Statistical Rethinking

2023

20. Horoscopes

Parametric assumptions:
(1) Independent samples
(2) Data normally distributed
(3) Equal variances
Type of data?
Discrete, categorical
Any counts < 5?
Continuous
No
Yes
Type of question?
Chi-square tests, one and two sample
Fisher's exact test
Relationships
Differences
One-sample t-test
Do you have dependent & independent variables?
Differences between what?
Means
Variances
Fmax test or Bartlett's test
Yes
No
Multiple means Single variable
Regression analysis
Correlation analysis
How many groups?
Parametric assumptions satisfied?
Parametric
Nonparametric
More than two
No
Yes
No
Pearson's r
Spearman's rank correlation
Two
Transform

Stargazing

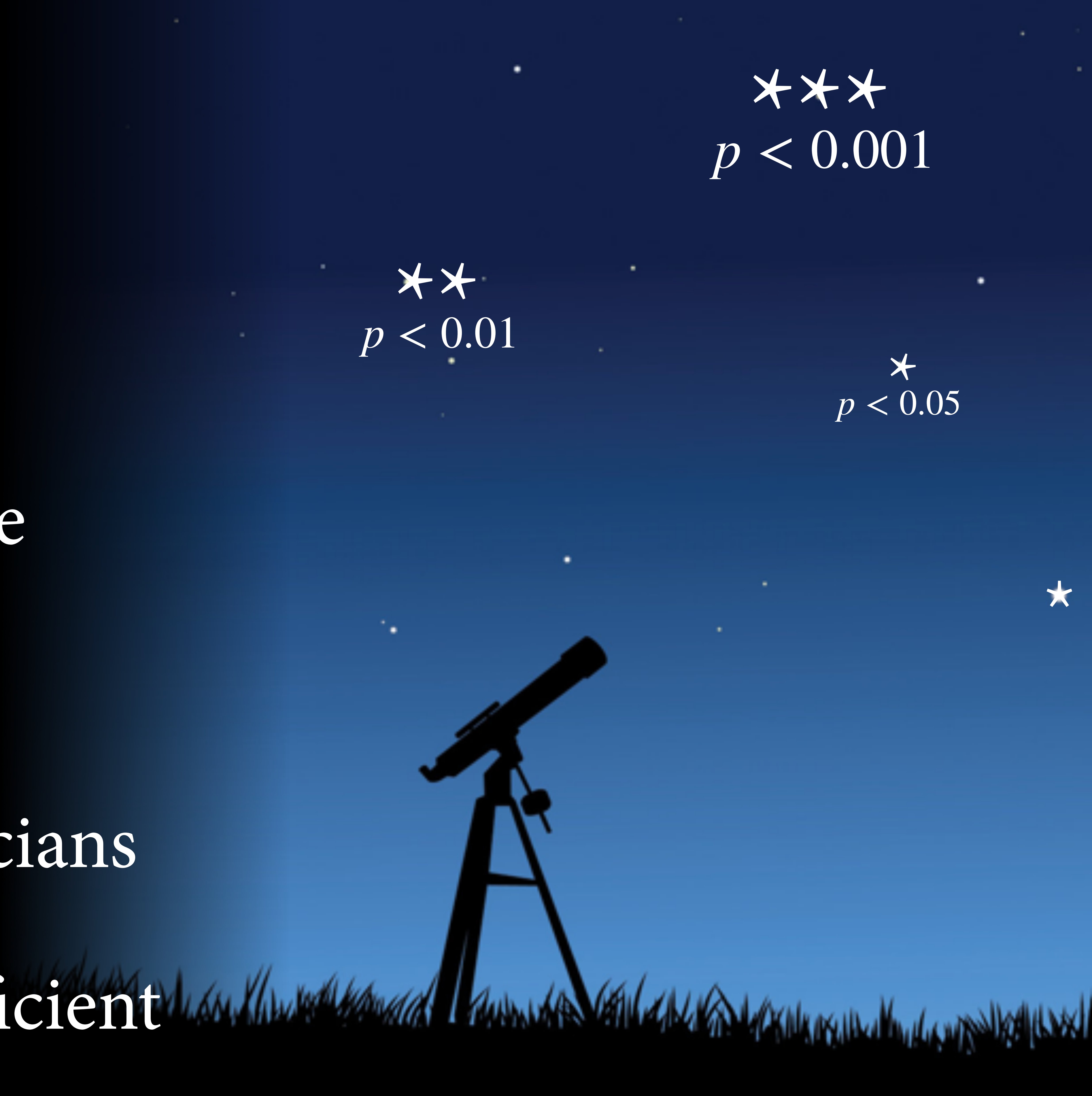
Fortune telling frameworks:

(1) From vague facts, vague advice

(2) Exaggerated importance

Applies to astrologers and statisticians

Valid vague advice exists, not sufficient

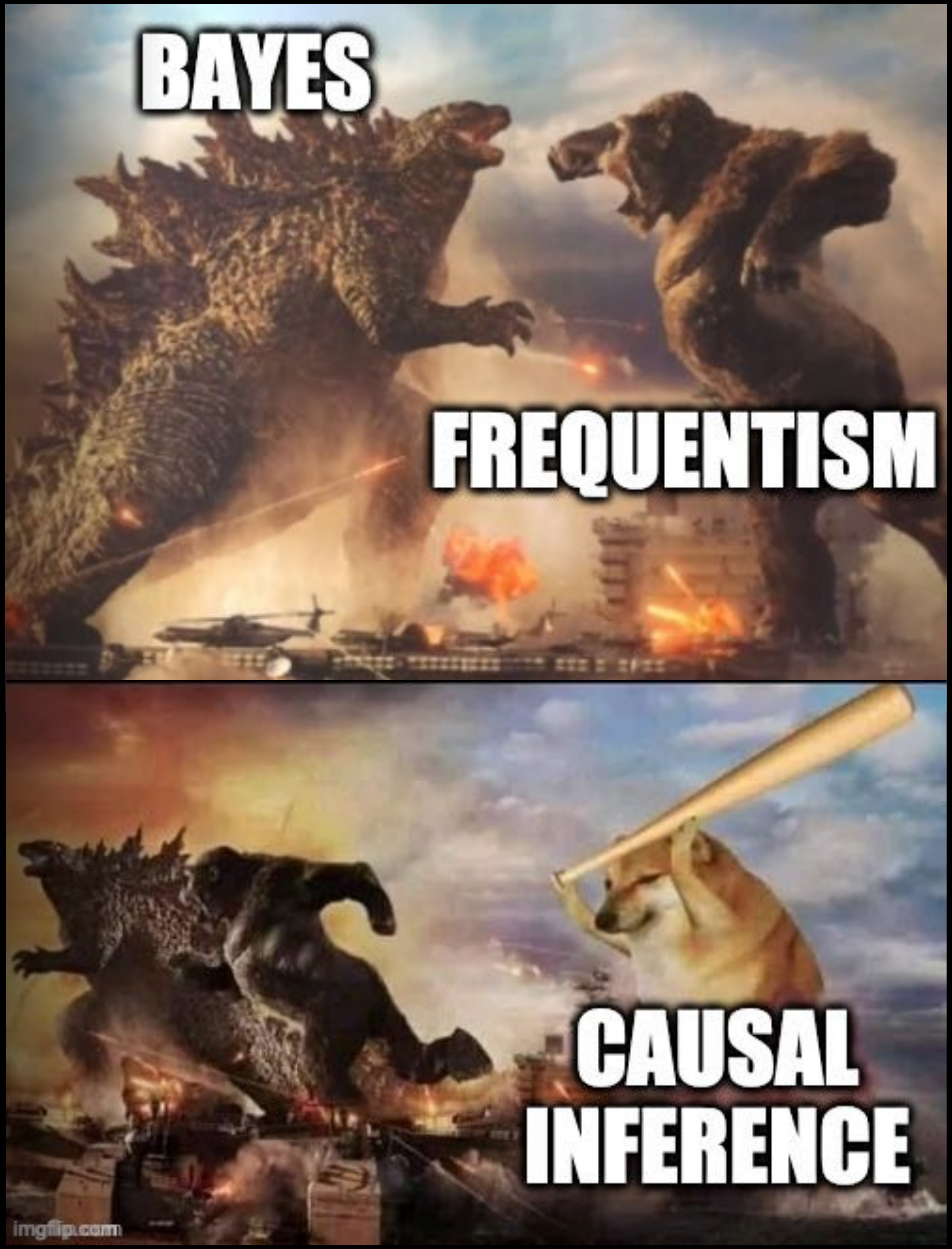


Stargazing

Statistical procedures acquire meaning from scientific models

Cannot offload **subjective** responsibility to an **objective** procedure

Many subjective responsibilities

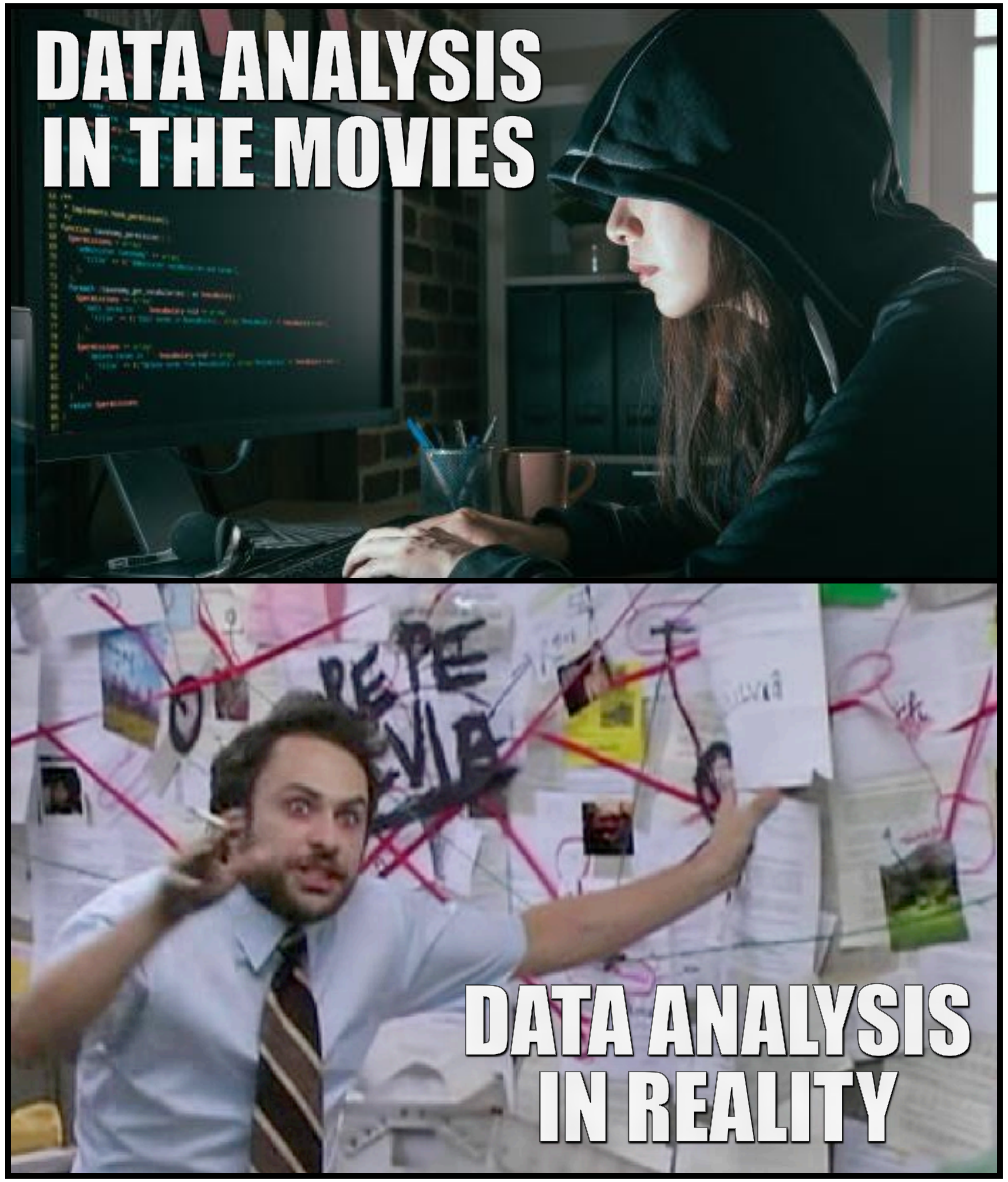


A Typical Scientific Laboratory

Planning

@StuartJRitchie JAKE-CLARK.TUMBLR

Working



Reporting

Planning

Goal setting

Theory building

Justified sampling plan

Justified analysis plan

Documentation

Open software & data formats

@StuartJRitchie

Planning

Goal setting – What for? Estimands

Theory building

Justified sampling plan

Justified analysis plan

Documentation

Open software & data formats

Planning

Goal setting – What for? Estimands

Theory building – Which assumptions?

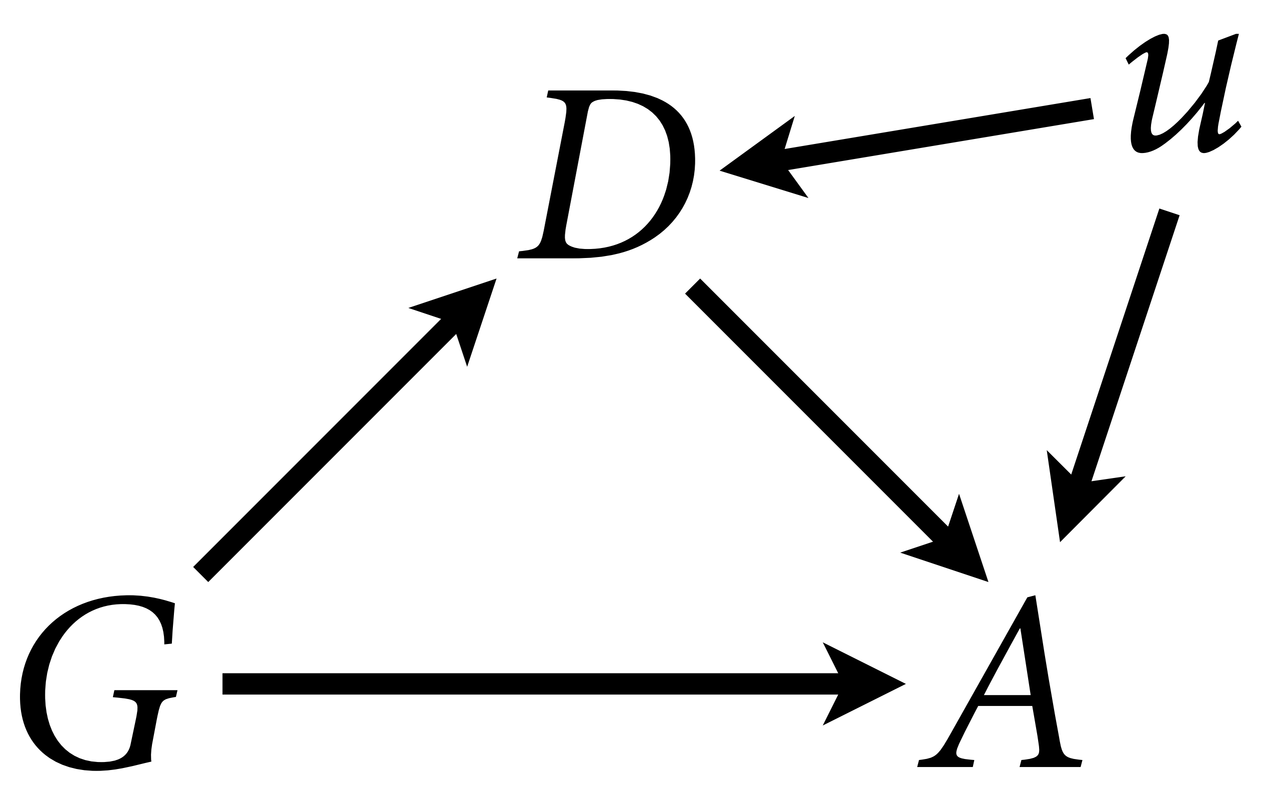
Justified sampling plan

Justified analysis plan

Documentation

Open software & data formats

Theory Building



Levels of theory building

(1) Heuristic causal models (DAGs)

(2) Structural causal models

(3) Dynamic models

(4) Agent-based models

$$\frac{dH}{dt} = H_t b_H - H_t(L_t m_H)$$

$$\frac{dL}{dt} = L_t(H_t b_L) - L_t m_L$$

Theory Building

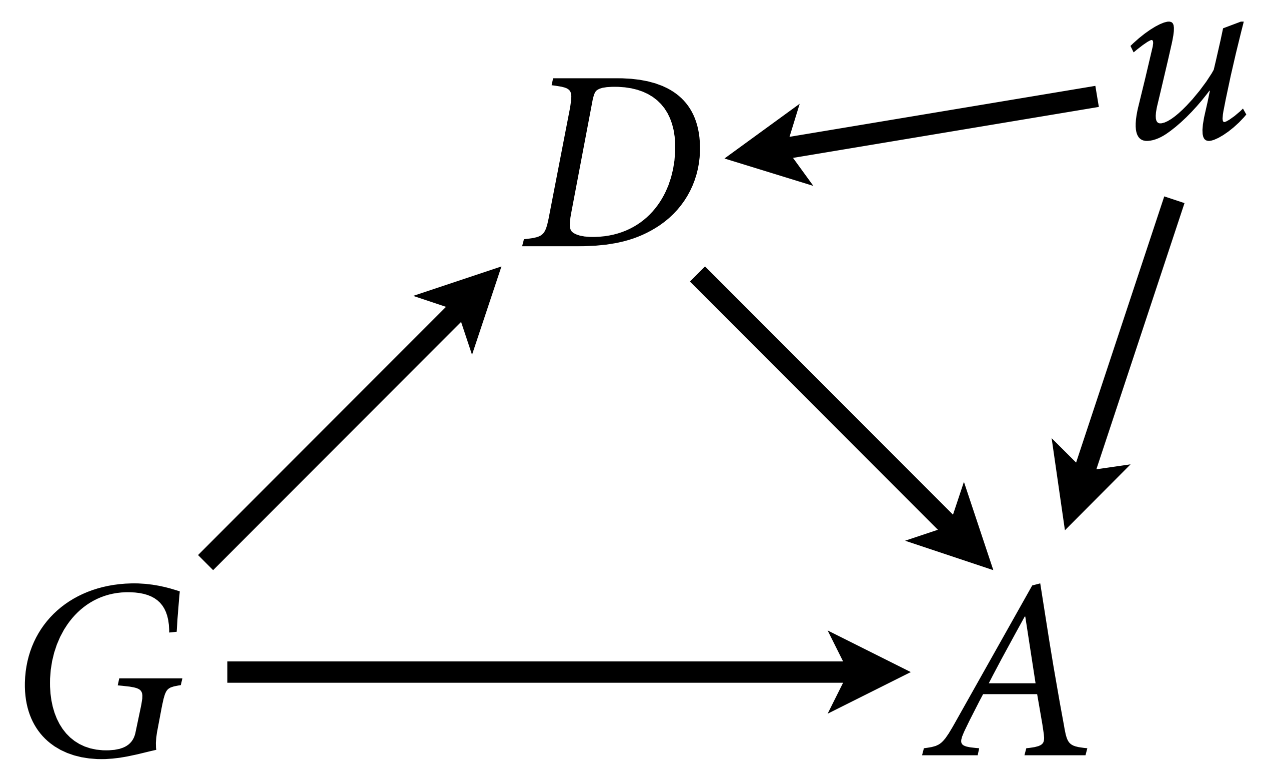
Heuristic causal models (DAGs)

(1) Treatment and outcome

(2) Other causes

(3) Other effects

(4) Unobserved causes



Theory Building

Heuristic causal models (DAGs)

(1) Treatment and outcome

(2) Other causes

(3) Other effects

(4) Unobserved causes

$$G \longrightarrow A$$

Theory Building

Heuristic causal models (DAGs)

