Computational Modeling of Social Behavior

Day 4

Networks, etc.

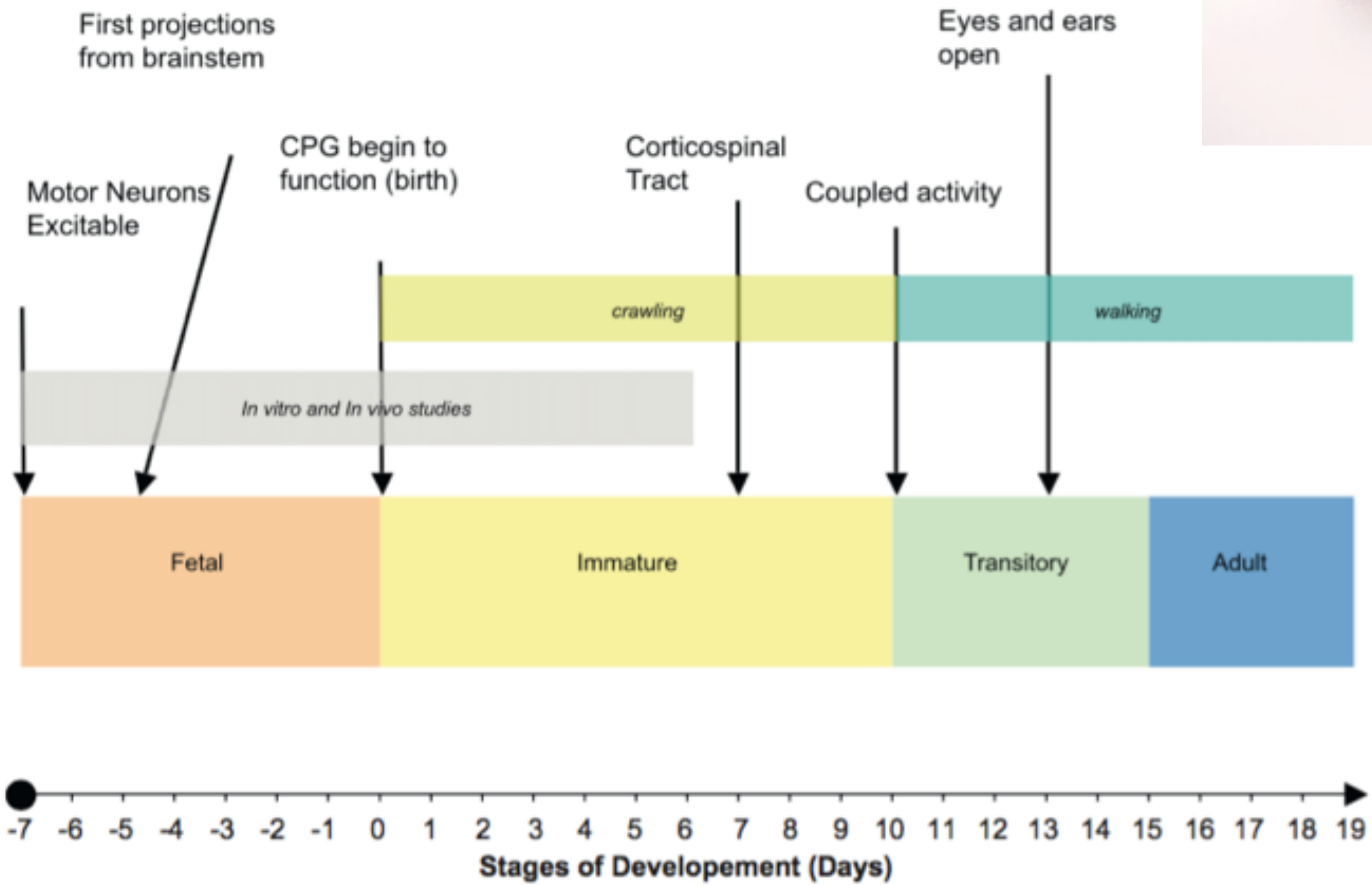
Paul Smaldino

Outline of the day

- **Morning**
  - Models and Empirical Data
  - Network Theory
- **Afternoon**
  - Modeling Agents on Networks
  - Coda: Why Model

What can we do with models?

- Scaffold theory development by creating mental models
- Explain generative mechanism behind existing data
- Predict future data

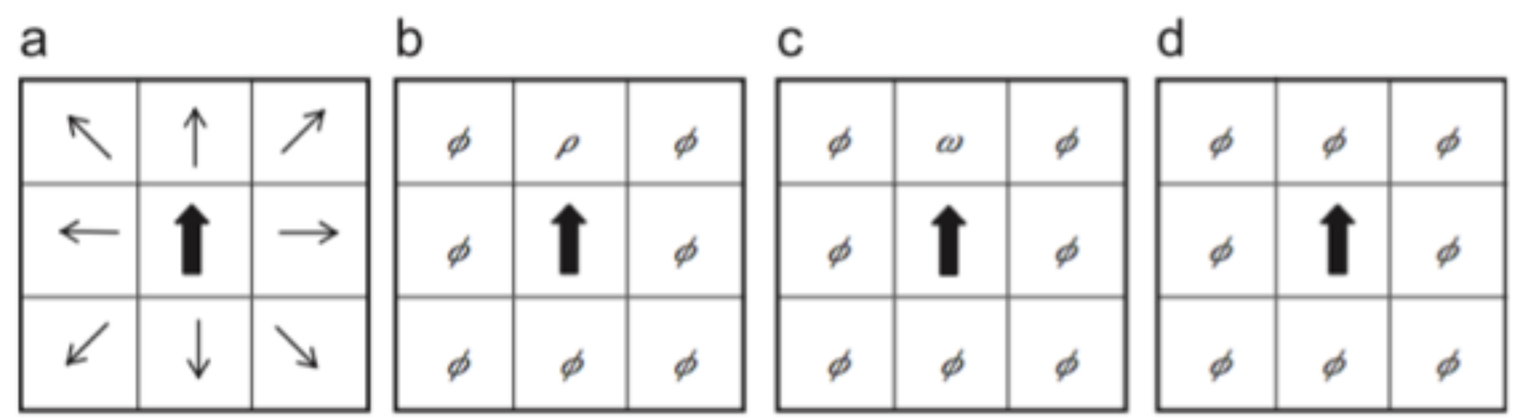


Schank & Alberts (2000) Proc R Soc B; Schank (2008) J. Theor. Biol.

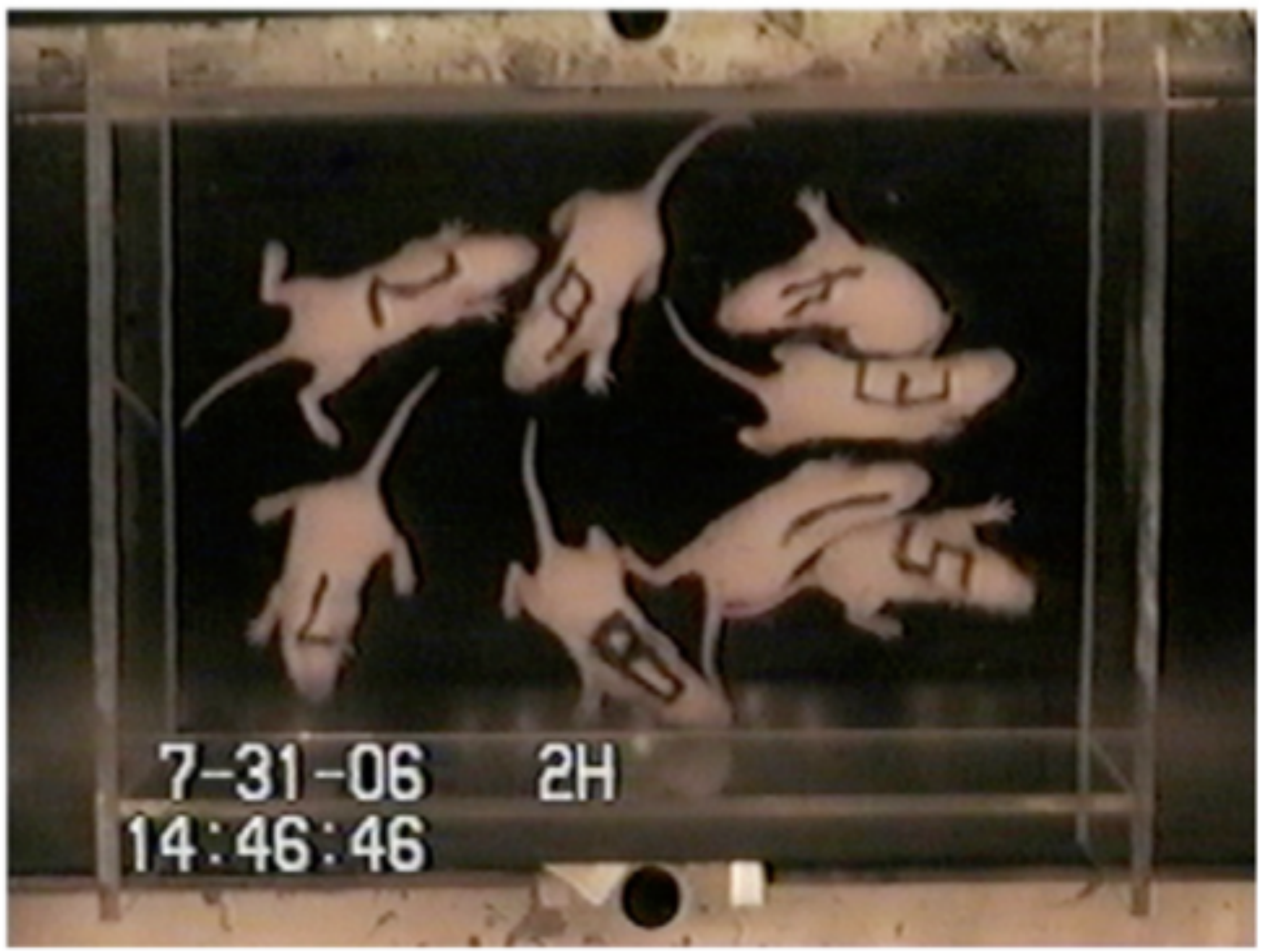
- Data collection:
  - Rat pups moved around in arena individually and in groups at 7 and 10 days old. Video capture.

