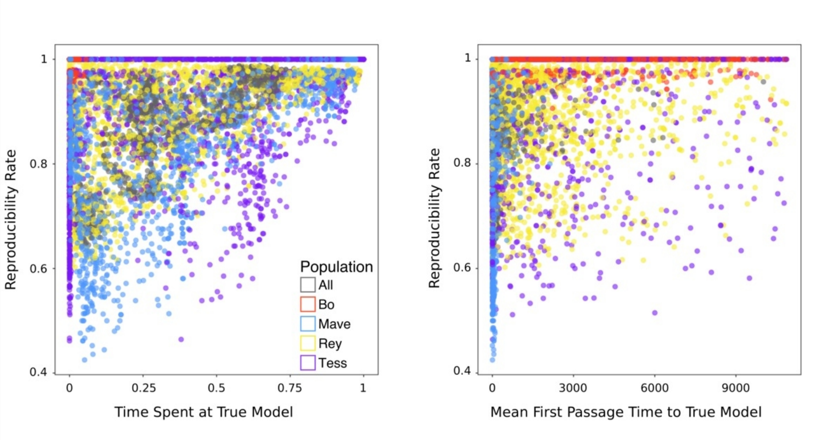
in a stochastic process, when truth is a state, these can be studied by

Results and conclusions

Results without replicators: Stickiness under low noise

How sticky a true model is varies with scientific population, statistical methods, and noise-to-signal ratio. Same is true for speed of discovery and time spent on true model in the long run.

Results with replicators: Reproducibility vs. scientific discovery



- All desirable properties vary with scientific population, M_T, and S.
- Reproducibility does not necessarily imply discovery of truth
- Innovative research speeds up scientific discovery

Results with replicators: Role of epistemic diversity

	Median					IQR				
	Rey	**Tess**	**Mave**	**Bo**	**All**	**Rey**	**Tess**	**Mave**	**Bo**	**All**
Time Spent at True Model	**23.9**	**30.6**	**40.1**	**0.3**	**26.0**	**41.9**	**60.8**	**66.5**	**0.7**	**43.7**
Stickiness	**90.8**	**97.7**	**93.0**	**90.0**	**91.9**	**17.2**	**16.0**	**17.0**	**16.2**	**14.9**
Mean First Passage Time to True Model	**820.5**	**403.5**	**16.0**	**1592.5**	**29.0**	**2137.0**	**2267.0**	**36.0**	**3384.5**	**73.0**
Reproducibility Rate	**89.9**	**93.1**	**90.3**	**100.0**	**90.2**	**12.9**	**20.0**	**17.6**	**<0.01**	**12.3**
Reproducibility Rate if True Model is Global Model	**90.8**	**100.0**	**97.5**	**100.0**	**91.7**	**17.2**	**8.3**	**11.1**	**<0.01**	**14.0**

Epistemic diversity optimizes the desirable properties of scientific discovery.

Concluding remarks

Reproducibility and truth have a complex relationship

Reproducibility and other desirable properties of scientific discovery have a complex relationship

Irreproducibility cannot be reduced down to methodological practices

Resources

Contacts:

- Bert Baumgaertner, U. of Idaho, bbaum@uidaho.edu
- Berna Devezer, U. of Idaho, bdevezer@uidaho.edu

Published at PLOS ONE: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0216125

Preprints at

https://arxiv.org/abs/1803.10118

https://arxiv.org/abs/1811.04525

All scripts and data at

https://github.com/gnardin/CRUST

Post- Session Evaluation

Please take 2 minutes to complete this brief but valuable post session evaluation. *Responses are anonymous and will be used to improve future programming.*

Type https:bit.ly/scits2019 into your browser to launch survey

or ….

Scan this QR code

Thank you!