(1) Treatment and outcome

(2) Other causes

(3) Other effects

(4) Unobserved causes

$$
\begin{array}{ccc}
 & & D \\
 & & \downarrow \\
G & \longrightarrow & A
\end{array}
$$

Theory Building

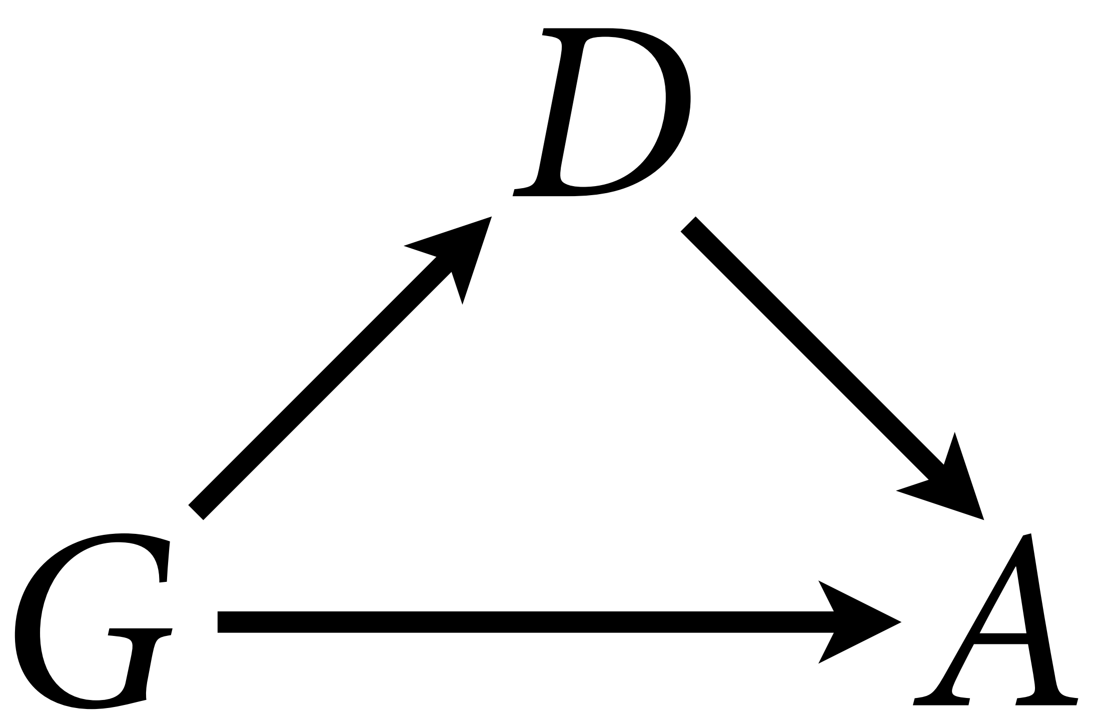
Heuristic causal models (DAGs)

(1) Treatment and outcome

(2) Other causes

(3) Other effects

(4) Unobserved causes



Theory Building

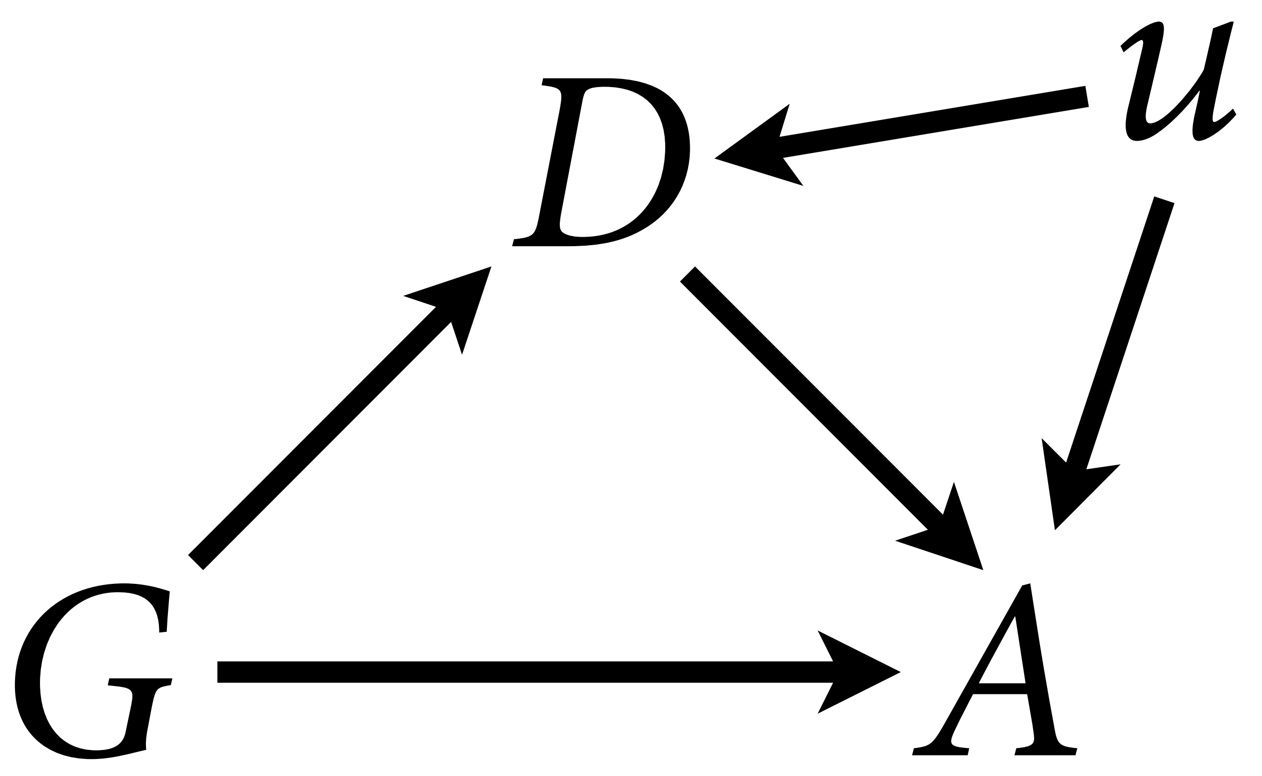
Heuristic causal models (DAGs)

(1) Treatment and outcome

(2) Other causes

(3) Other effects

(4) Unobserved causes



Planning

Goal setting – What for? Estimands

Theory building – Which assumptions?

Justified sampling plan – Which data?

Justified analysis plan

Documentation

Open software & data formats

Planning

Goal setting – What for? Estimands

Theory building – Which assumptions?

Justified sampling plan – Which data?

Justified analysis plan – Which golems?

Documentation

Open software & data formats

Planning

Goal setting – What for? Estimands

Theory building – Which assumptions?

Justified sampling plan – Which data?

Justified analysis plan – Which golems?

Documentation – How did it happen?

Open software & data formats

Planning

Goal setting – What for? Estimands

Theory building – Which assumptions?

Justified sampling plan – Which data?

Justified analysis plan – Which golems?

Documentation – How did it happen?

Open software & data formats

Pre-Registration

Pre-registration: Prior public documentation of research design and analysis plan

Goal: Make transparent which decisions are sample-dependent

Does little to improve data analysis

Lots of pre-registered causal salad

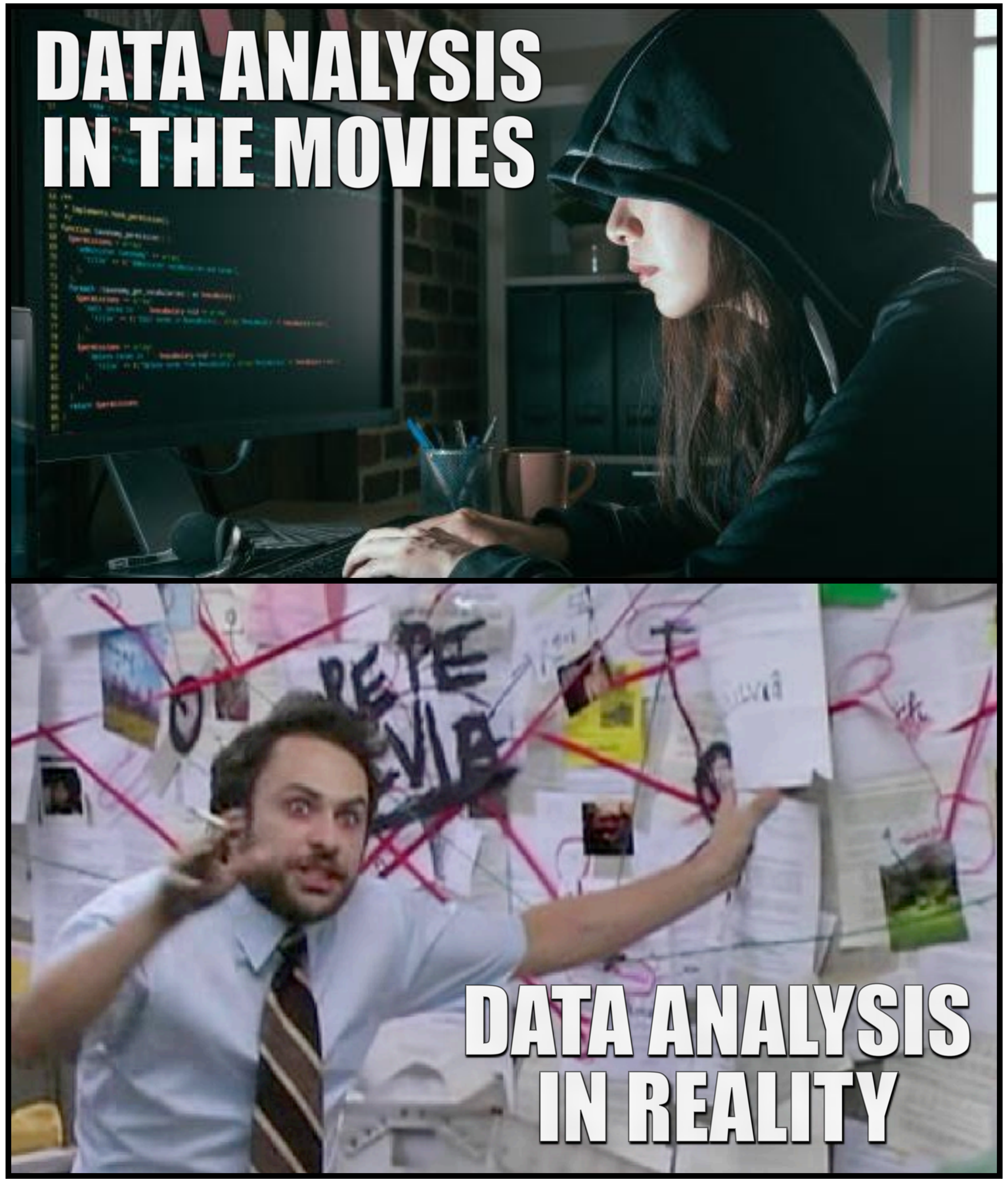
Working

Control

Incremental testing

Documentation

Review



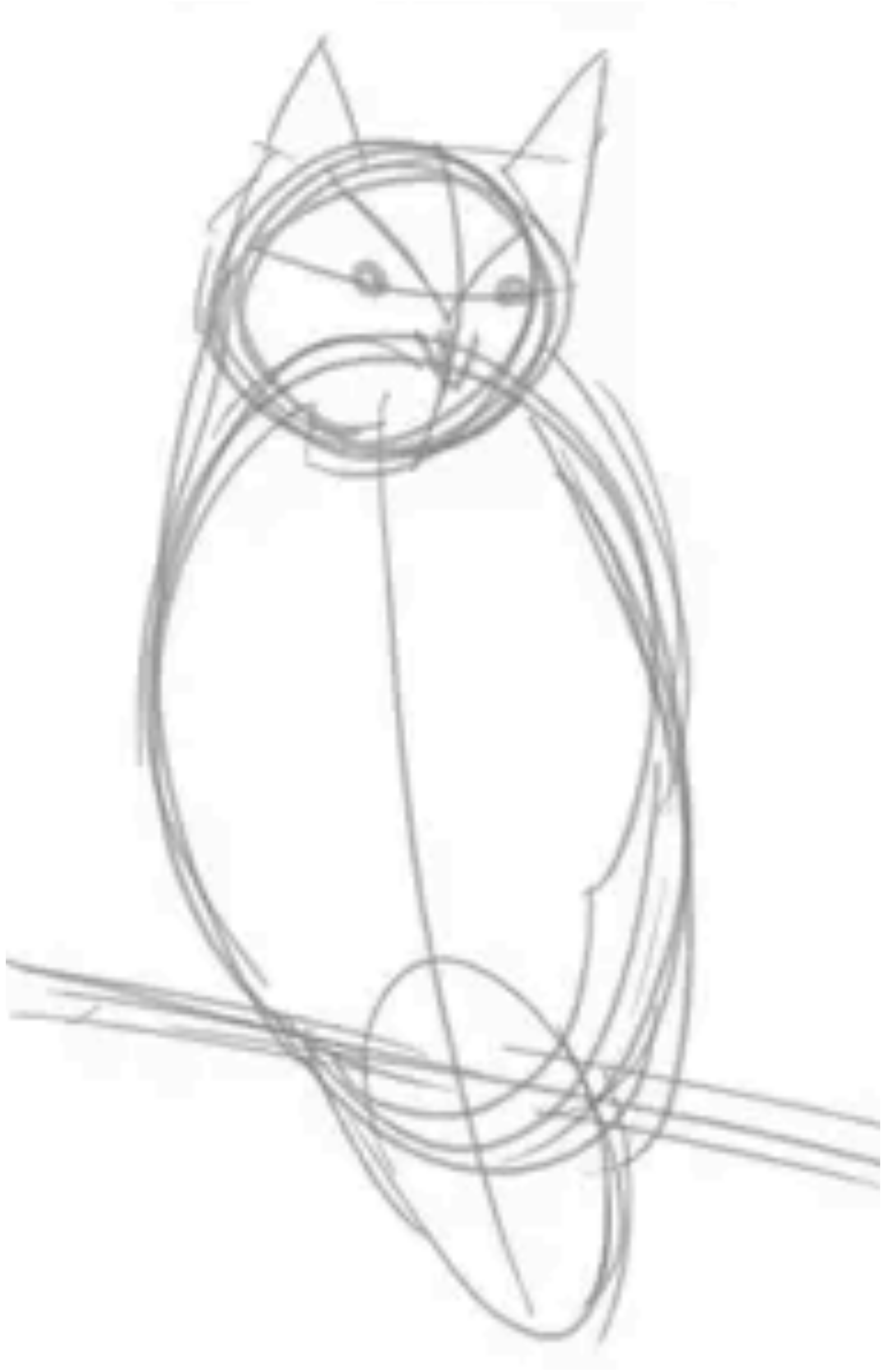
1 Express theory as probabilistic program

2 Prove planned analysis could work (conditionally)

3 Test pipeline on synthetic data

4 Run pipeline on empirical data

entire history open



Professional Norms

Dangerous lack of professional norms in scientific computing

Often impossible to figure out what was done

Often impossible to know if code works as intended

Like pipetting by mouth

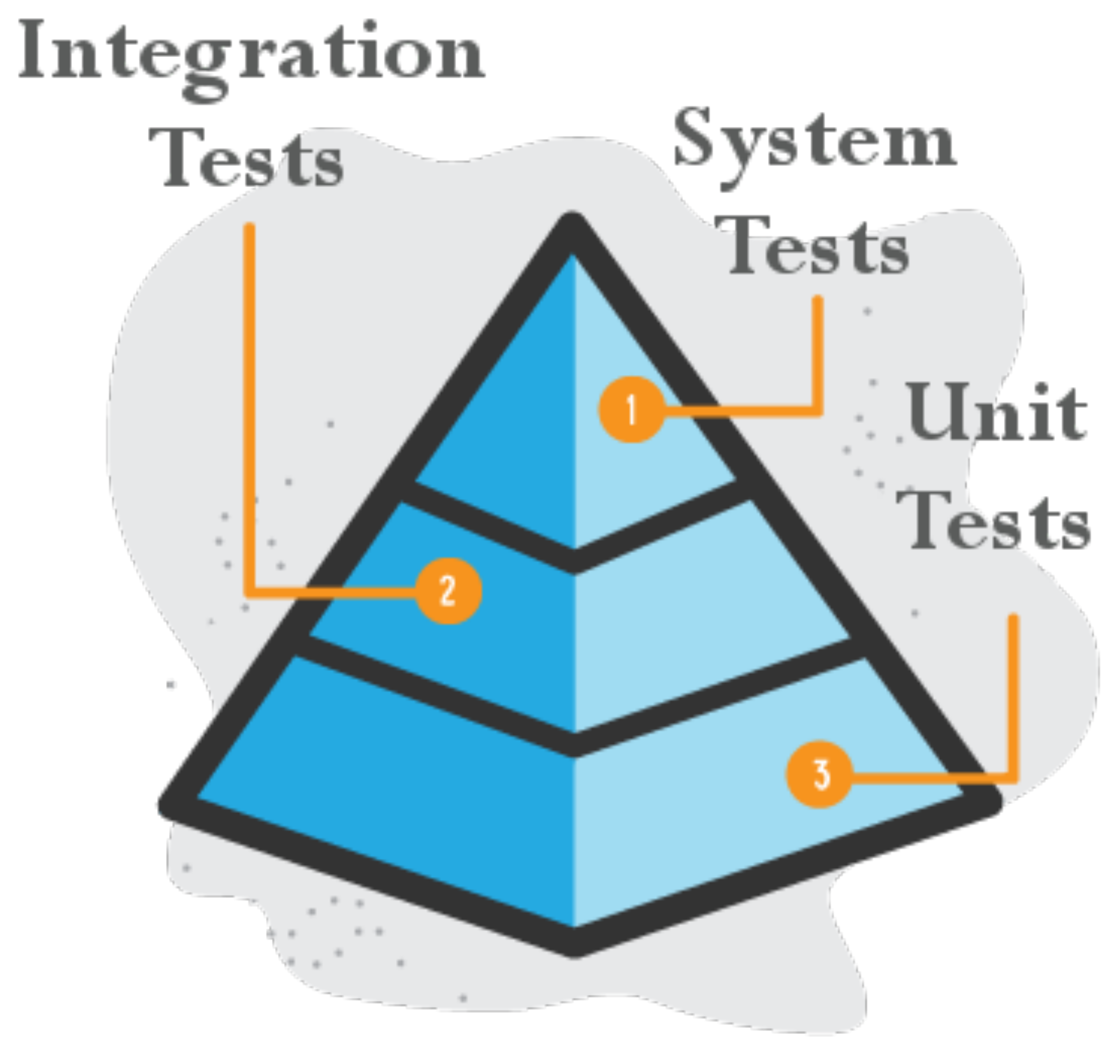
Research Engineering

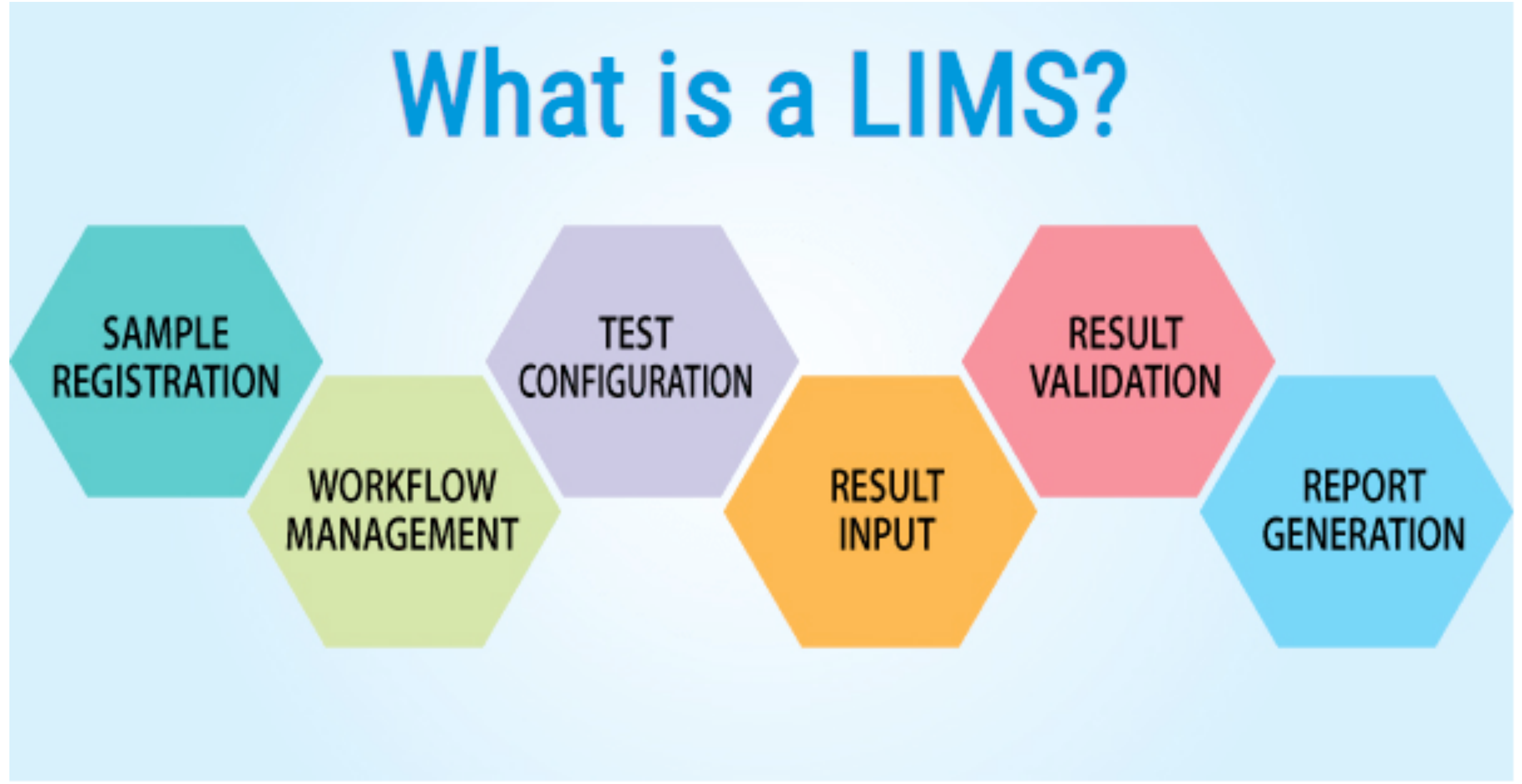
Control: Versioning, back-up, accountability