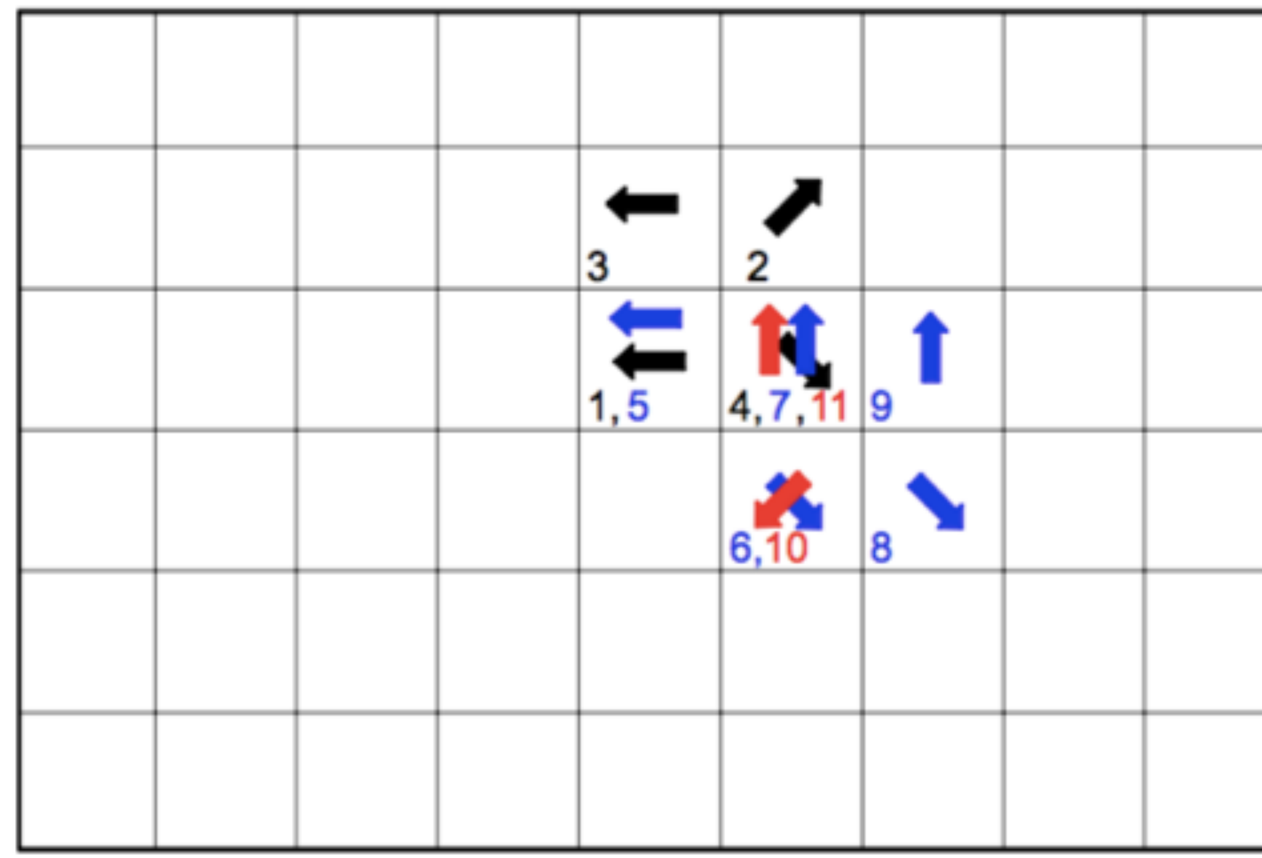
- Model:
  - Agents move through simulated arena,
  - Evolved contingent movement behaviors in response to nothing, wall, and other pups

$$R_1: \quad o_{12} = \begin{cases} \rho & \text{if any of the front three cells contain a pup} \\ \omega & \text{if any of the front three cells contain a wall and none contain a pup} \\ \phi & \text{if all three front cells are empty} \end{cases}$$



Schank (2008) J. Theor. Biol.

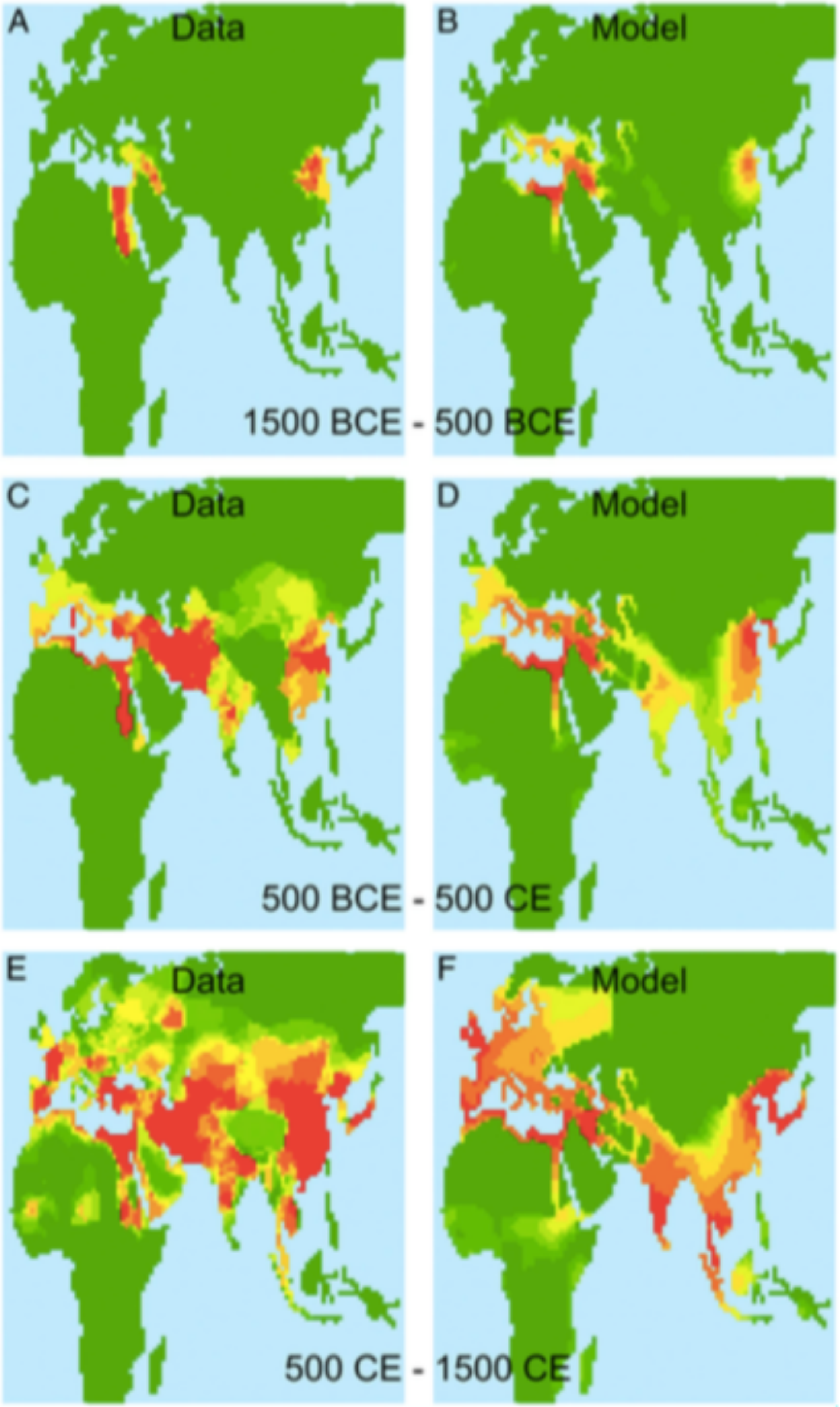




Schank (2008) J. Theor. Biol.

- Results:
    - Evolved models fit data than any null model
    - At 7 days, individual-evolved model fit group data with other pups treated as wall
    - At 10 days, individual-evolved model was terrible fit, required social contingent movement.
    - Supports conclusion that social awareness is not present at 7 days old, but is by 10 days old.

Schank (2008) J. Theor. Biol.