Incremental testing: Piece by piece

Documentation: Comment everything

Review: 4 eyes on code and materials





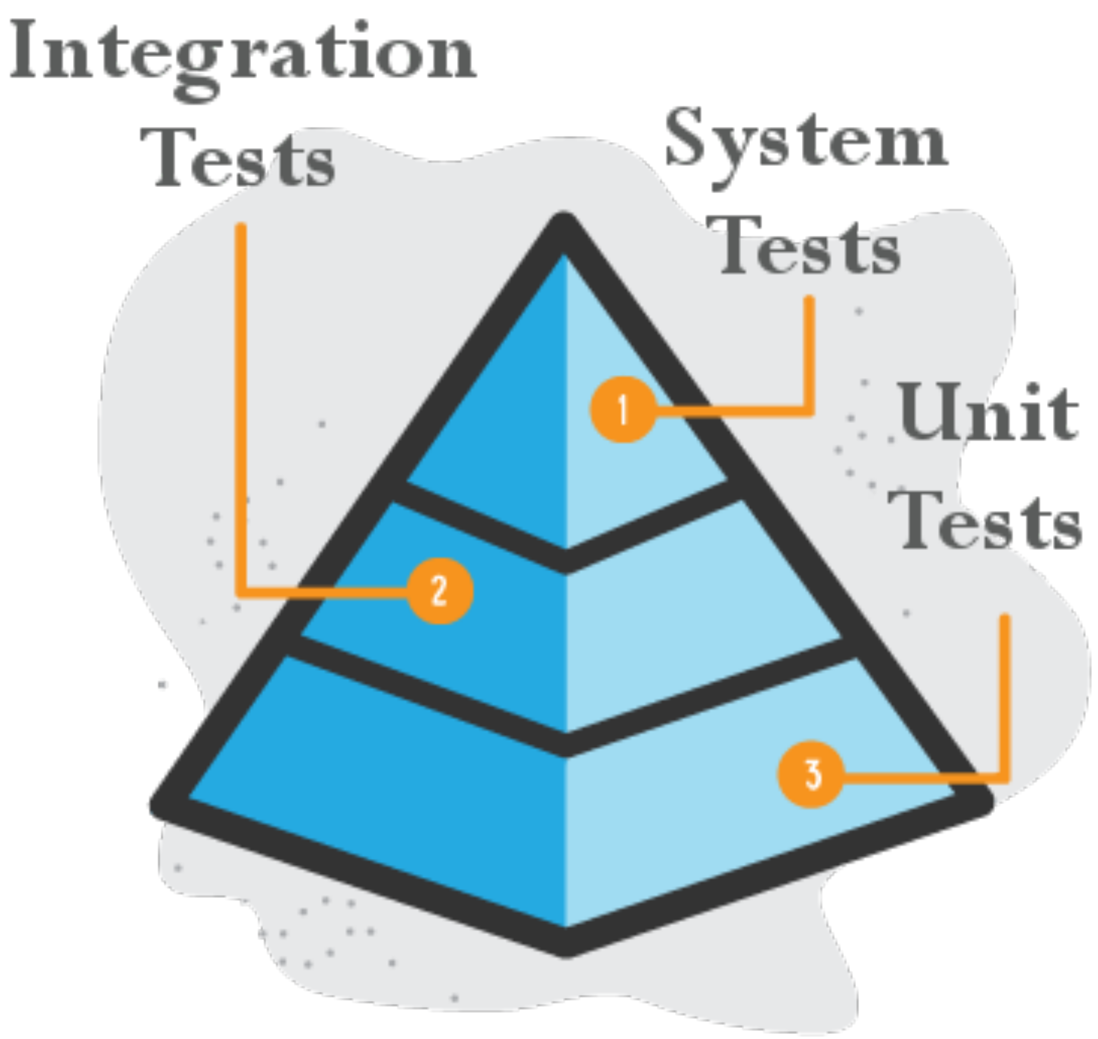
Research Engineering

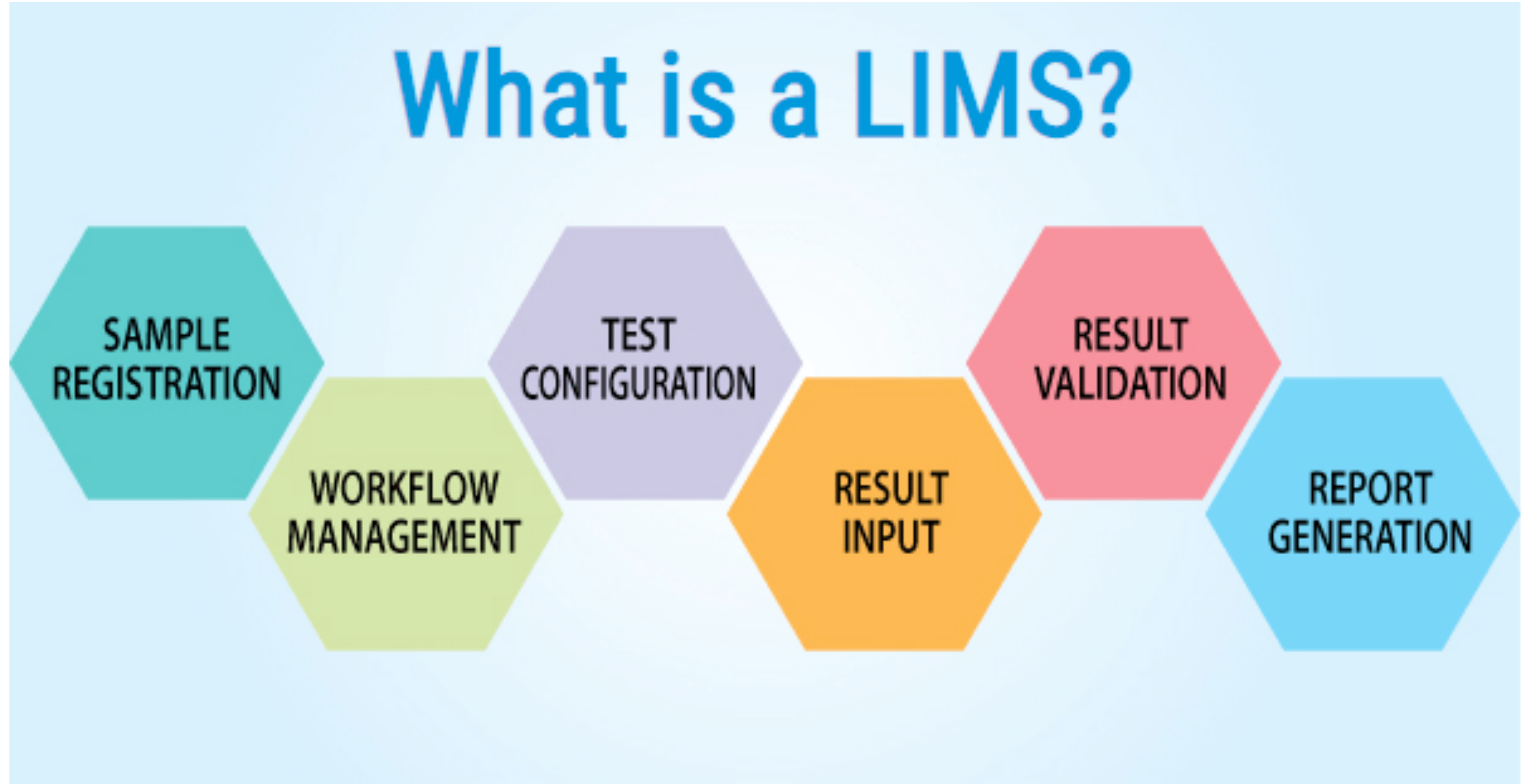
Control: Versioning, back-up, accountability

Incremental testing: Piece by piece

Documentation: Comment everything

Review: 4 eyes on code and materials

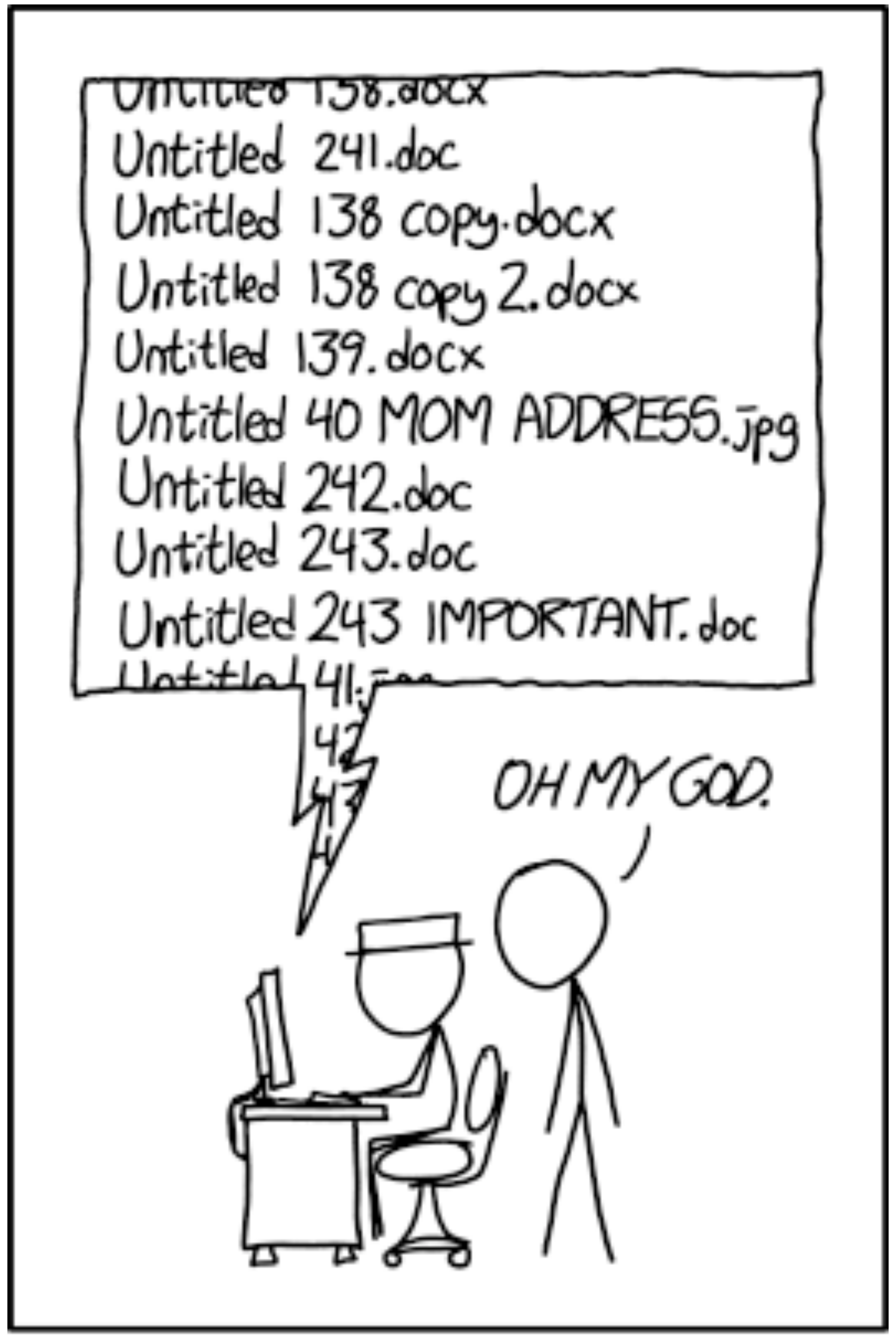




Versioning and Testing

Version control: Database of changes to project files, managed history

Testing: Incremental milestones, test each before moving to next



PROTIP: NEVER LOOK IN SOMEONE ELSE'S DOCUMENTS FOLDER.

rmcelreath / stat_rethinking_2022 Public

Unpin | Unwatch 189 | Fork 212 | Star 2.8k

Code | Issues 4 | Pull requests 1 | Actions | Projects | Wiki | Security | Insights | ...

main | 1 branch | 0 tags | Go to file | Add file | Code

rmcelreath week 9 solutions 49baa56 3 days ago 64 commits

homework	week 9 solutions	3 days ago
scripts_animation	lecture 10 script	7 days ago
README.md	lecture 19 links	7 days ago
additional_reading.md	poststrat link	22 days ago
title.gif	README update	4 months ago

README.md

About

Statistical Rethinking course winter 2022

Readme
2.8k stars
189 watching
212 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

R 97.8% Stan 2.2%

2 README.md

```
@@ -40,7 +40,7 @@ Lecture playlist on Youtube: <[Statistical Rethinking 2022](https://www.youtube.
40 40   | Week 07 | 18 February | Chapters 13 and 14 | [13] <[Multi-Multilevel Models](https://www.youtube.com/watch?v=n2aJYtuGu54&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=13)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-13)> <br> [14] <[Correlated varying effects](https://www.youtube.com/watch?v=XDoAglqd7ss&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=14)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-14)>
41 41   | Week 08 | 25 February | Chapter 14 | [15] <[Social Networks](https://www.youtube.com/watch?v=L_QumFUv7C8&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=15)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-15)> <br> [16] <[Gaussian Processes](https://www.youtube.com/watch?v=PIuqxOBJqLU&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=16)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-16)>
42 42   | Week 09 | 04 March | Chapter 15 | [17] <[Measurement Error](https://www.youtube.com/watch?v=lTFAB6QmwHM&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=17)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-17)> <br> [18] <[Missing Data](https://www.youtube.com/watch?v=oMiSb8GKR0o&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=18)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-18)>
43    - | Week 10 | 11 March | Chapters 16 and 17 | [19] Beyond GLMs: State-space Models, ODEs <br> [20] Horoscopes
   43 + | Week 10 | 11 March | Chapters 16 and 17 | [19] <[Beyond GLMs](https://www.youtube.com/watch?v=Doaod09YitA&list=PLDcUM9US4XdMROZ57-OIRtIK0aOynbgZN&index=19)> <[(Slides)](https://speakerdeck.com/rmcelreath/statistical-rethinking-2022-lecture-19)> <br> [20] Horoscopes
44 44
45 45
46 46   # Coding
```

Versioning and Testing

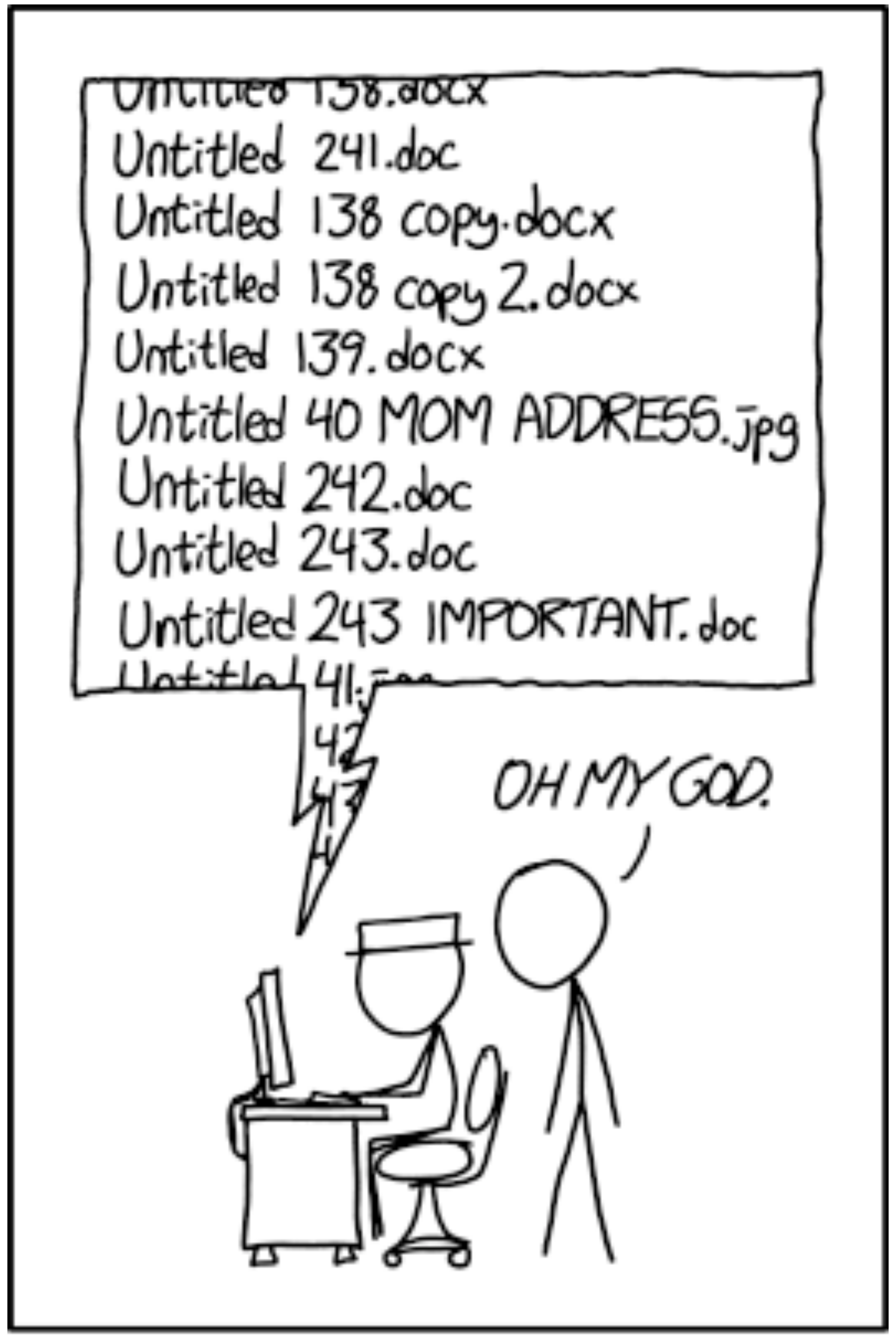
Most researchers don't need all git's features

But do:

Commit changes after each milestone
Maintain test code in project

Do not:

Replace raw data with processed data



PROTIP: NEVER LOOK IN SOMEONE ELSE'S DOCUMENTS FOLDER.