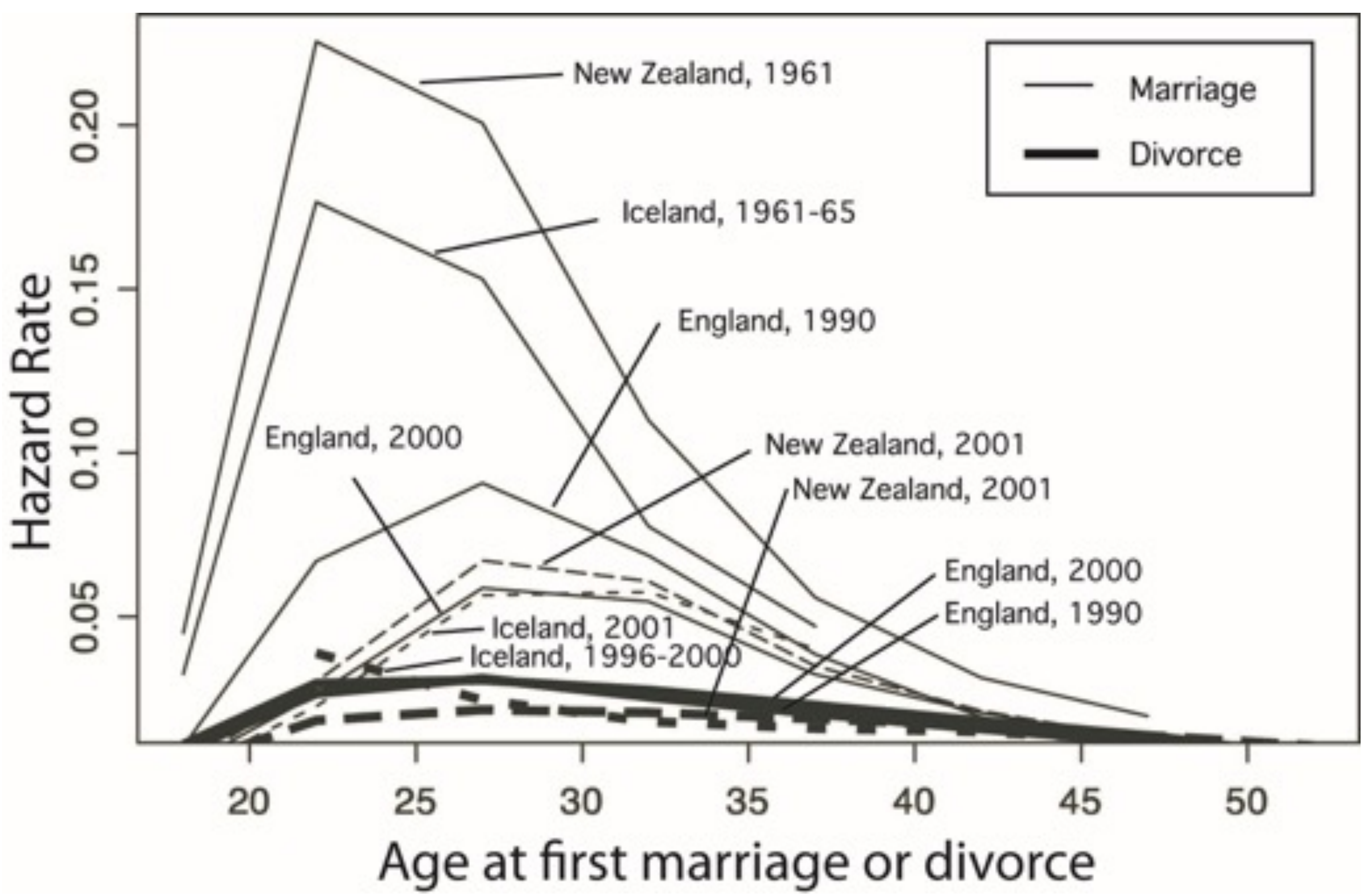


Turchin et al. (2013) PNAS

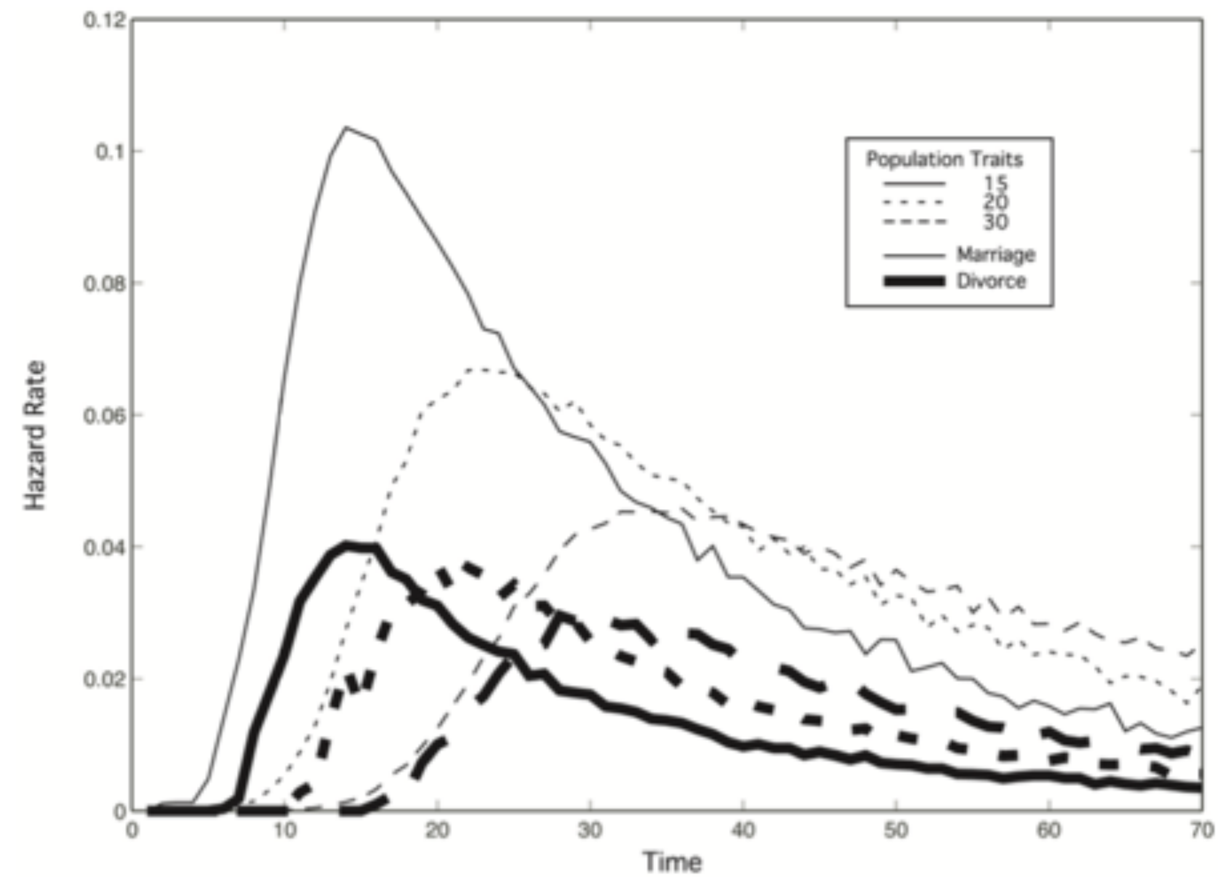
- Epstein: If you didn't generate it, you didn't explain it
- But, if you *did* generate it, you have only generated a *candidate* explanation


ALIENS
HD
HISTORY.COM

Data



Model



Hills & Todd (2008) JASSS

Mate choice model

Male and female agents vary in “attractiveness” on 1-10 scale and have opportunities to form pairs.

Two decision rules:

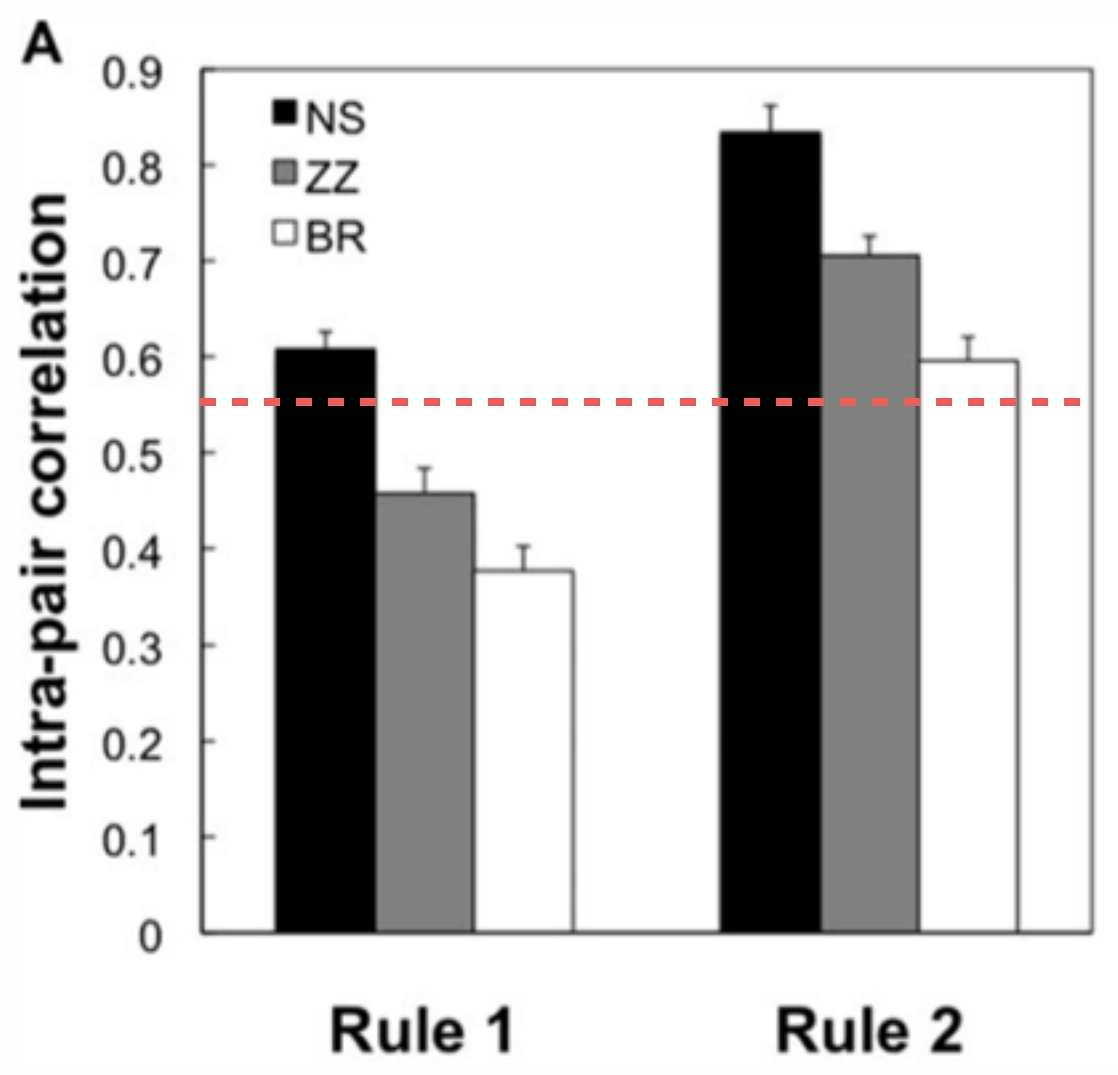
1. Prefer the most attractive
2. Prefer the most similar