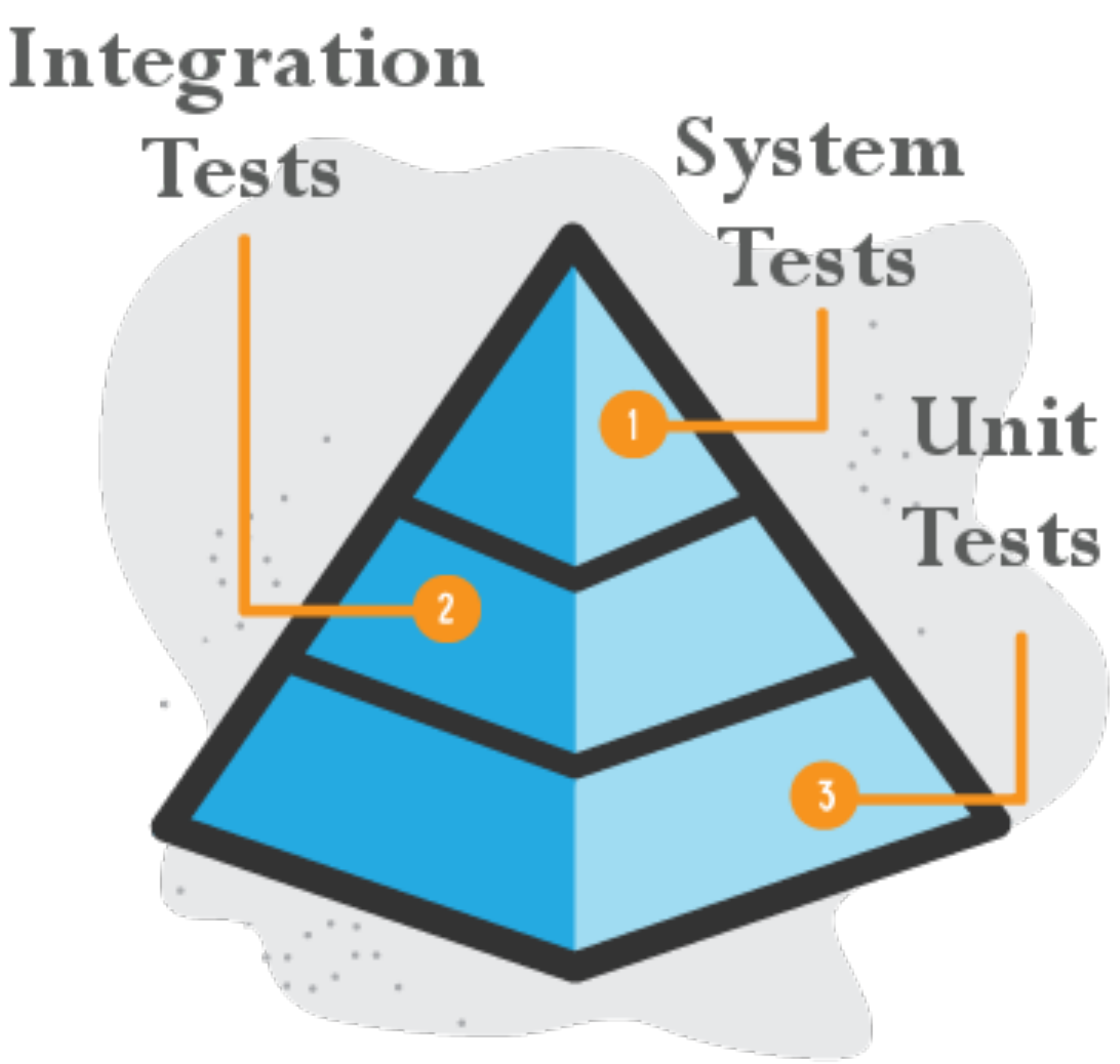
More on Testing

Complex analyses must be built in steps

Test each step

Social networks lecture (#15) as example

Milestones:
(1) Synthetic data simulation
(2) Dyadic reciprocity model
(3) Add generalized giving/receiving
(4) Add wealth, association index



math

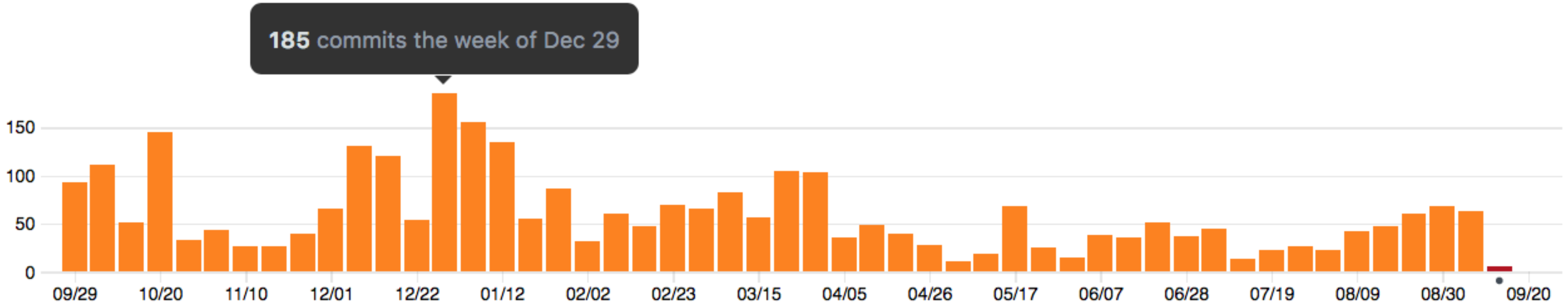
The Stan Math Library is a C++ template library for automatic differentiation of any order using forward, reverse, and mixed modes. It includes a range of built-in functions for probabilistic modeling, linear algebra, and equation solving.

math · automatic-differentiation · stan · stan-math-library

C++ · BSD-3-Clause · 127 · 449 · 214 (20 issues need help) · 23

Updated 5 minutes ago

5.1 MB of library code   8.2 MB of test code



https://github.com/stan-dev/math

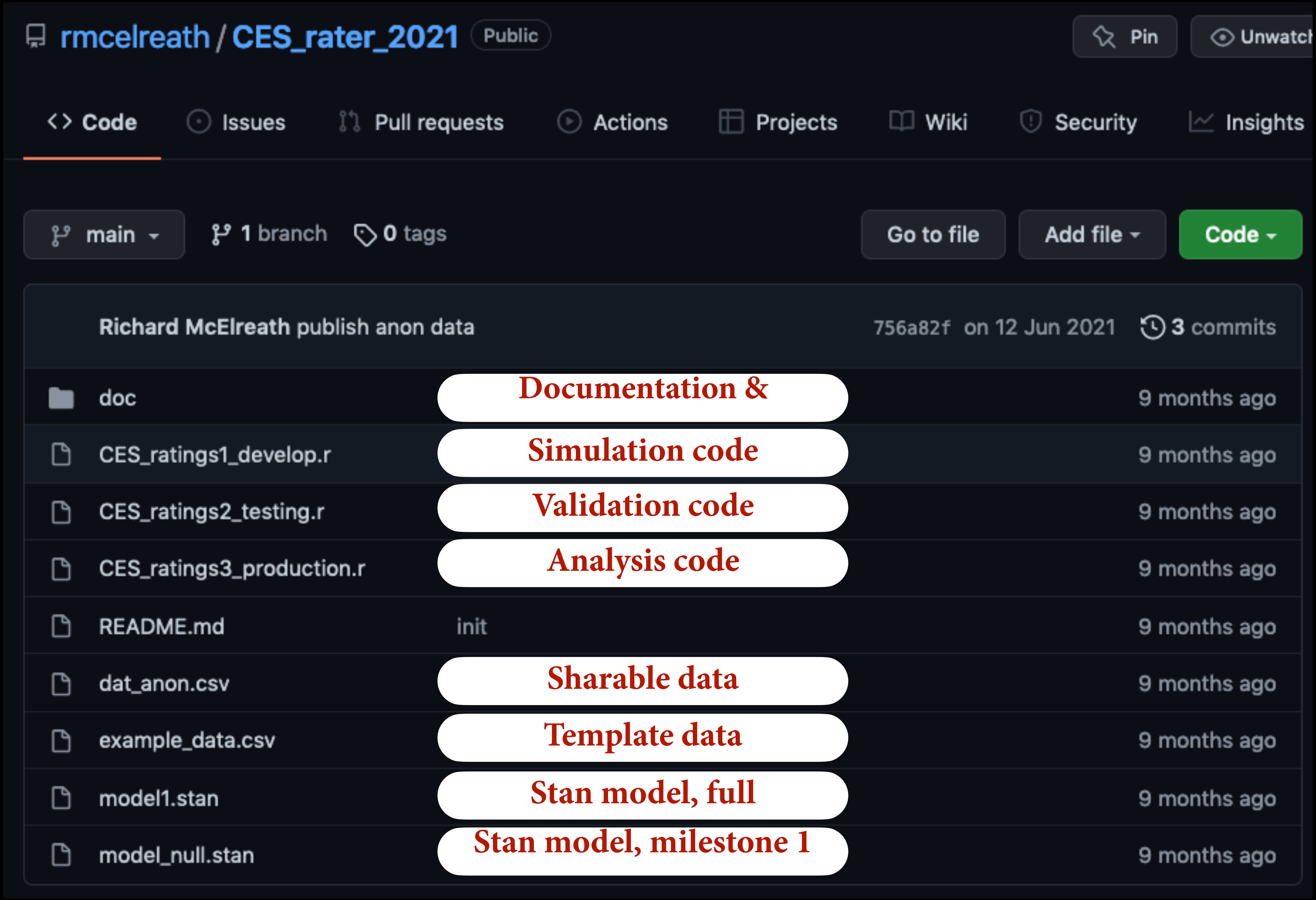
rmcelreath / CES_rater_2021 Public

Pin Unwatch

Code Issues Pull requests Actions Projects Wiki Security Insights

main 1 branch 0 tags Go to file Add file Code

Richard McElreath publish anon data		756a82f on 12 Jun 2021 3 commits
doc	init	9 months ago
CES_ratings1_develop.r	report stage	9 months ago
CES_ratings2_testing.r	report stage	9 months ago
CES_ratings3_production.r	publish anon data	9 months ago
README.md	init	9 months ago
dat_anon.csv	publish anon data	9 months ago
example_data.csv	init	9 months ago
model1.stan	init	9 months ago
model_null.stan	init	9 months ago


rmcelreath / CES_rater_2021
Public
Pin
Unwatch
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
main
1 branch
0 tags
Go to file
Add file
Code
Richard McElreath publish anon data
756a82f on 12 Jun 2021
3 commits
doc
Documentation &
9 months ago
CES_ratings1_develop.r
Simulation code
9 months ago
CES_ratings2_testing.r
Validation code
9 months ago
CES_ratings3_production.r
Analysis code
9 months ago
README.md
init
9 months ago
dat_anon.csv
Sharable data
9 months ago
example_data.csv
Template data
9 months ago
model1.stan
Stan model, full
9 months ago
model_null.stan
Stan model, milestone 1
9 months ago

DATA CARPENTRY

BUILDING COMMUNITIES TEACHING UNIVERSAL DATA LITERACY

What is Data Carpentry?

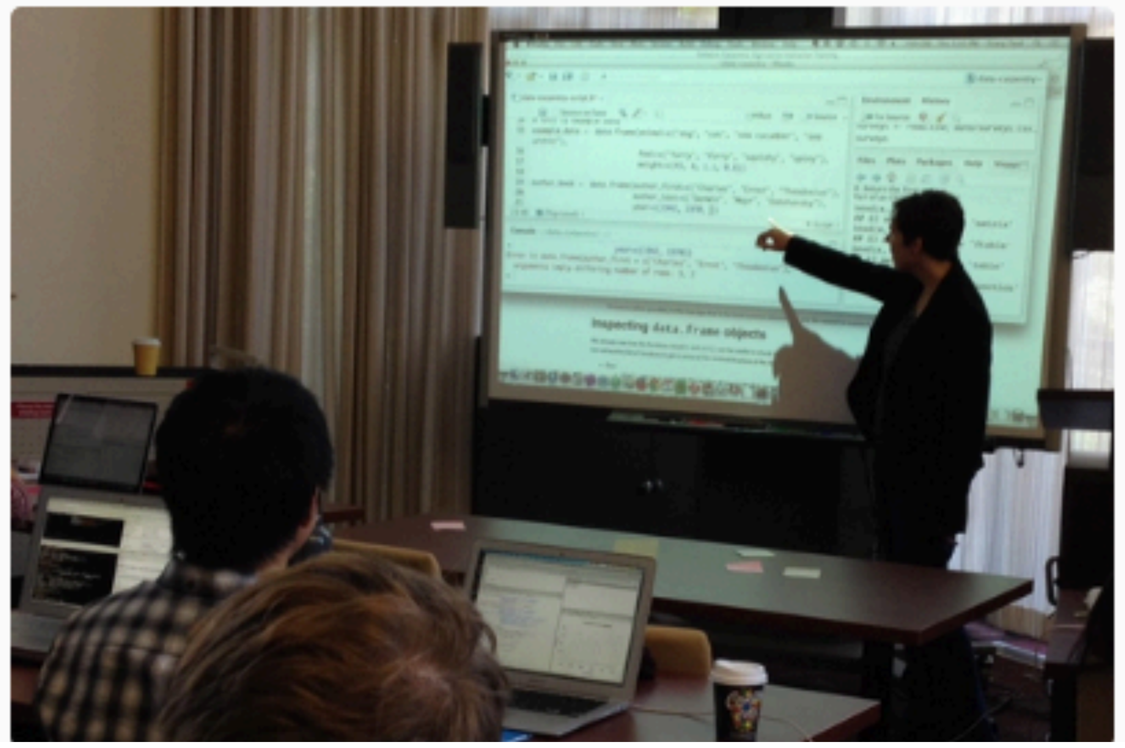
Data Carpentry develops and teaches workshops on the fundamental data skills needed to conduct research. Our mission is to provide researchers high-quality, domain-specific training covering the full lifecycle of data-driven research.

Data Carpentry is now a lesson program within **The Carpentries**, having merged with **Software Carpentry** in January, 2018. Data Carpentry's focus is on the introductory computational skills needed for data management and analysis in all domains of research. Our lessons are domain-specific, and build on the existing knowledge of learners to enable them to quickly apply skills learned to their own research. *Our initial target audience is learners who have little to no prior computational experience.* We create a friendly environment for learning to empower researchers and enable data driven discovery.

Host a Workshop

Attend a Workshop

Get Involved



Lessons in English

Lesson	Site	Repository	Reference	Instructor Notes
Ecology Workshop Overview				
Data Organization in Spreadsheets for Ecologists				
Data Cleaning with OpenRefine for Ecologists				
Data Management with SQL for Ecologists				
Data Analysis and Visualization in R for Ecologists				
Data Analysis and Visualization in Python for Ecologists				

https://datacarpentry.org/

Scientists rename human genes to stop Microsoft Excel from misreading them as dates

Sometimes it's easier to rewrite genetics than update Excel

By James Vincent | Aug 6, 2020, 8:44am EDT

STUDIES FOUND A FIFTH OF GENETIC DATA IN PAPERS WAS AFFECTED BY EXCEL ERRORS

WHY DID MICROSOFT WIN IN A FIGHT AGAINST HUMAN GENETICS?

Janna Hutz
@jannahutz

THRILLED by this announcement by the Human Gene Nomenclature Committee.

- **Symbols that affect data handling and retrieval.** For example, all symbols that autoconverted to dates in Microsoft Excel have been changed (for example, SEPT1 is now SEPTIN1; MARCH1 is now MARCHF1); tRNA synthetase symbols that were also common words have been changed (for example, WARS is now WARS1; CARS is now CARS1).

5:08 PM · Aug 4, 2020

1.5K   611 people are Tweeting about this

https://www.theverge.com/2020/8/6/21355674

Scientists rename human genes to stop Microsoft Excel from misreading them as dates

Sometimes it's easier to rewrite genetics than update Excel

By James Vincent | Aug 6, 2020, 8:44am EDT

STUDIES F
GENETIC D
AFFECTED

Careful primary data entry, okay with rules, tests

Never process data in Excel; use code

WHY DID MICROSOFT WIN IN A FIGHT AGAINST HUMAN GENETICS?

Microsoft Excel have been changed (for example, SEPT1 is now SEPTIN1; MARCH1 is now MARCHF1); tRNA synthetase symbols that were also common words have been changed (for example, WARS is now WARS1; CARS is now CARS1).

5:08 PM · Aug 4, 2020

1.5K 611 people are Tweeting about this

https://www.theverge.com/2020/8/6/21355674

PAUSE

Reporting

Sharing materials

Describing methods

Describing data

Describing results

Making decisions

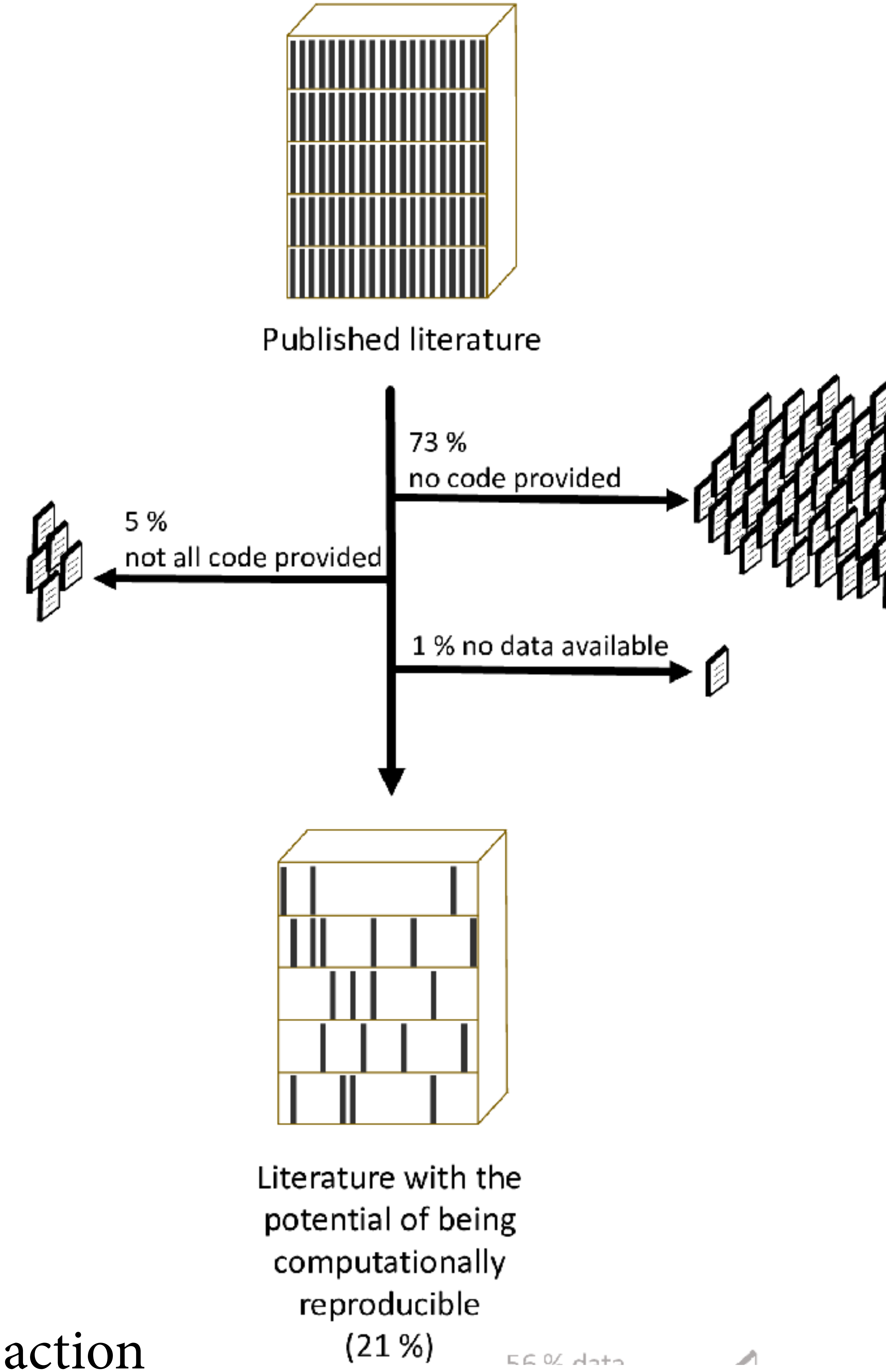
Sharing Materials

The paper is an advertisement; the data and its analysis are the product

Make code and data available through a link, not “by request”

Some data not shareable; code always shareable

Archived code & data will be required



Culina et al 2020 Low availability of code in ecology: A call for urgent action

Describing Methods

Minimal information:

(1) Math-stats notation of stat model

(2) Explanation of how (1) provides estimand

(3) Algorithm used to produce estimate

(4) Diagnostics, code tests

(5) Cite software packages

$$G_{AB} \sim \text{Poisson}(\lambda_{AB})$$

$$\log(\lambda_{AB}) = \alpha + T_{AB} + G_A + R_B$$

$$G_{BA} \sim \text{Poisson}(\lambda_{BA})$$

$$\log(\lambda_{BA}) = \alpha + T_{BA} + G_B + R_A$$

$$\begin{pmatrix} T_{AB} \\ T_{BA} \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right)$$

$$\rho \sim \text{LKJCorr}(2)$$

$$\sigma \sim \text{Exponential}(1)$$

$$\alpha \sim \text{Normal}(0,1)$$

$$\begin{pmatrix} G_A \\ R_A \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{R}_{GR}, \mathbf{S}_{GR} \right)$$

$$\mathbf{R}_{GR} \sim \text{LKJCorr}(2)$$

$$\mathbf{S}_{GR} \sim \text{Exponential}(1)$$

To estimate the reciprocity within dyads, we model the correlation within dyads in giving, using a multilevel mixed-membership model (*textbook citation*). To control for confounding from generalized giving and receiving, as indicated by the DAG in the previous section, we stratify giving and receiving by household. The full model with priors is presented at right. We estimated the posterior distribution using Hamiltonian Monte Carlo as implemented in Stan version 2.29 (*citation*). We validated the model on simulated data and assessed convergence by inspection of trace plots, R-hat values, and effective sample sizes. Diagnostics are reported in Appendix B and all results can be replicated using the code available at **LINK**.

$$G_{AB} \sim \text{Poisson}(\lambda_{AB})$$

$$\log(\lambda_{AB}) = \alpha + T_{AB} + G_A + R_B$$

$$G_{BA} \sim \text{Poisson}(\lambda_{BA})$$

$$\log(\lambda_{BA}) = \alpha + T_{BA} + G_B + R_A$$

$$\begin{pmatrix} T_{AB} \\ T_{BA} \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right)$$

$$\rho \sim \text{LKJCorr}(2)$$

$$\sigma \sim \text{Exponential}(1)$$

$$\alpha \sim \text{Normal}(0,1)$$

$$\begin{pmatrix} G_A \\ R_A \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{R}_{GR}, \mathbf{S}_{GR} \right)$$

$$\mathbf{R}_{GR} \sim \text{LKJCorr}(2)$$

$$\mathbf{S}_{GR} \sim \text{Exponential}(1)$$

To estimate the reciprocity within dyads, we model the correlation within dyads in giving, using a multilevel mixed-membership model (*textbook citation*). To control for confounding from generalized giving and receiving, as indicated by the DAG in the previous section, we stratify giving and receiving by household. The full model with priors is presented at right. We estimated the posterior distribution using Hamiltonian Monte Carlo as implemented in Stan version 2.29 (*citation*). We validated the model on simulated data and assessed convergence by inspection of trace plots, R-hat values, and effective sample sizes. Diagnostics are reported in Appendix B and all results can be replicated using the code available at **LINK**.

$$G_{AB} \sim \text{Poisson}(\lambda_{AB})$$

$$\log(\lambda_{AB}) = \alpha + T_{AB} + G_A + R_B$$

$$G_{BA} \sim \text{Poisson}(\lambda_{BA})$$

$$\log(\lambda_{BA}) = \alpha + T_{BA} + G_B + R_A$$

$$\begin{pmatrix} T_{AB} \\ T_{BA} \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right)$$

$$\rho \sim \text{LKJCorr}(2)$$

$$\sigma \sim \text{Exponential}(1)$$

$$\alpha \sim \text{Normal}(0,1)$$

$$\begin{pmatrix} G_A \\ R_A \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{R}_{GR}, \mathbf{S}_{GR} \right)$$

$$\mathbf{R}_{GR} \sim \text{LKJCorr}(2)$$

$$\mathbf{S}_{GR} \sim \text{Exponential}(1)$$

To estimate the reciprocity within dyads, we model the correlation within dyads in giving, using a multilevel mixed-membership model (*textbook citation*). To control for confounding from generalized giving and receiving, as indicated by the DAG in the previous section, we stratify giving and receiving by household. The full model with priors is presented at right. We estimated the posterior distribution using Hamiltonian Monte Carlo as implemented in Stan version 2.29 (*citation*). We validated the model on simulated data and assessed convergence by inspection of trace plots, R-hat values, and effective sample sizes. Diagnostics are reported in Appendix B and all results can be replicated using the code available at **LINK**.

$$G_{AB} \sim \text{Poisson}(\lambda_{AB})$$
$$\log(\lambda_{AB}) = \alpha + T_{AB} + G_A + R_B$$
$$G_{BA} \sim \text{Poisson}(\lambda_{BA})$$
$$\log(\lambda_{BA}) = \alpha + T_{BA} + G_B + R_A$$
$$\begin{pmatrix} T_{AB} \\ T_{BA} \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right)$$
$$\rho \sim \text{LKJCorr}(2)$$
$$\sigma \sim \text{Exponential}(1)$$
$$\alpha \sim \text{Normal}(0,1)$$
$$\begin{pmatrix} G_A \\ R_A \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{R}_{GR}, \mathbf{S}_{GR} \right)$$
$$\mathbf{R}_{GR} \sim \text{LKJCorr}(2)$$
$$\mathbf{S}_{GR} \sim \text{Exponential}(1)$$

To estimate the reciprocity within dyads, we model the correlation within dyads in giving, using a multilevel mixed-membership model (*textbook citation*). To control for confounding from generalized giving and receiving, as indicated by the DAG in the previous section, we stratify giving and receiving by household. The full model with priors is presented at right. We estimated the posterior distribution using Hamiltonian Monte Carlo as implemented in Stan version 2.29 (*citation*). We validated the model on simulated data and assessed convergence by inspection of trace plots, R-hat values, and effective sample sizes. Diagnostics are reported in Appendix B and all results can be replicated using the code available at **LINK**.

$$G_{AB} \sim \text{Poisson}(\lambda_{AB})$$

$$\log(\lambda_{AB}) = \alpha + T_{AB} + G_A + R_B$$

$$G_{BA} \sim \text{Poisson}(\lambda_{BA})$$

$$\log(\lambda_{BA}) = \alpha + T_{BA} + G_B + R_A$$

$$\begin{pmatrix} T_{AB} \\ T_{BA} \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right)$$

$$\rho \sim \text{LKJCorr}(2)$$

$$\sigma \sim \text{Exponential}(1)$$

$$\alpha \sim \text{Normal}(0,1)$$

$$\begin{pmatrix} G_A \\ R_A \end{pmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{R}_{GR}, \mathbf{S}_{GR} \right)$$

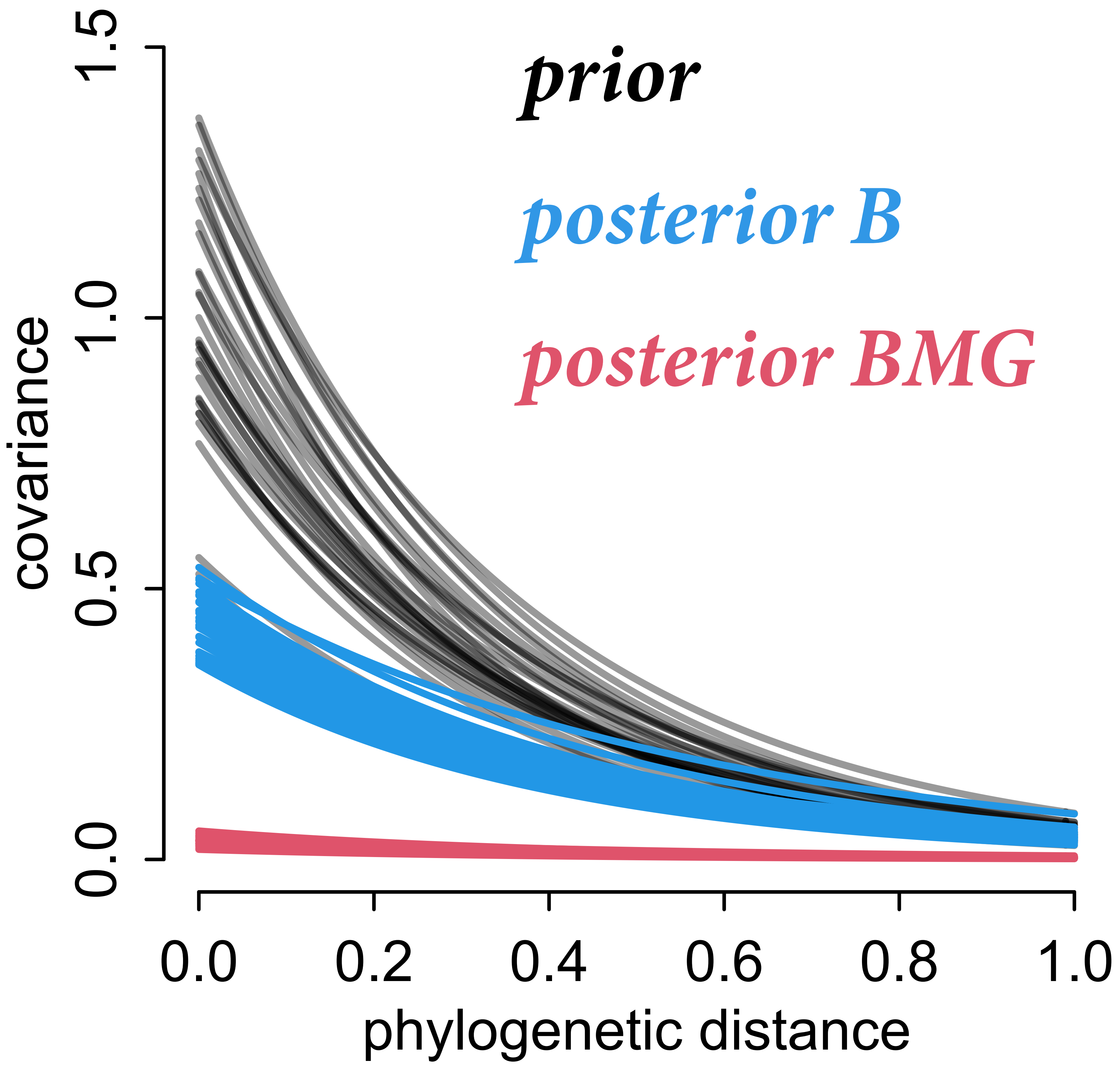
$$\mathbf{R}_{GR} \sim \text{LKJCorr}(2)$$

$$\mathbf{S}_{GR} \sim \text{Exponential}(1)$$

Justify Priors

"Priors were chosen through prior predictive simulation so that pre-data predictions span only the range of scientifically plausible outcomes.

In the results, we explicitly compare the posterior distribution to the prior, so that the impact of the sample is obvious."



Justifying Methods

Naive reviewers: “*Good science doesn’t need complex stats*”

Causal model often requires complexity

Big data => unit heterogeneity

Ethical responsibility to do our best

Change discussion from statistics to causal models

“Pooh?” said Piglet.
“Yes, Piglet?” said Pooh.
“27417 parameters,” said Piglet.
“Oh, bother,” said Pooh.

Justifying Methods

Write for the editor, not the reviewer

Find other papers in discipline/journal that have used Bayesian methods or similar models (Bayesian or not)

Explain results in Bayesian terms, show densities, cite disciplinary guides

Bayes is ancient, normative, often the only practical way to estimate complex models

"Pooh?" said Piglet.
"Yes, Piglet?" said Pooh.
"27417 parameters," said Piglet.
"Oh, bother," said Pooh.

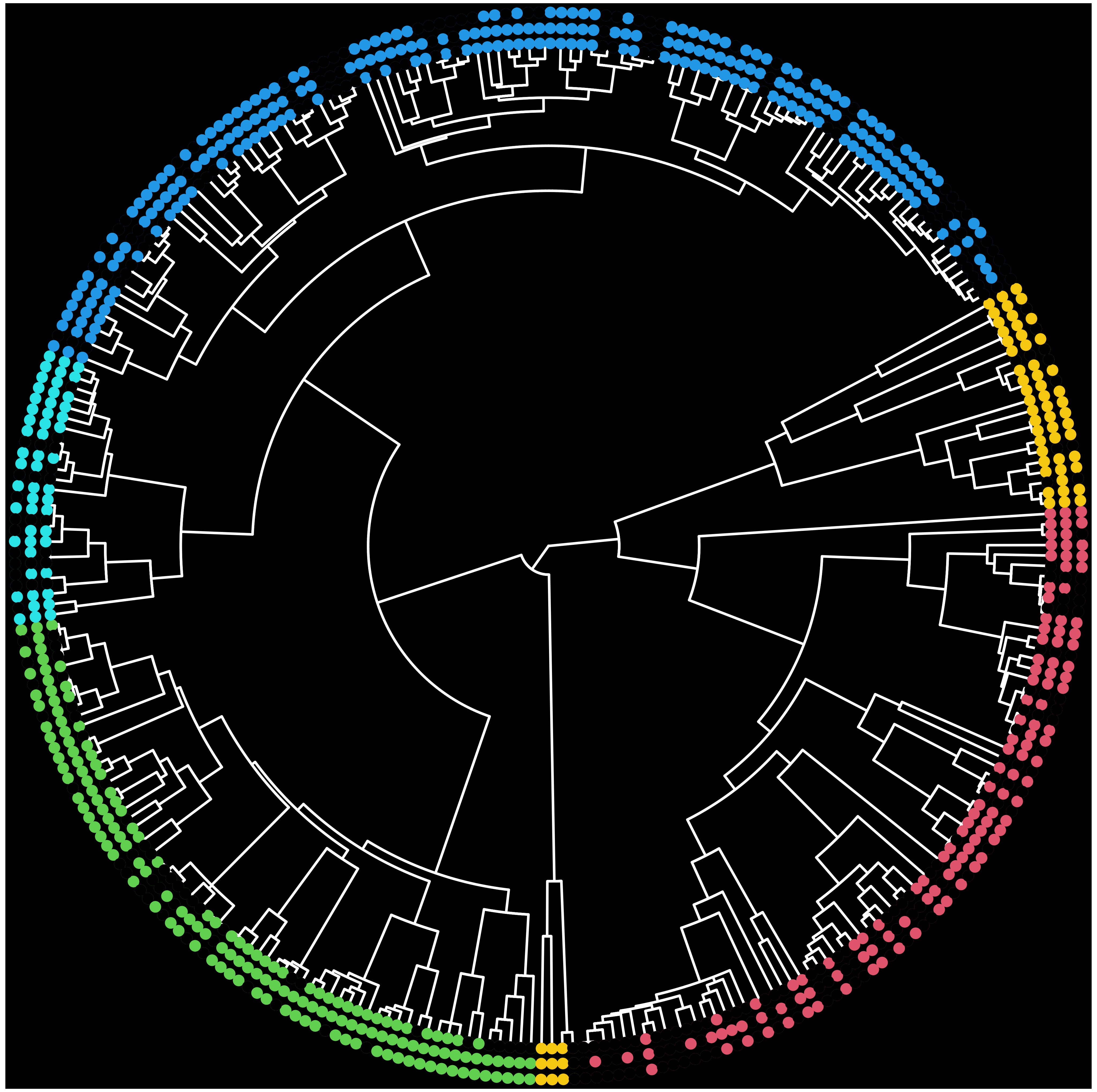
Describing Data

1k observations of 1 person
-vs-
1 observation of each of 1k people

“Effective” sample size function of estimand and hierarchical structure

Variables measured at which levels?

Missing values!