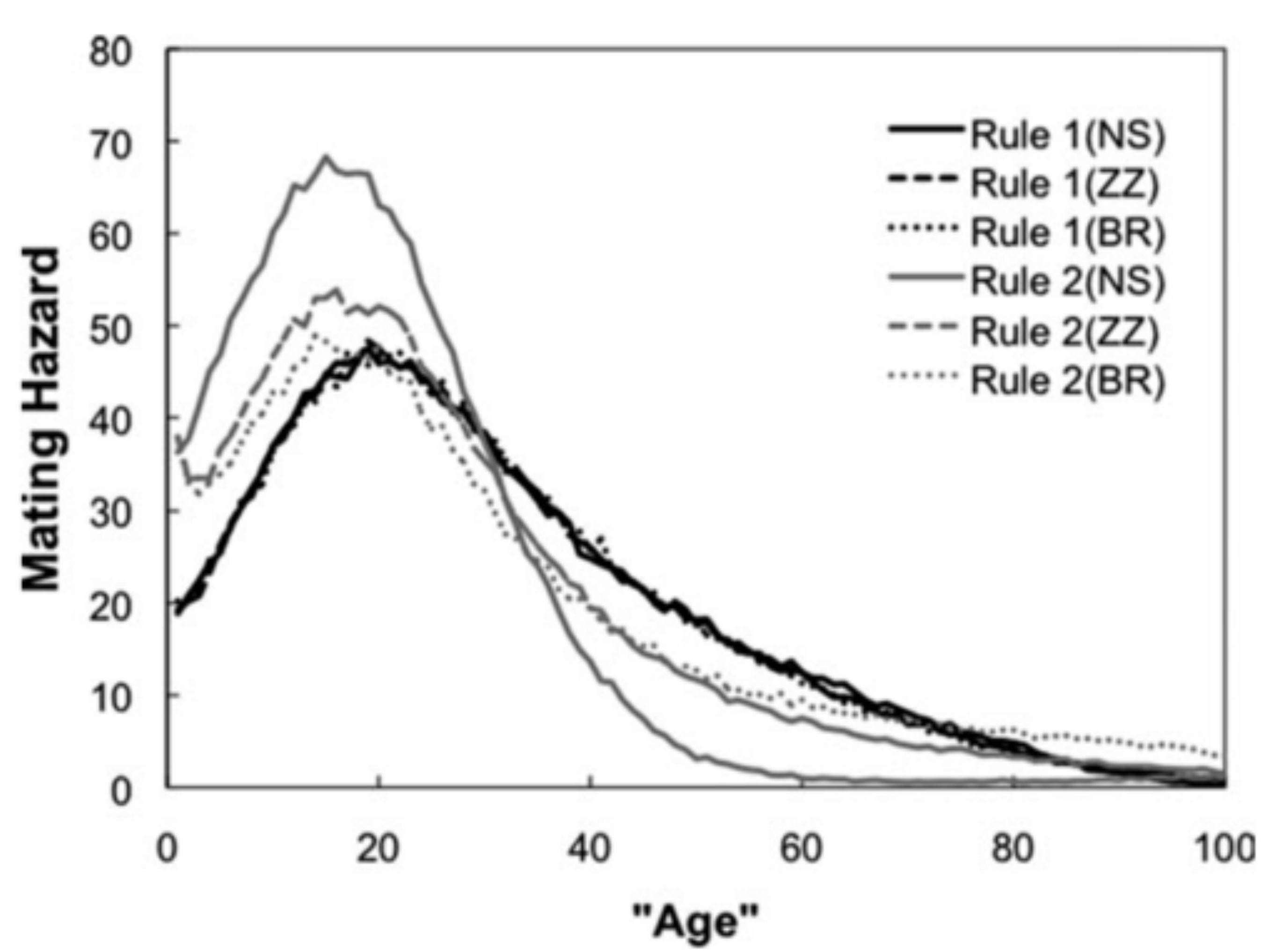
Three movement rules

1. Non-spatial/well-mixed (NS)
2. Zigzag (ZZ): move rapidly through space
3. Brownian (BR): move slowly through space

Smaldino & Schank (2013) *Complexity*

Assumptions about interaction networks can make two very different decision rules each fit the data