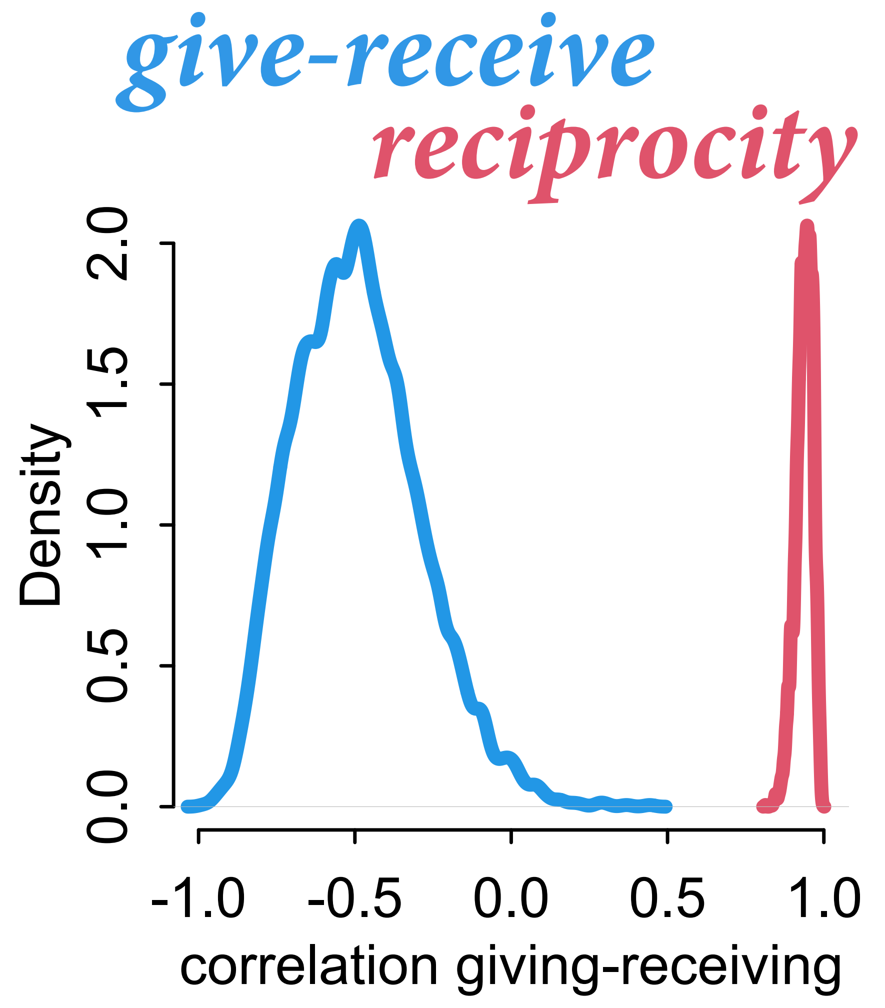
Describing Results

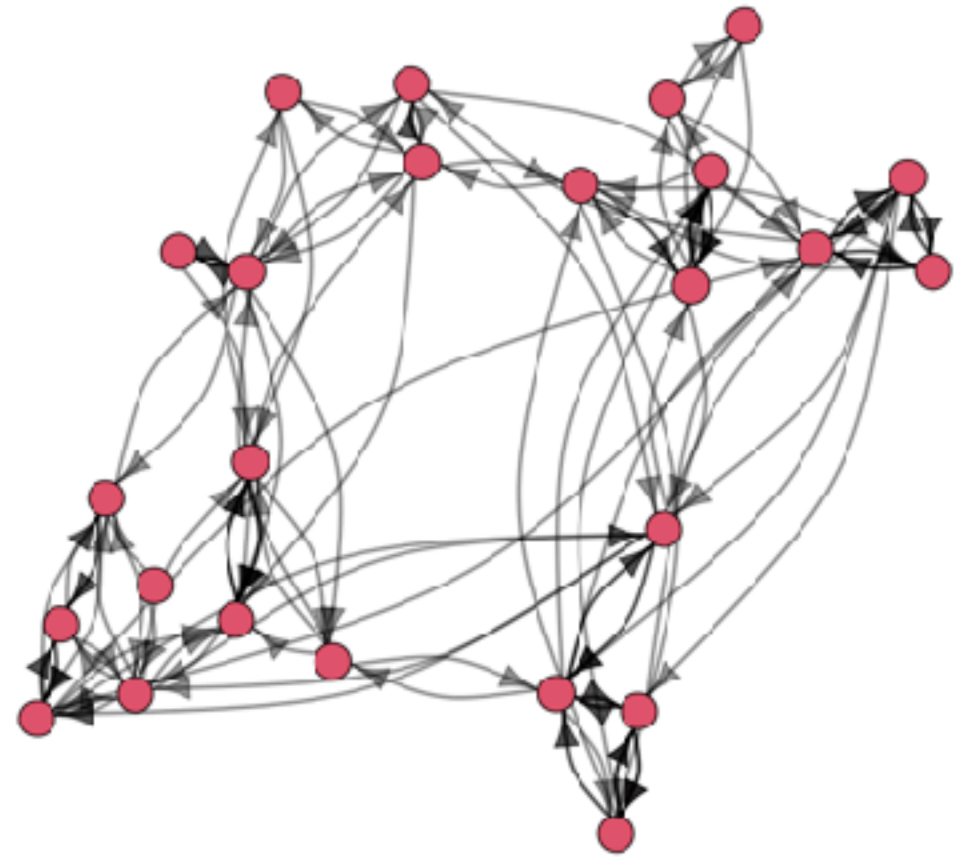
Estimands, marginal causal effects

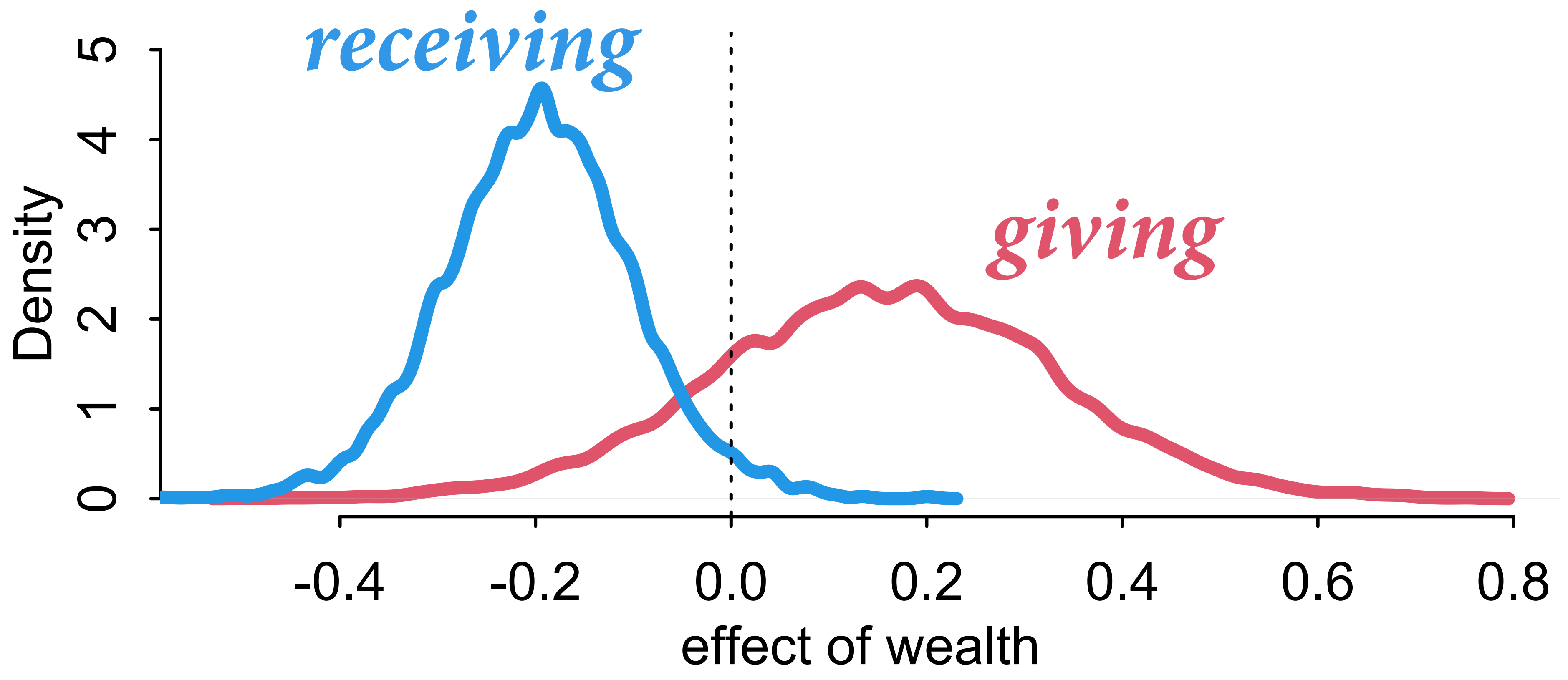
Warn against causal interpretation of control variables (Table 2 fallacy)

Densities better than intervals; Sample realizations often better than densities

Figures assist comparisons

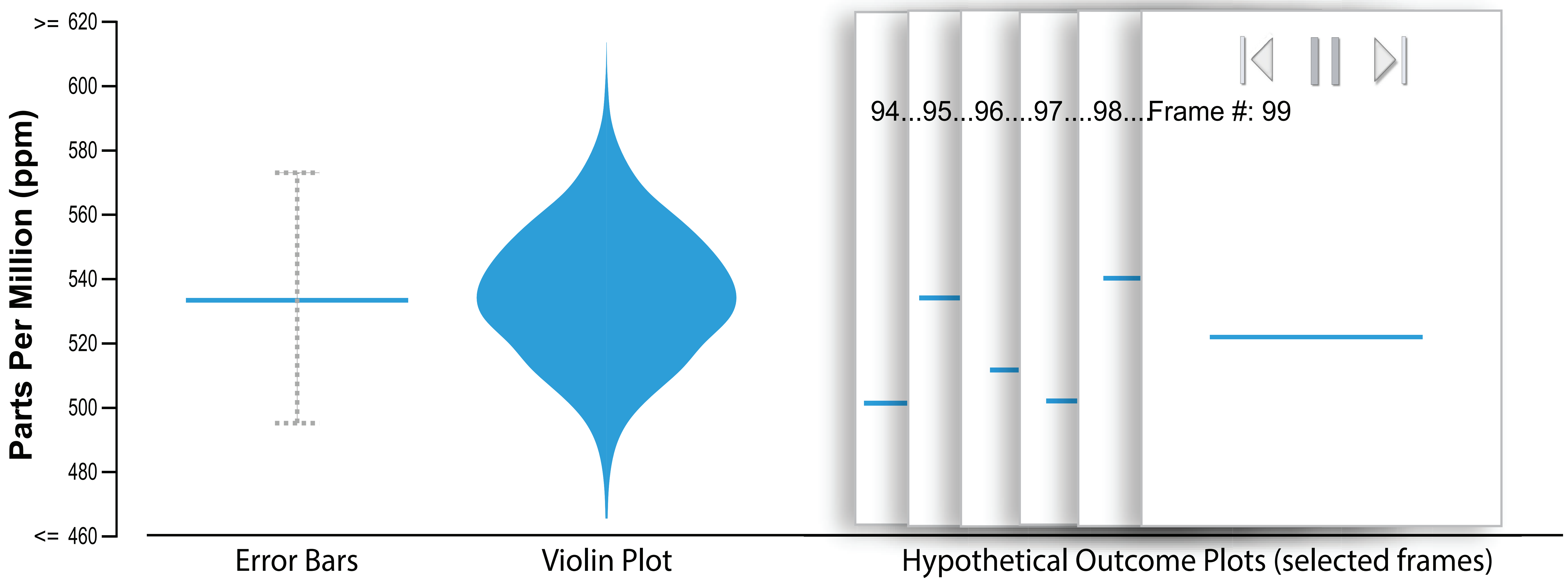


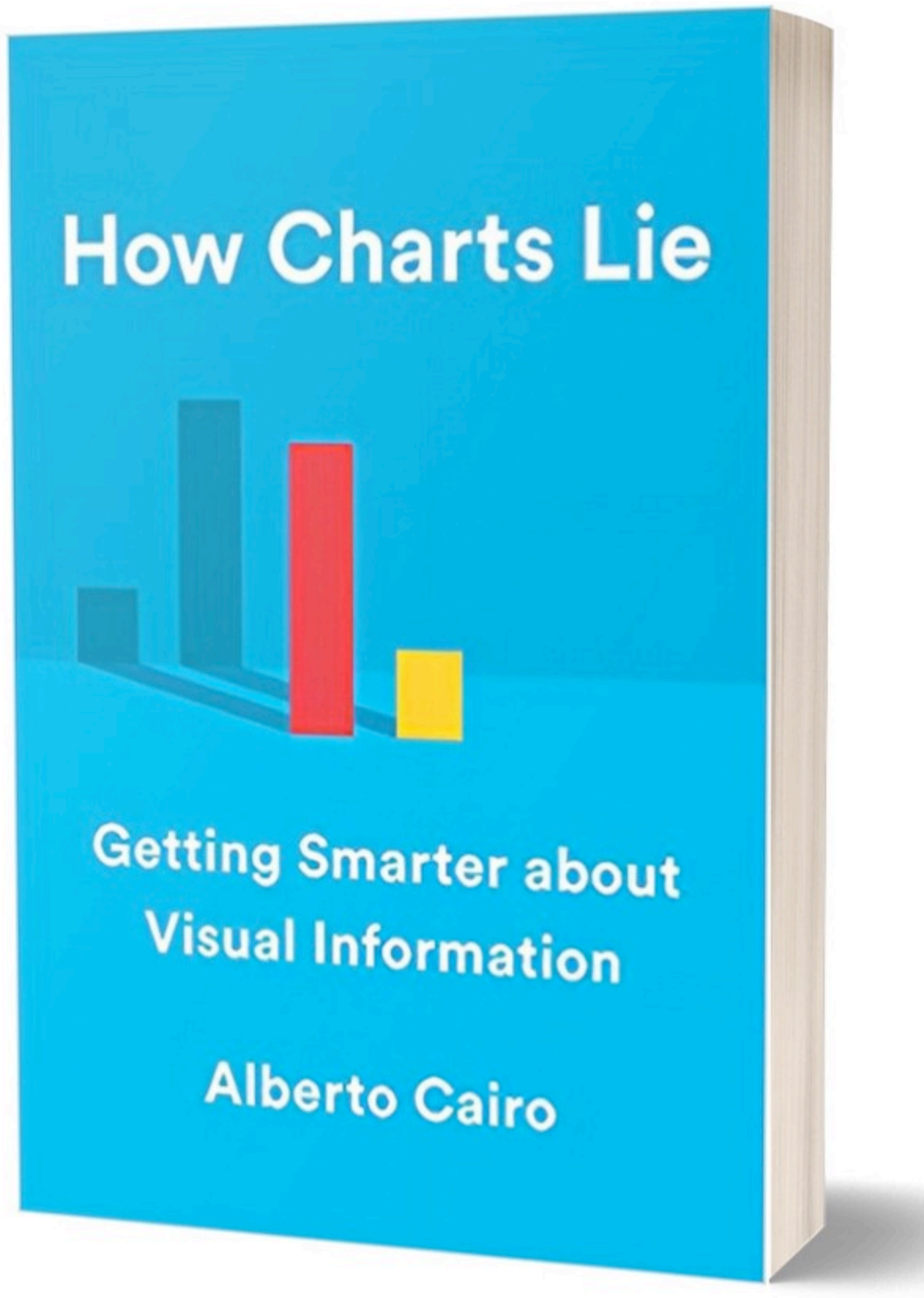




Hypothetical Outcome Plots Outperform Error Bars and Violin Plots for Inferences About Reliability of Variable Ordering

Jessica Hullman[1,*], Paul Resnick[2], Eytan Adar[2],





Making Decisions

Academic research: Communicate uncertainty, conditional on sample & models

Industry research: What should we do, given the uncertainty, conditional on sample & models?

Also: "Does my boss have any idea what 'uncertainty' means, or does he think that's the refuge of cowards?"

Making Decisions

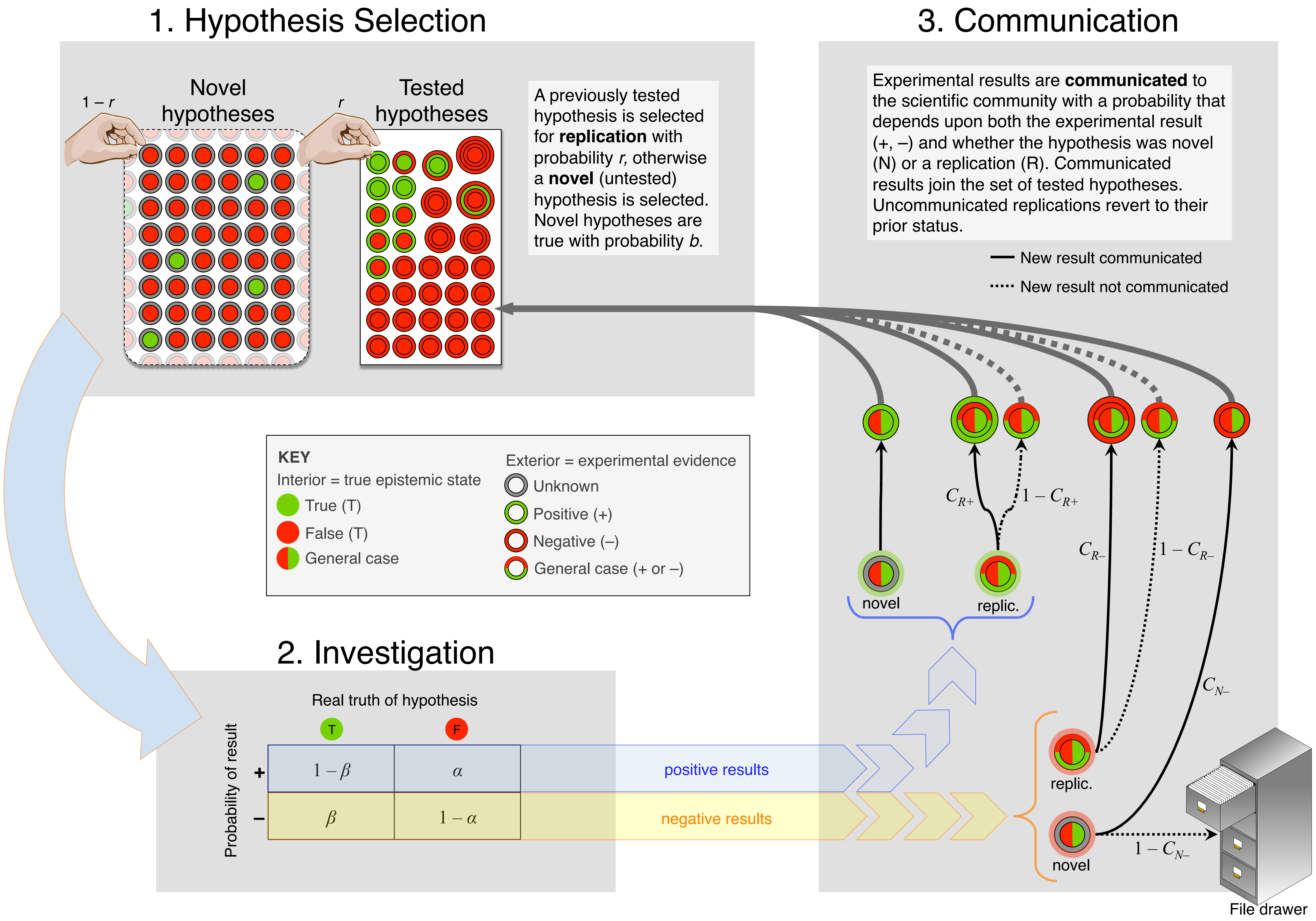
Bayesian decision theory:

(1) State costs & benefits of outcomes
(2) Compute posterior benefits of hypothetical policy choices