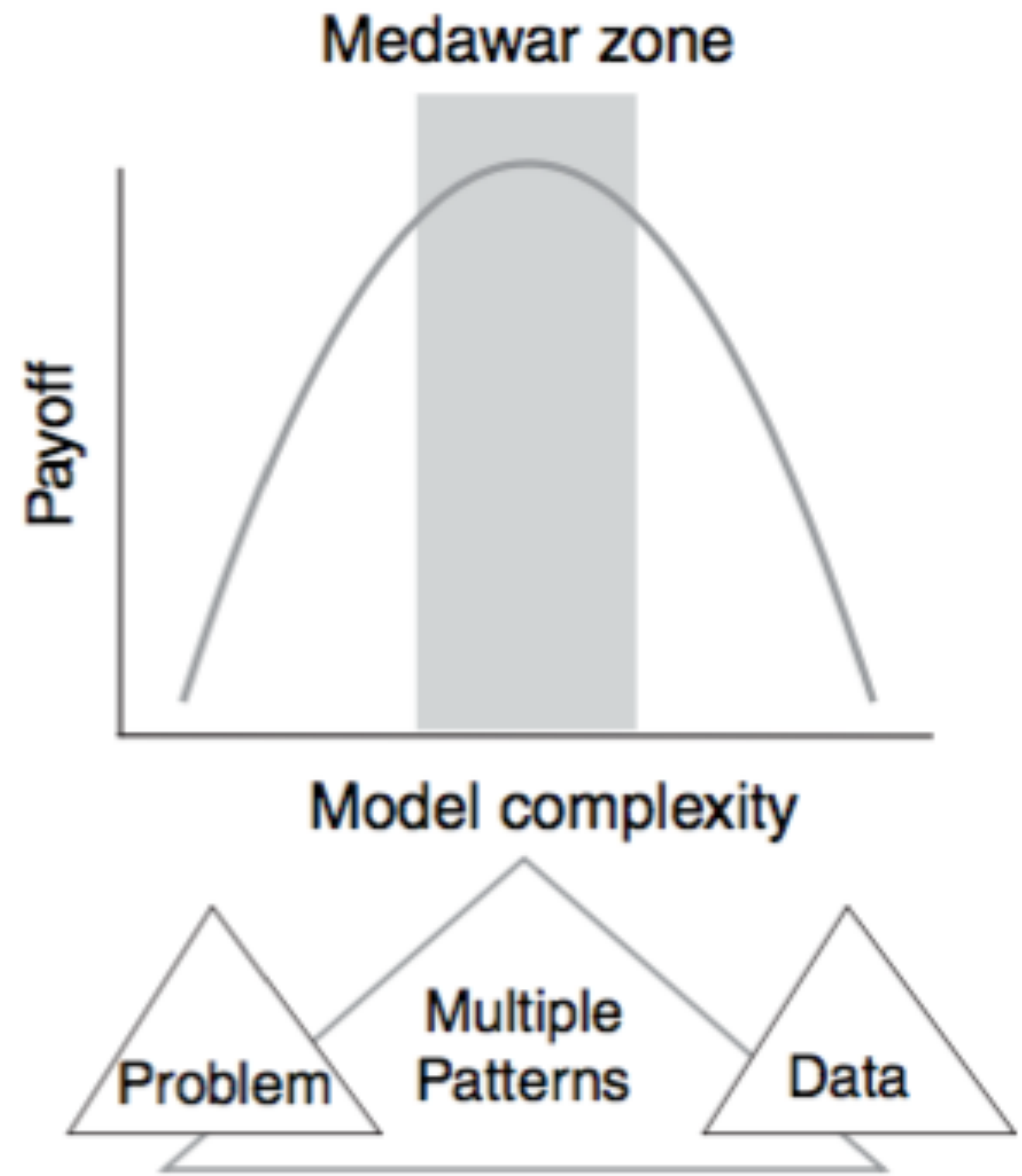




Smaldino & Schank (2013) *Complexity*

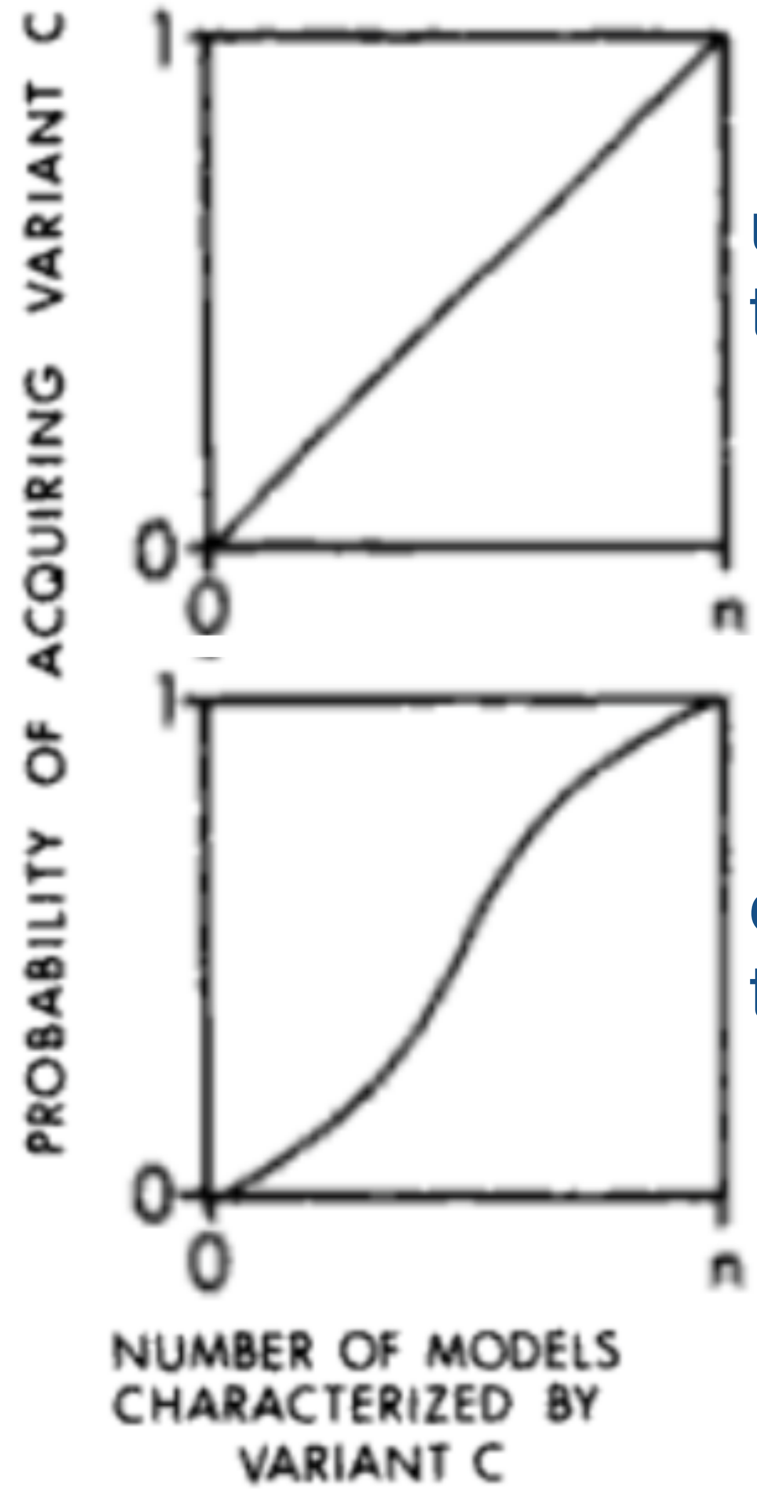
Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology

Volker Grimm,[1*] Eloy Revilla,[2] Uta Berger,[3] Florian Jeltsch,[4] Wolf M. Mooij,[5] Steven F. Railsback,[6]
Hans-Hermann Thulke,[1] Jacob Weiner,[7] Thorsten Wiegand,[1] Donald L. DeAngelis[8]



Model assumptions are important

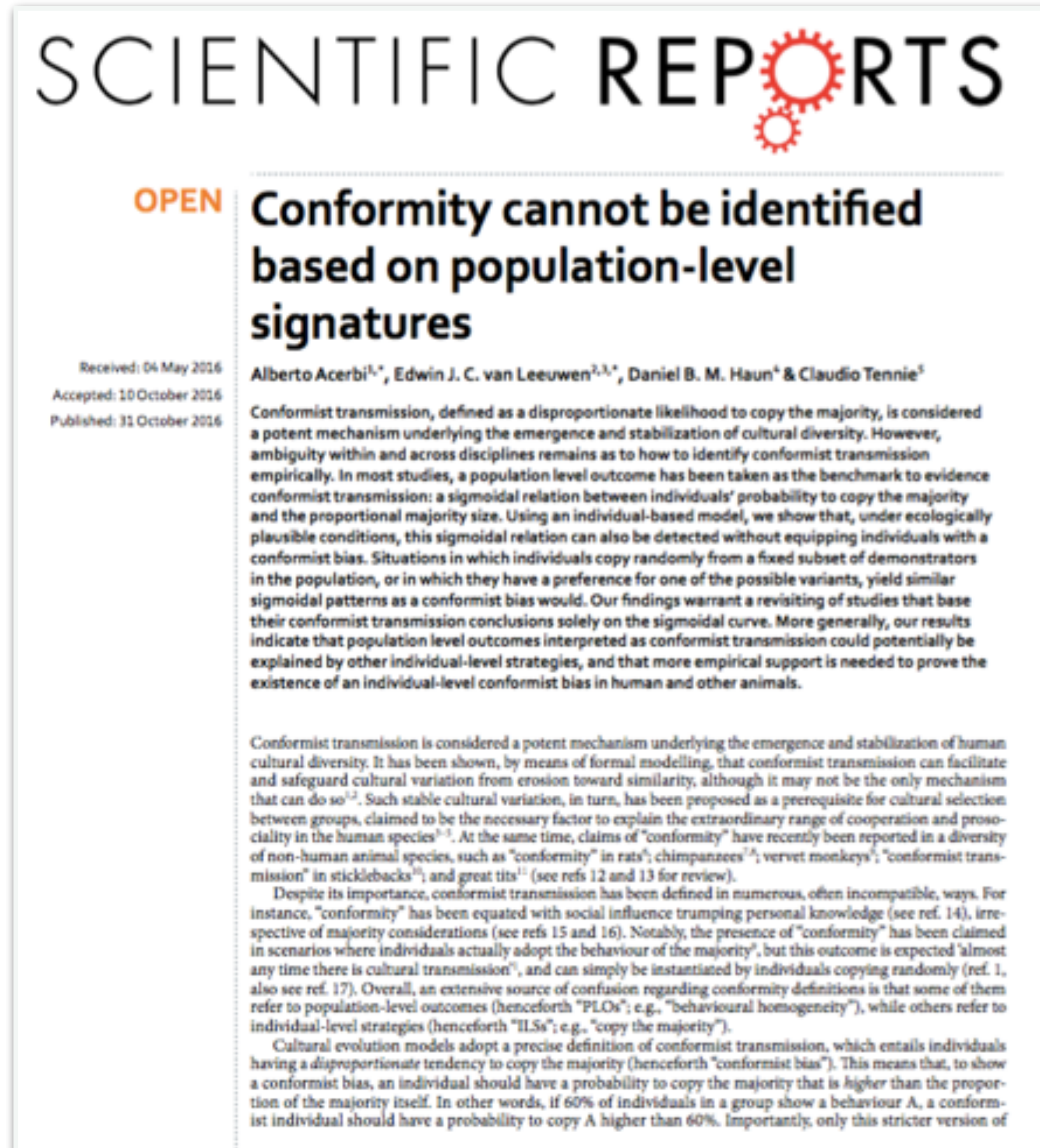
- Conformity: an above-baseline probability of adopting the common behavioral variant



Boyd & Richerson (1985) *Culture and the Evolutionary Process*

Model assumptions are important

- Two variants: A and B
- Initial: 50% each
- In each run, one variant was preferred by all (direct bias)
- Each time step:
  - Each individual paired with randomly chosen demonstrator
  - If demonstrator had preferred variant, copy
  - Else, copy with probability *pLess* = 0.2



SCIENTIFIC REPORTS

OPEN

Conformity cannot be identified based on population-level signatures

Received: 04 May 2016
Accepted: 10 October 2016
Published: 31 October 2016

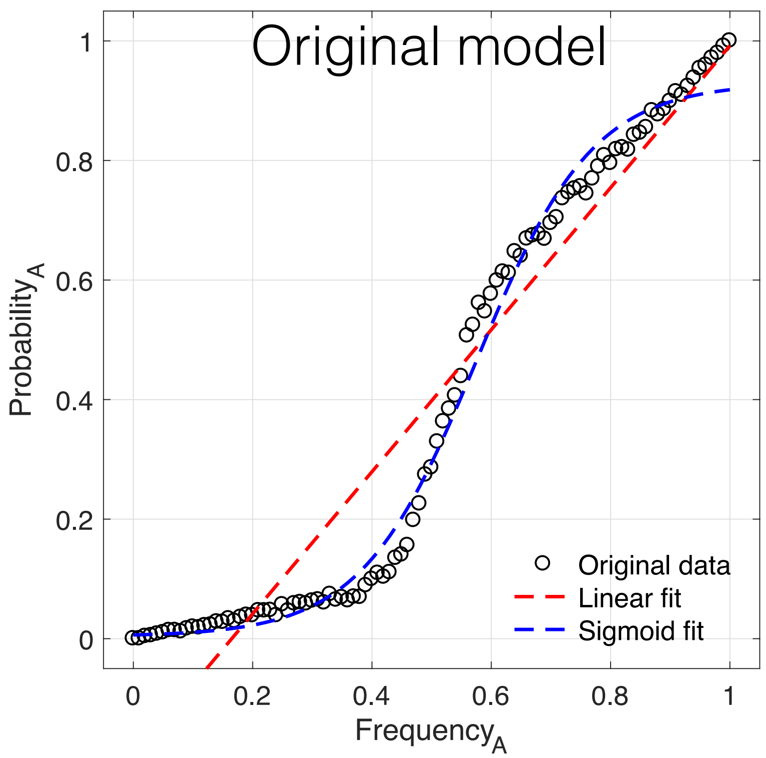
Alberto Acerbi[1,*], Edwin J. C. van Leeuwen[2,3,*], Daniel B. M. Haun[4] & Claudio Tennie[5]

Conformist transmission, defined as a disproportionate likelihood to copy the majority, is considered a potent mechanism underlying the emergence and stabilization of cultural diversity. However, ambiguity within and across disciplines remains as to how to identify conformist transmission empirically. In most studies, a population level outcome has been taken as the benchmark to evidence conformist transmission: a sigmoidal relation between individuals' probability to copy the majority and the proportional majority size. Using an individual-based model, we show that, under ecologically plausible conditions, this sigmoidal relation can also be detected without equipping individuals with a conformist bias. Situations in which individuals copy randomly from a fixed subset of demonstrators in the population, or in which they have a preference for one of the possible variants, yield similar sigmoidal patterns as a conformist bias would. Our findings warrant a revisiting of studies that base their conformist transmission conclusions solely on the sigmoidal curve. More generally, our results indicate that population level outcomes interpreted as conformist transmission could potentially be explained by other individual-level strategies, and that more empirical support is needed to prove the existence of an individual-level conformist bias in human and other animals.

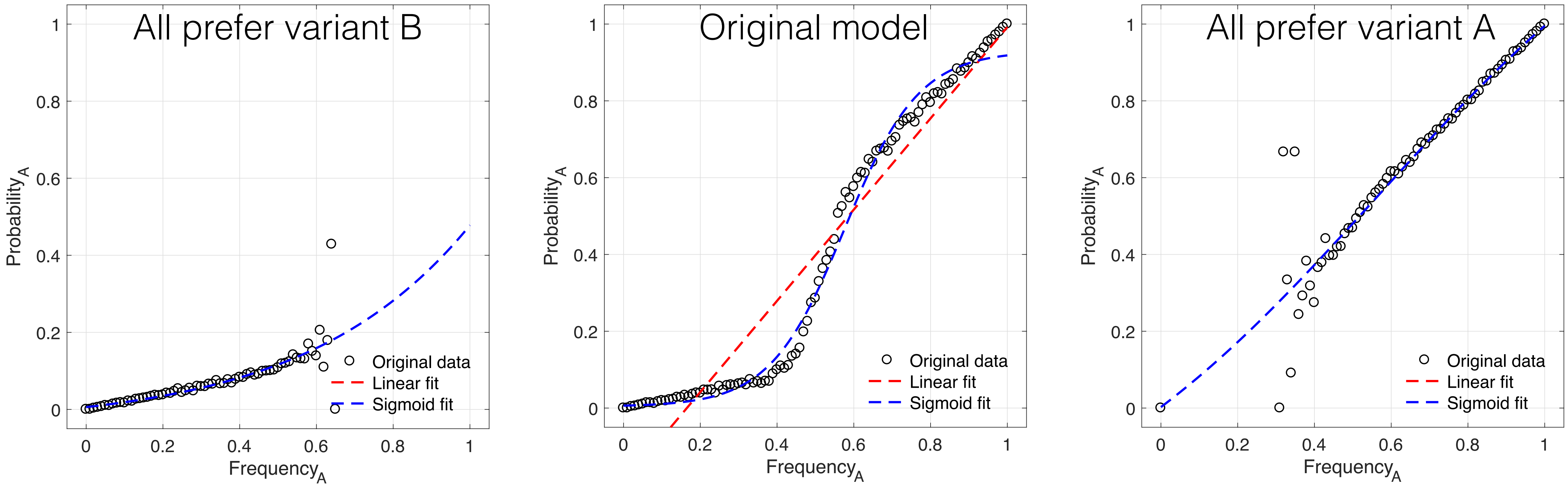
Conformist transmission is considered a potent mechanism underlying the emergence and stabilization of human cultural diversity. It has been shown, by means of formal modelling, that conformist transmission can facilitate and safeguard cultural variation from erosion toward similarity, although it may not be the only mechanism that can do so[1,2]. Such stable cultural variation, in turn, has been proposed as a prerequisite for cultural selection between groups, claimed to be the necessary factor to explain the extraordinary range of cooperation and prosociality in the human species[3–5]. At the same time, claims of "conformity" have recently been reported in a diversity of non-human animal species, such as "conformity" in rats[6]; chimpanzees[7,8]; vervet monkeys[9]; "conformist transmission" in sticklebacks[10]; and great tits[11] (see refs 12 and 13 for review).

Despite its importance, conformist transmission has been defined in numerous, often incompatible, ways. For instance, "conformity" has been equated with social influence trumping personal knowledge (see ref. 14), irrespective of majority considerations (see refs 15 and 16). Notably, the presence of "conformity" has been claimed in scenarios where individuals actually adopt the behaviour of the majority", but this outcome is expected 'almost any time there is cultural transmission'", and can simply be instantiated by individuals copying randomly (ref. 1, also see ref. 17). Overall, an extensive source of confusion regarding conformity definitions is that some of them refer to population-level outcomes (henceforth "PLOs"; e.g., "behavioural homogeneity"), while others refer to individual-level strategies (henceforth "ILSs"; e.g., "copy the majority").

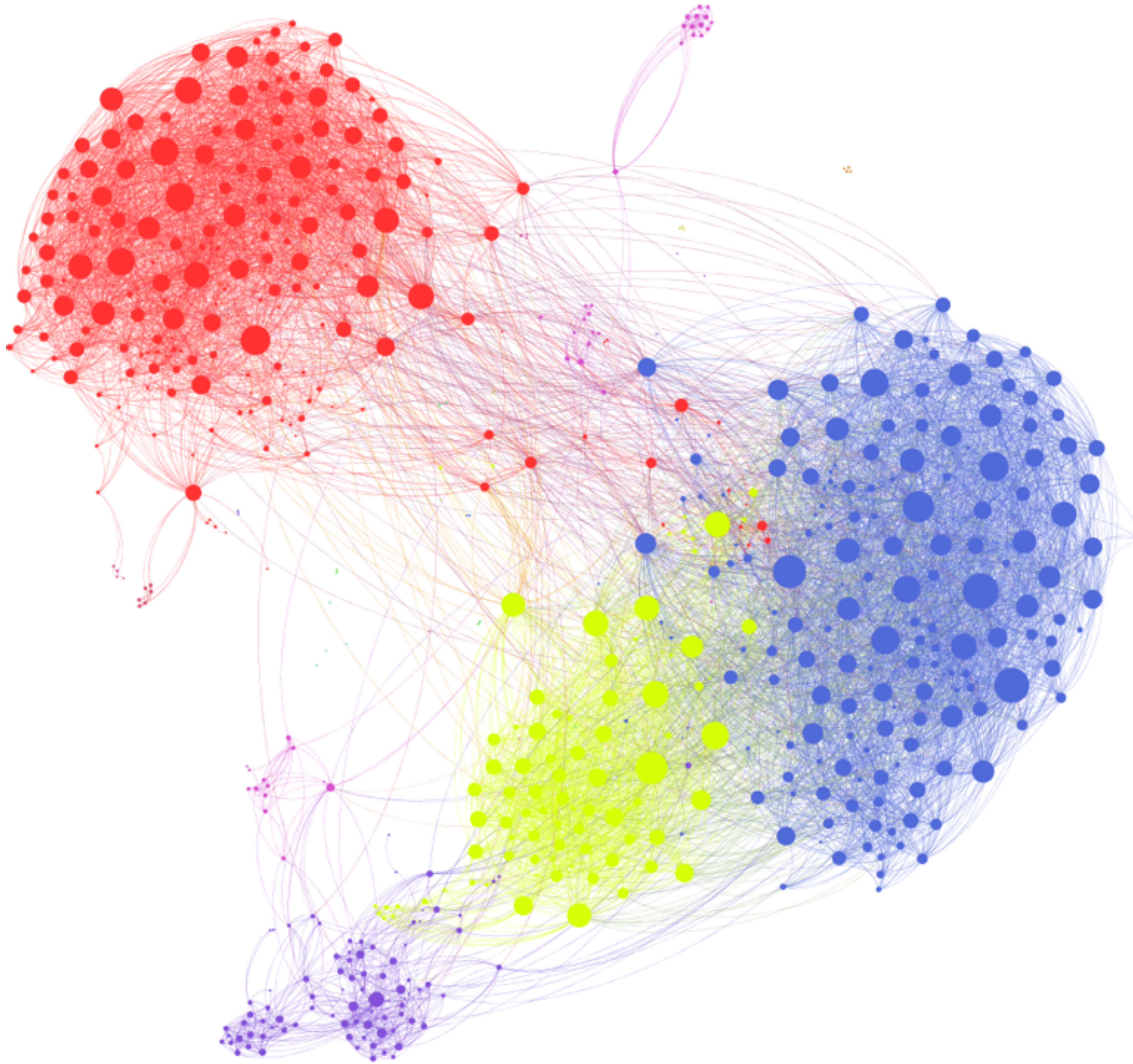
Cultural evolution models adopt a precise definition of conformist transmission, which entails individuals having a *disproportionate* tendency to copy the majority (henceforth "conformist bias"). This means that, to show a conformist bias, an individual should have a probability to copy the majority that is *higher* than the proportion of the majority itself. In other words, if 60% of individuals in a group show a behaviour A, a conformist individual should have a probability to copy A higher than 60%. Importantly, only this stricter version of