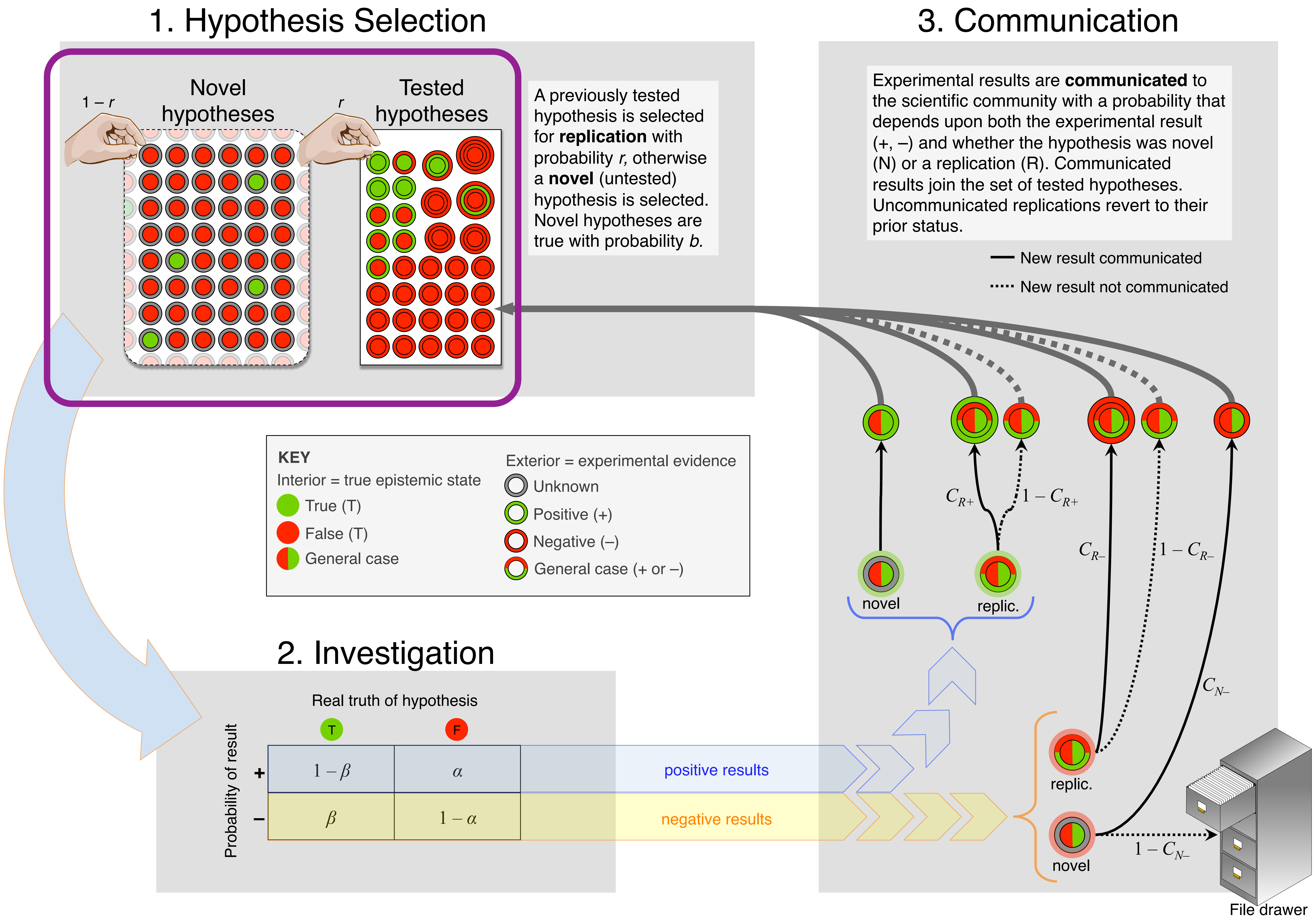
Simple example in Chapter 3

Can be integrated with dynamic optimization

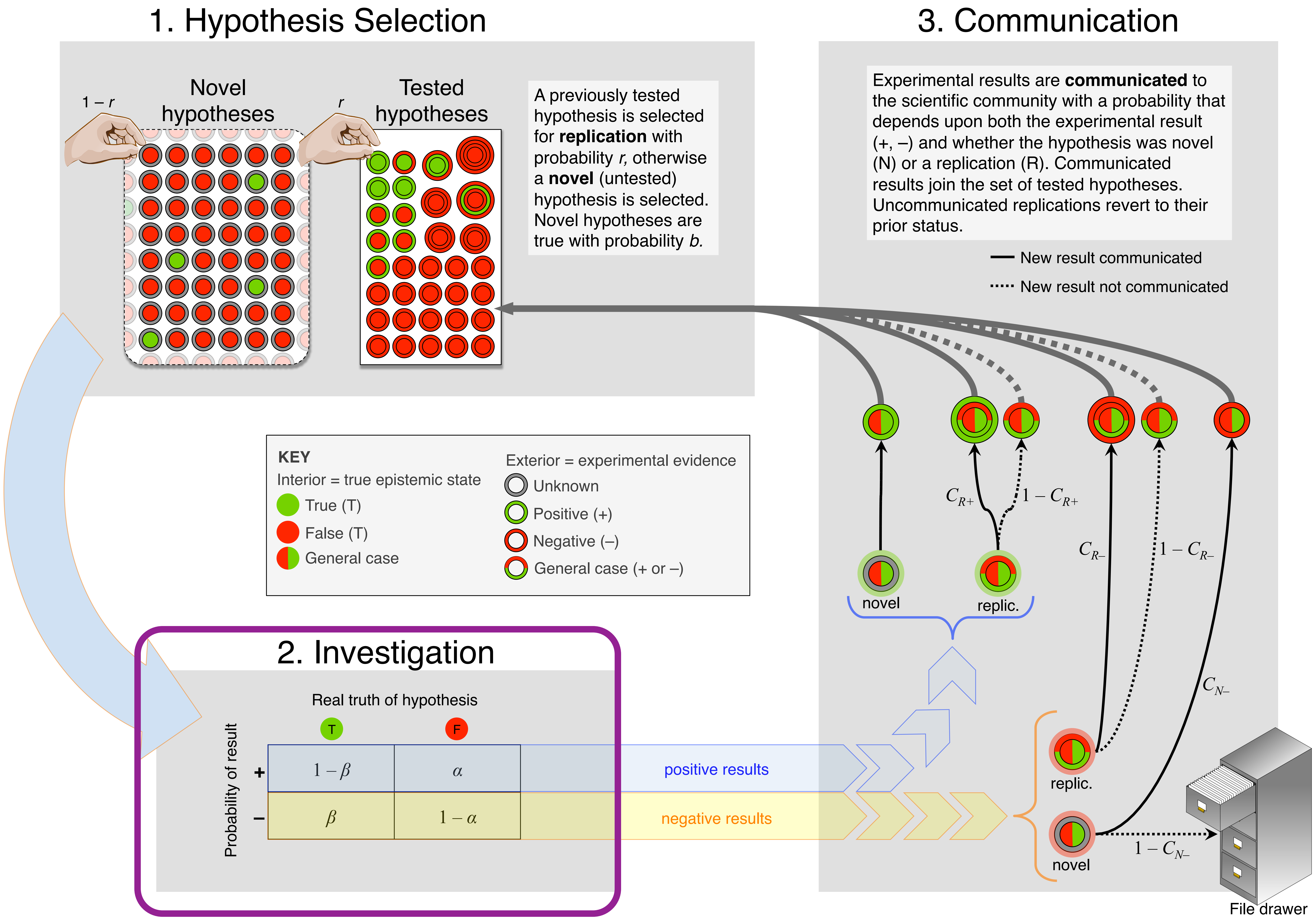

SCIENCE
ME DISCUSSING SCIENCE REFORM



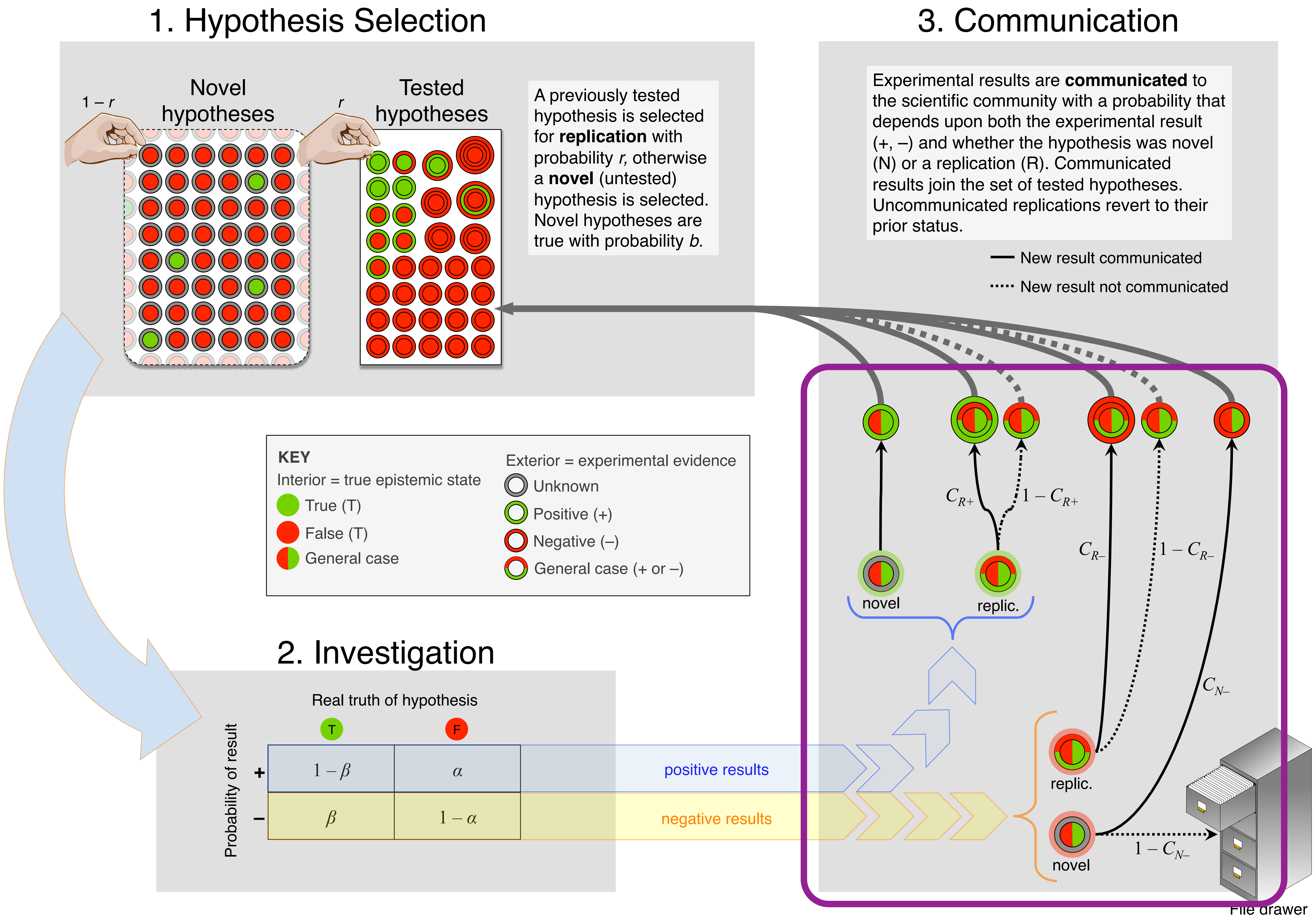
McElreath & Smaldino. 2015. Replication, communication, and the population dynamics of scientific discovery.



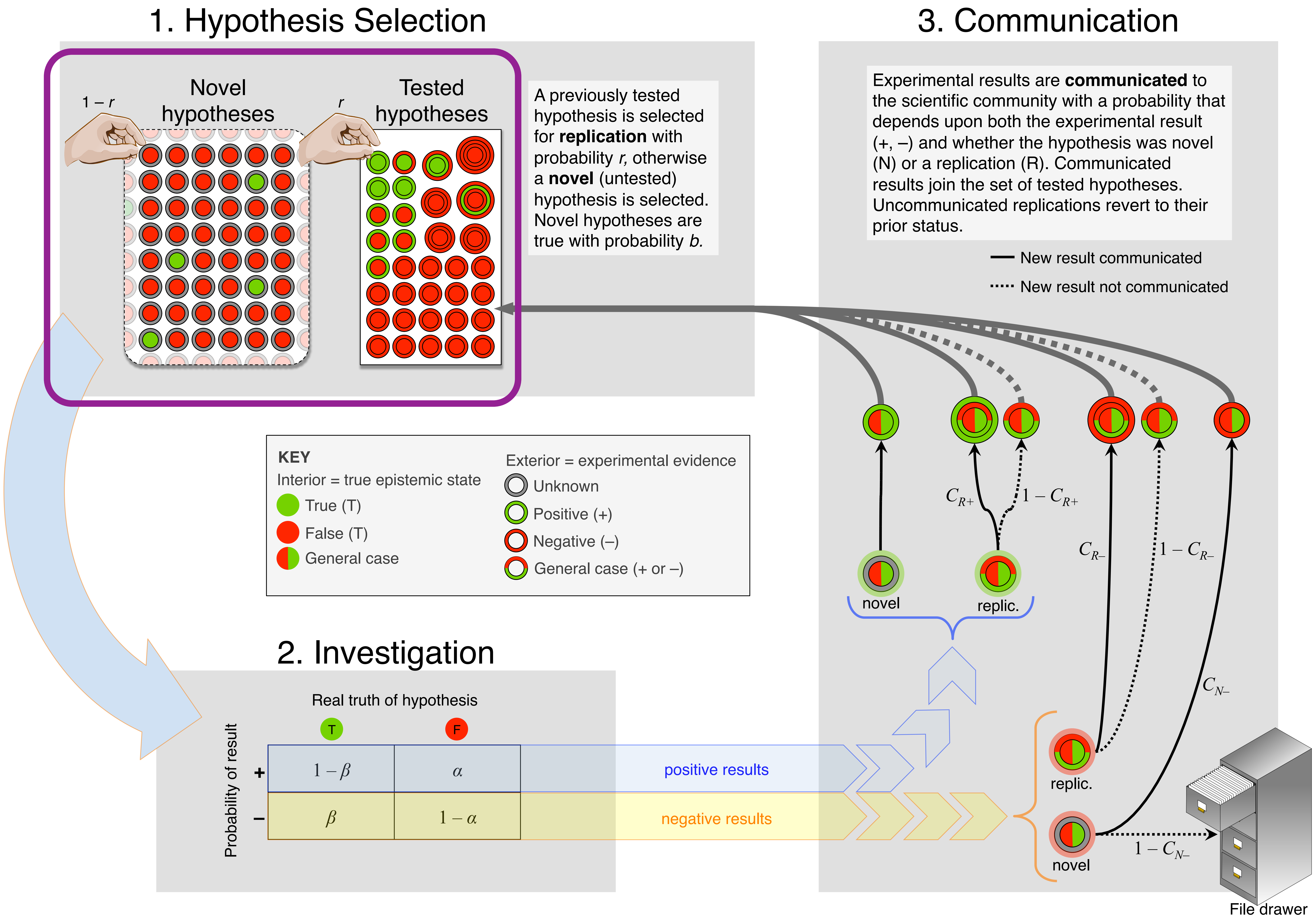
McElreath & Smaldino. 2015. Replication, communication, and the population dynamics of scientific discovery.



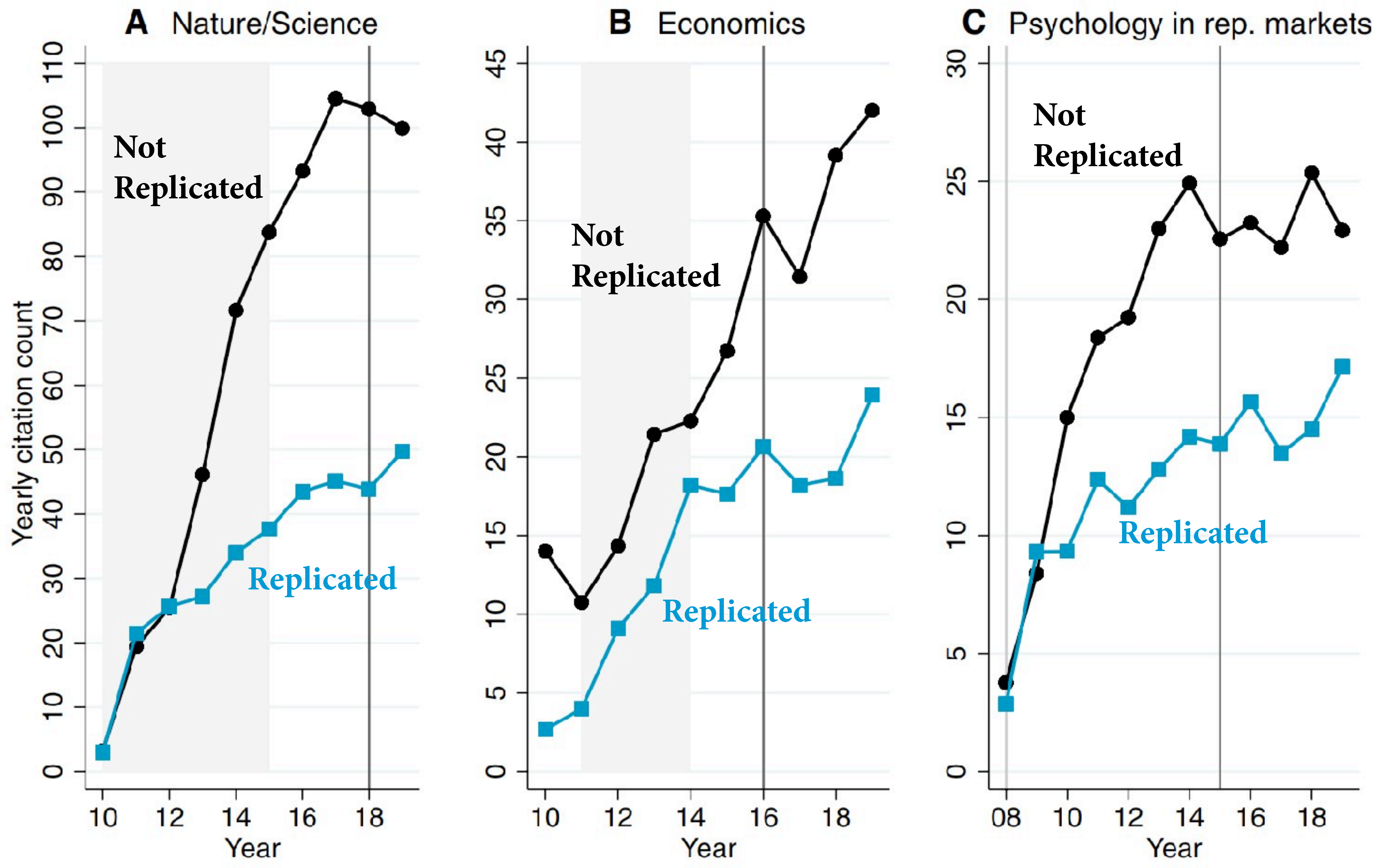
McElreath & Smaldino. 2015. Replication, communication, and the population dynamics of scientific discovery.



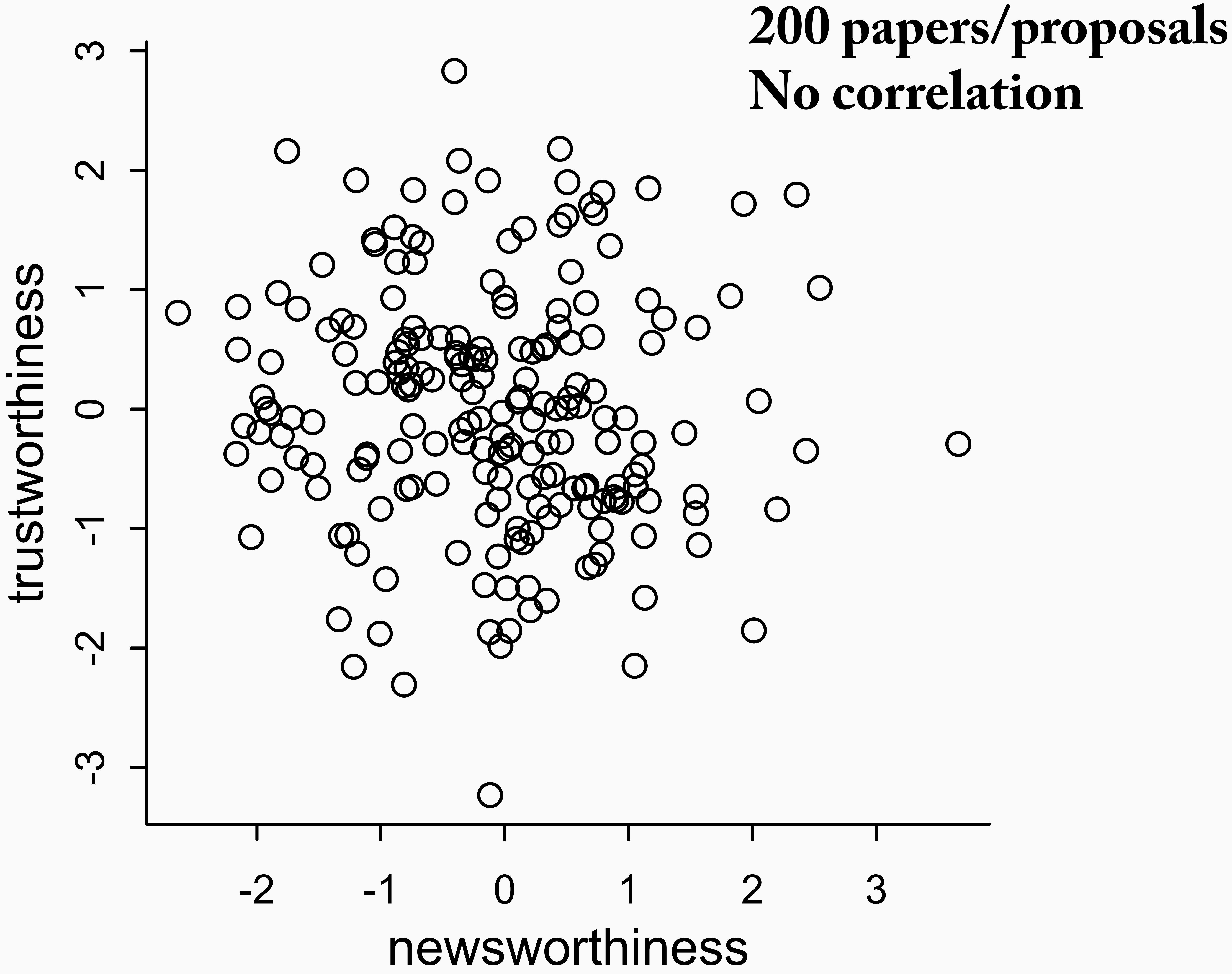
McElreath & Smaldino. 2015. Replication, communication, and the population dynamics of scientific discovery.