Original model
Probability$_A$
Frequency$_A$
0
0.2
0.4
0.6
0.8
1
Original data
Linear fit
Sigmoid fit

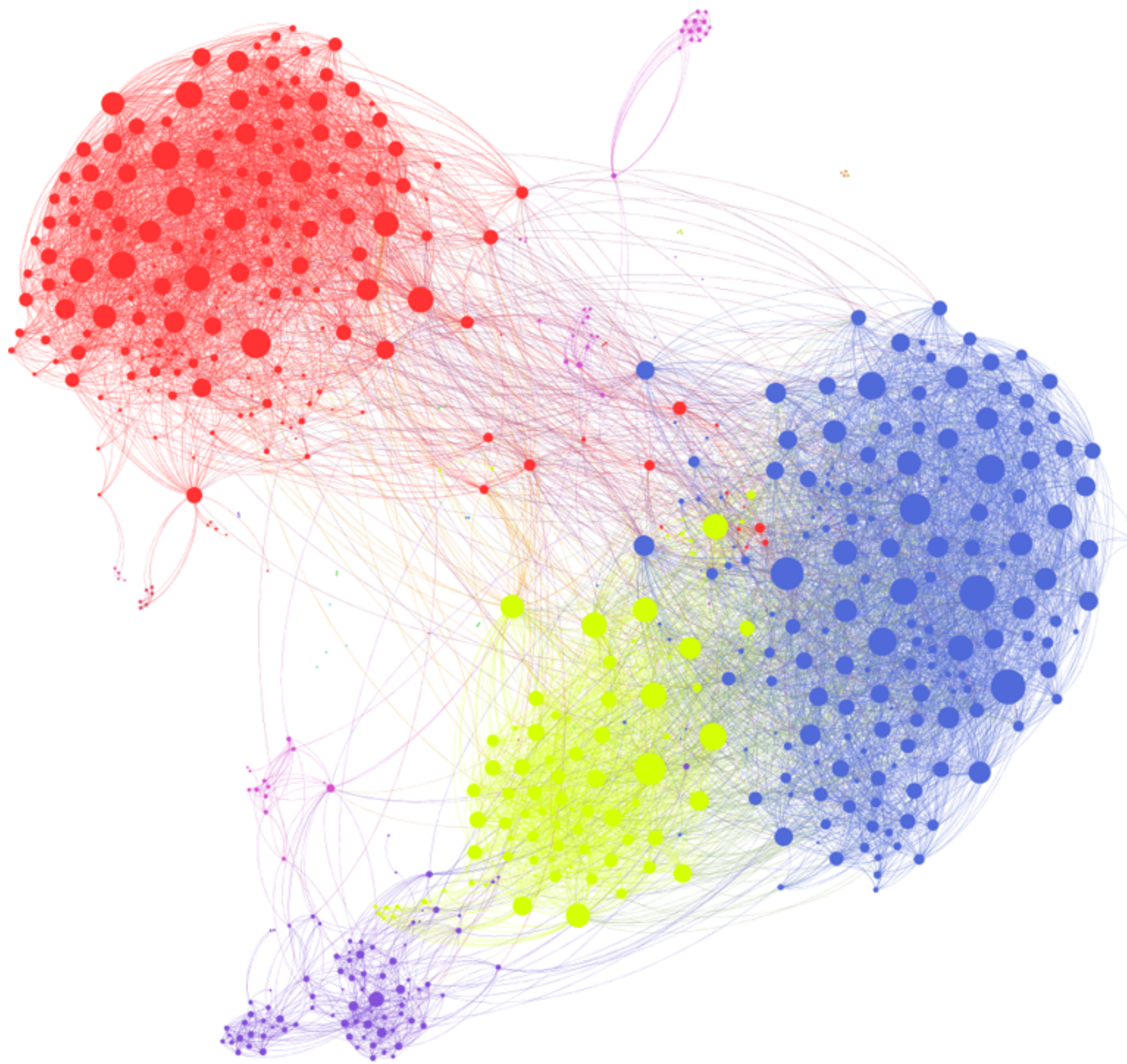
Initial frequency$_A$ always 50%



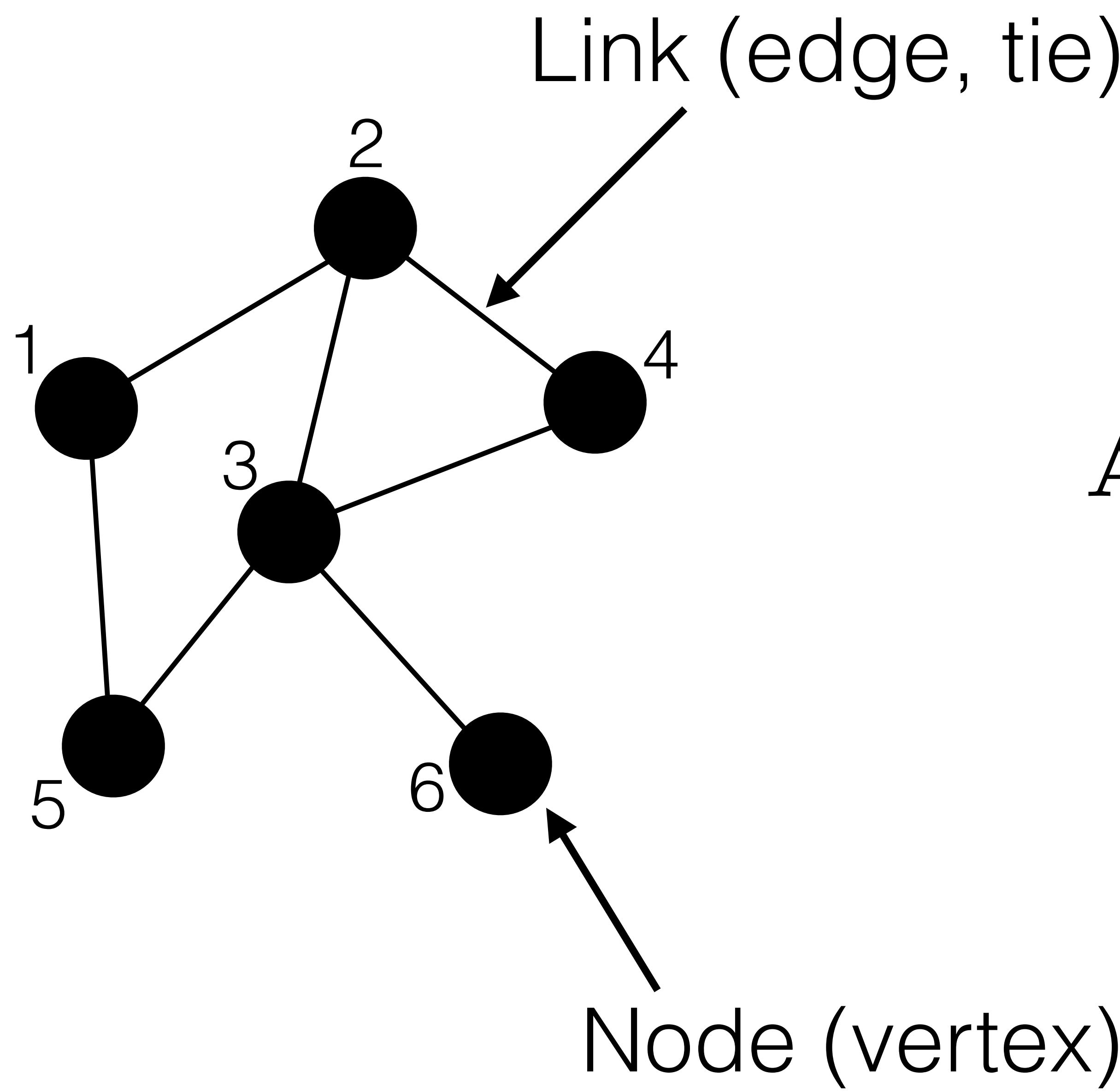
Smaldino, Aplin, & Farine (2017) bioRxiv:159038



Networks



What is a network?