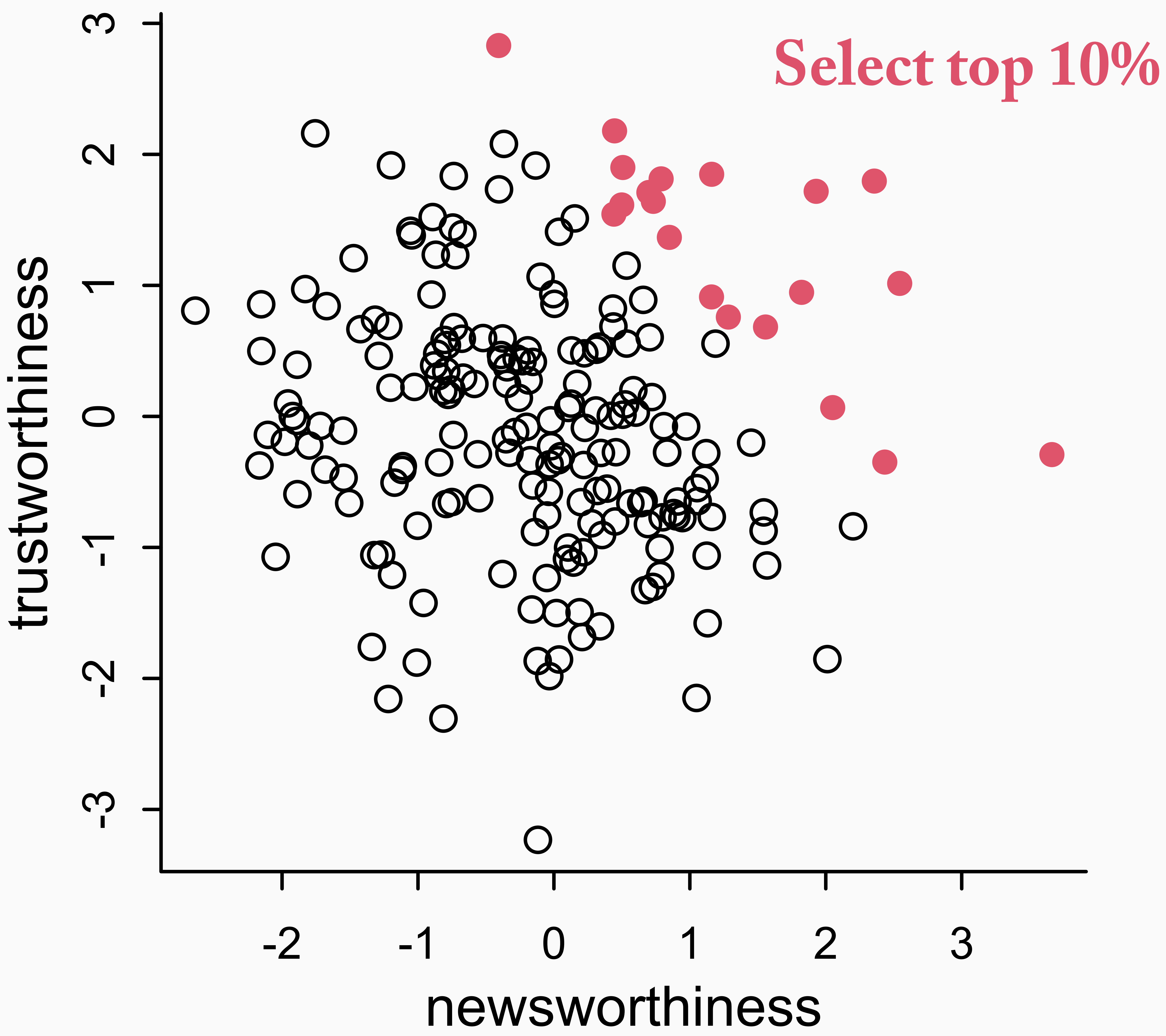
McElreath & Smaldino. 2015. Replication, communication, and the population dynamics of scientific discovery.



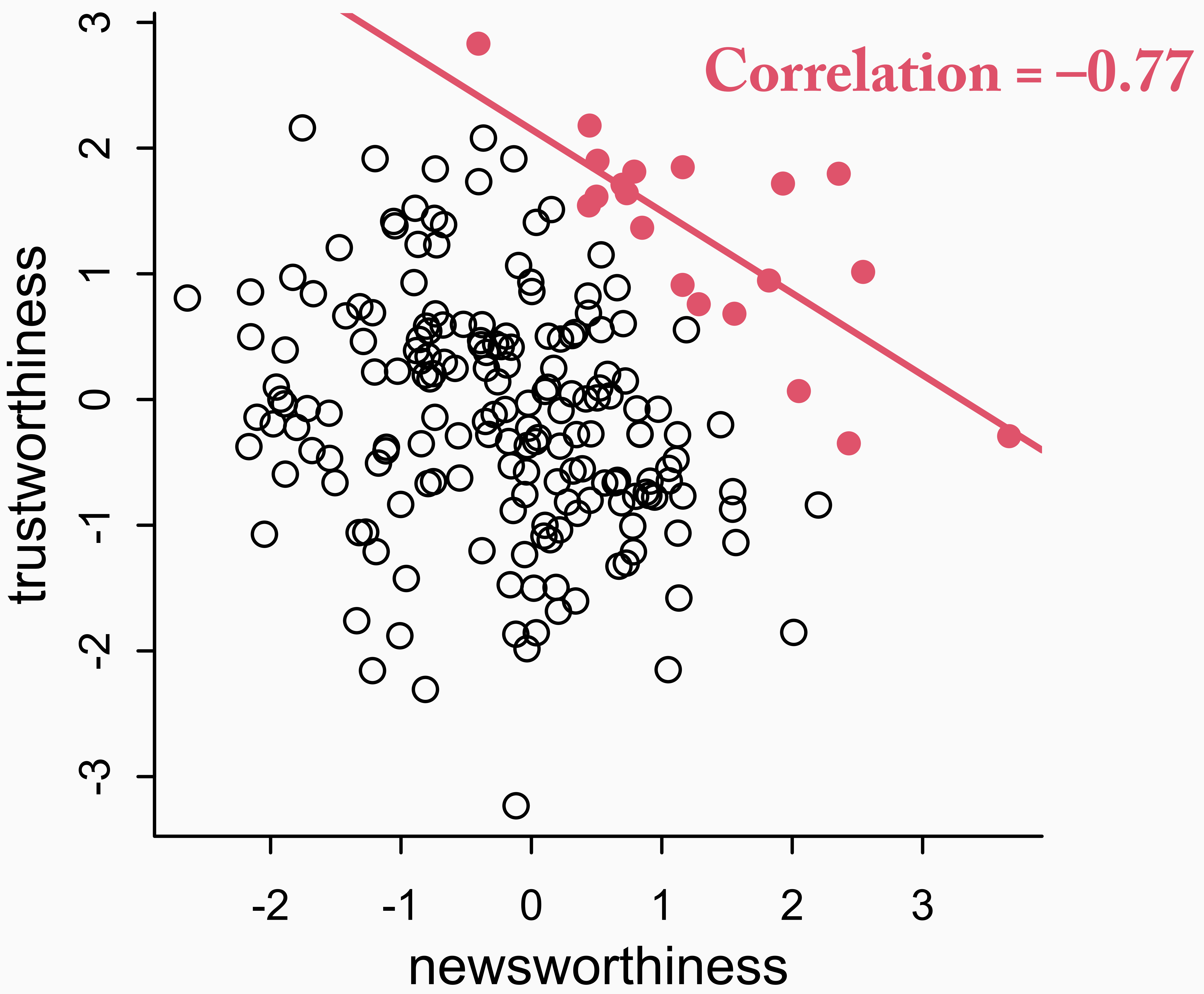
Serra-Garcia & Gneezy 2021 Nonreplicable publications are cited more than replicable ones



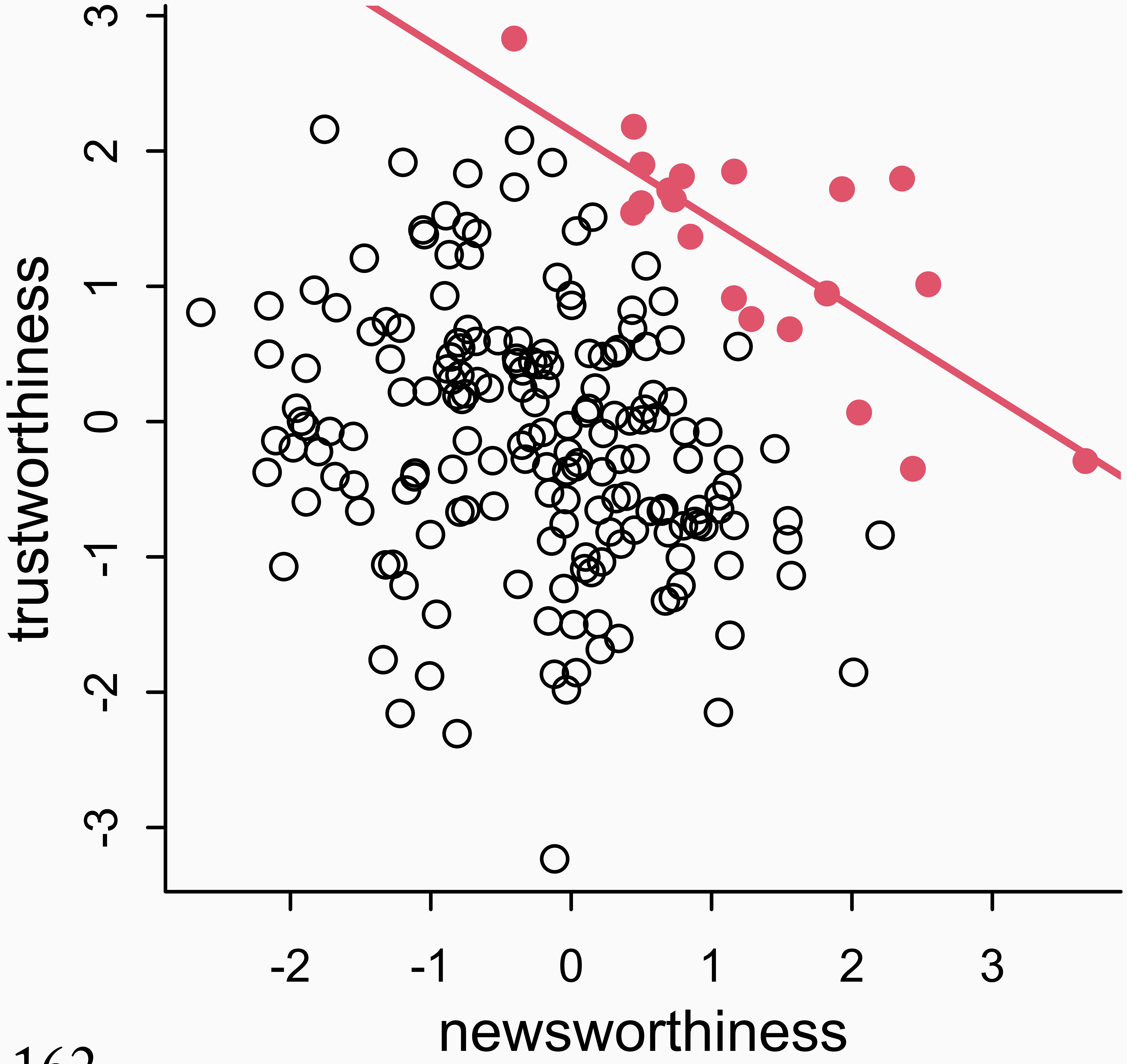
200 papers/proposals
No correlation
trustworthiness
newsworthiness
-3
-2
-1
0
1
2
3



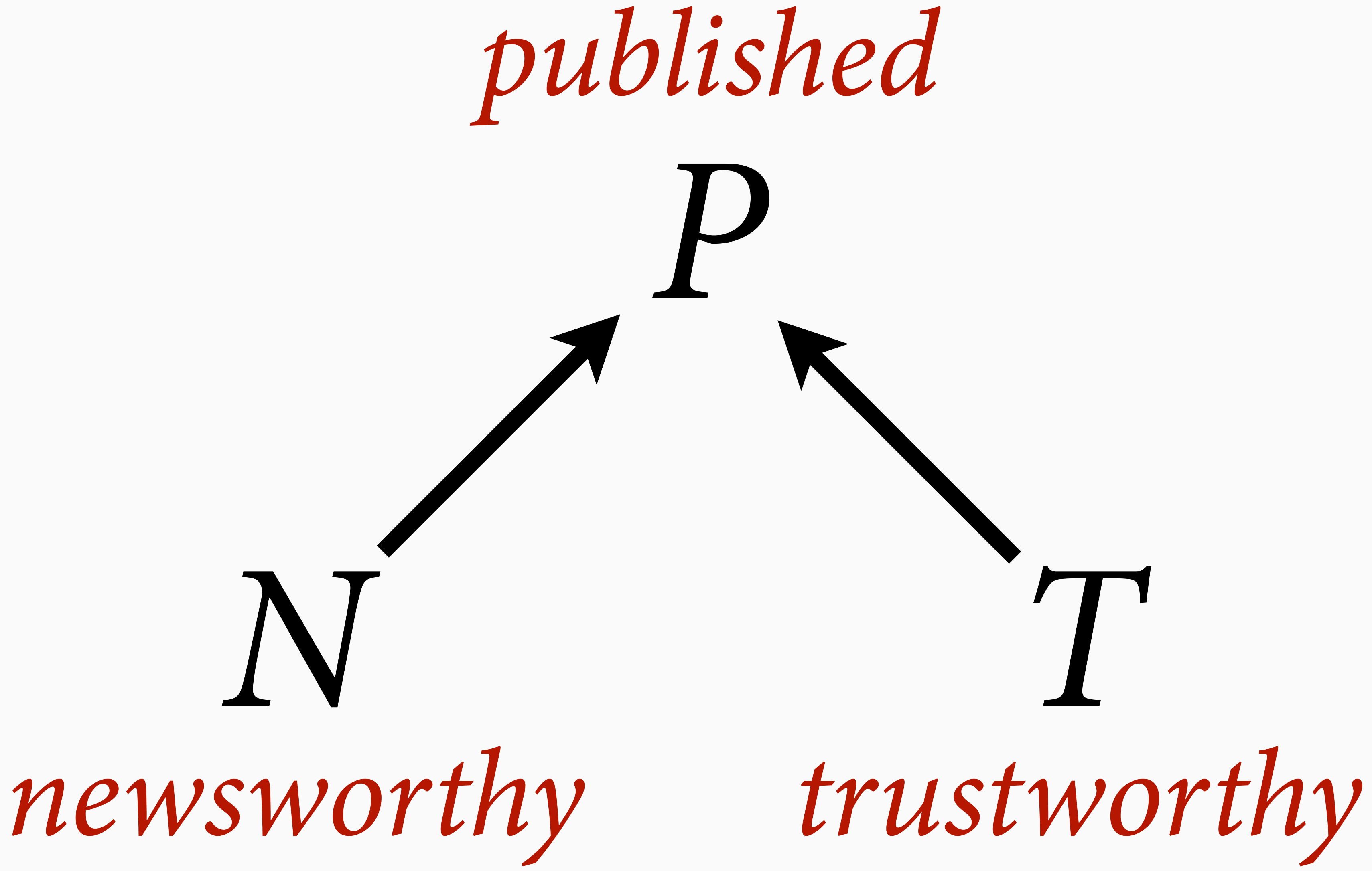
Select top 10%
trustworthiness
newsworthiness
3
2
1
0
-1
-2
-3
-2
-1
0
1
2
3



Correlation = –0.77
trustworthiness
newsworthiness
-3
-2
-1
0
1
2
3



trustworthiness
newsworthiness
3
2
1
0
-1
-2
-3
-2
-1
0
1
2
3



published
P
N
newsworthy
T
trustworthy

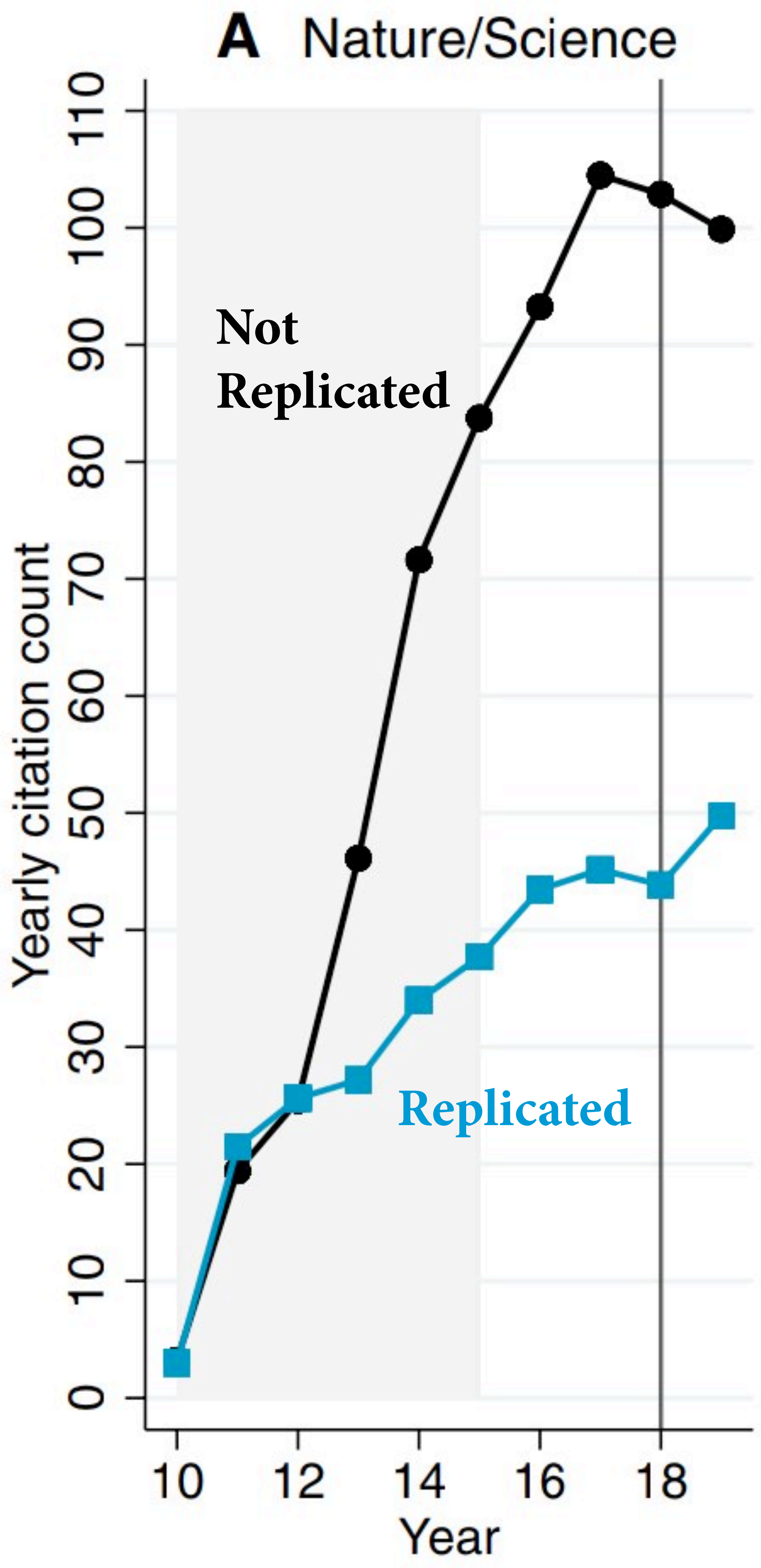
Horoscopes for Research

No one knows how research works

But many easy fixes at hand

(1) No stats without associated causal model
(2) Prove that your code works (in principle)
(3) Share as much as possible
(4) Beware proxies of research quality

Many things you dislike about academia were once well-intentioned reforms



END