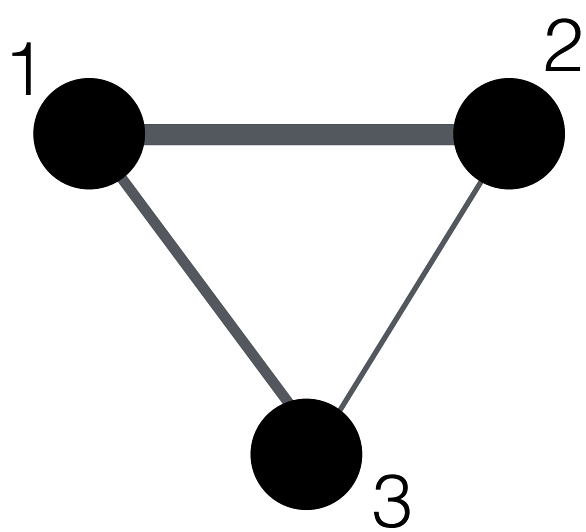


Adjacency matrix:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

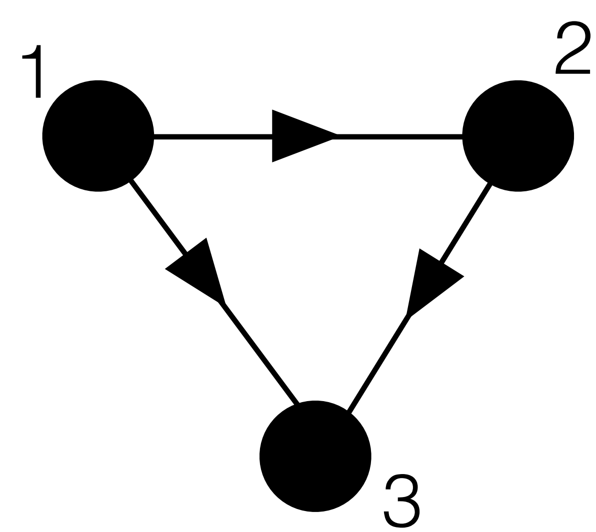
Weighted networks

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 0.5 \\ 1 & 0.5 & 0 \end{pmatrix}$$

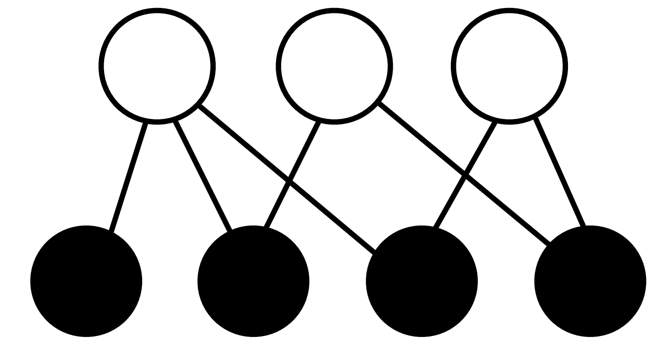


Directed networks

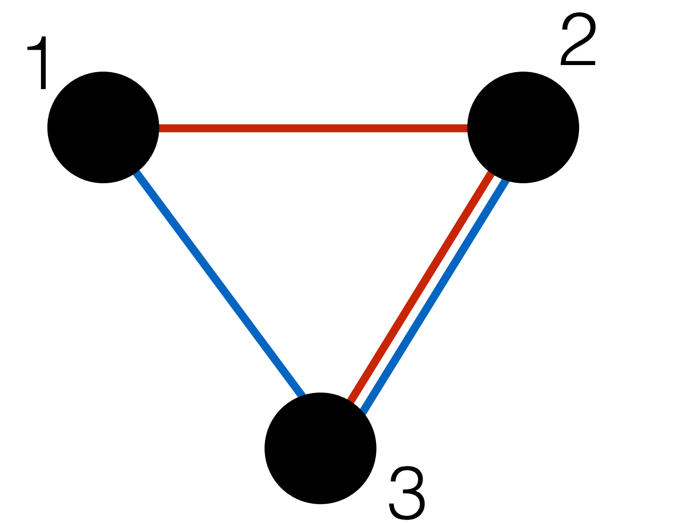
$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$



Bipartite networks

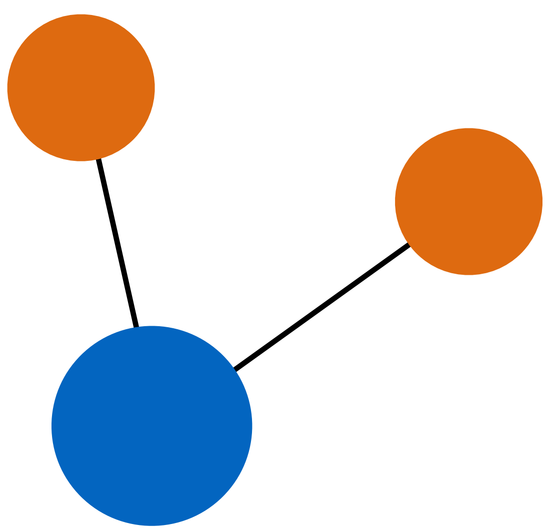


Multiplex networks



Degree and density

- Which are the most important or central nodes in a network?



degree: $$k_i = \sum_{j=1}^{n} A_{ij}$$

(also called 'degree centrality')

The **density** of a network is the proportion of possible edges that actually exist.

Eigenvector Centrality

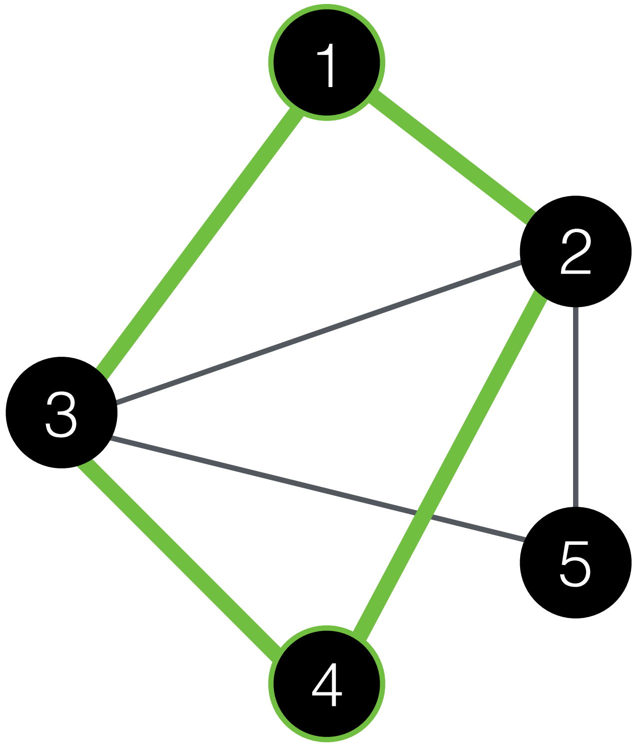
- Give more weight to edges that connect to highly-connected nodes
- Requires computing the eigenvectors of the adjacency matrix (requires linear algebra)
- **Google's** PageRank algorithm is a variant of this

Paths

- A **path** between two nodes is any sequence of non-repeating connected nodes that connects the two nodes

- The **shortest path** between two nodes is one that connects the two nodes with the smallest number of edges (also called the **distance** between the nodes)

- The **average path length** is the average distance between all pairs of nodes in a network



Euler and the Seven Bridges of Königsberg

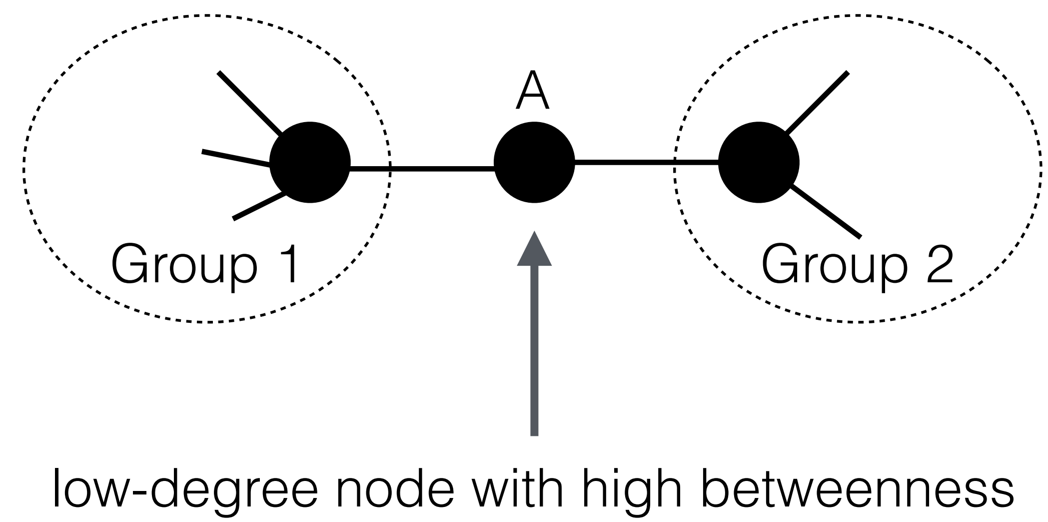
Is there any walking route that crosses all seven bridges exactly once?

Betweenness Centrality

- Extent to which a node lies on paths between other nodes

- Let n^i_{st} be 1 if node i lies on the shortest path from node s to node t, and 0 if it doesn't (or if there is no such path). The betweenness centrality of node i is:

$$x_i = \sum_{st} n^i_{st}$$



Closeness Centrality

- Based on mean distance from a node to other nodes.
- Take reciprocal so higher values indicate higher closeness

Mean distance from node *i* to all other nodes

$$\ell_i = \frac{1}{n}\sum_j d_{ij}$$

Closeness centrality:

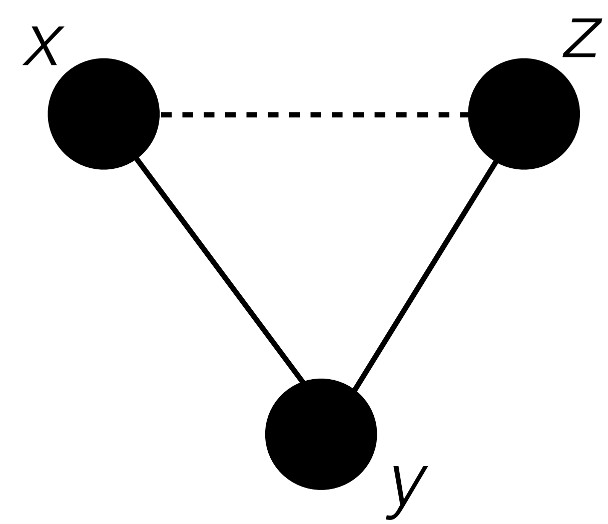
$$C_i = \frac{1}{\ell_i} = \frac{n}{\sum_j d_{ij}}$$

Interpretation of centrality measures

Centrality measure	Interpretation in social networks
Degree	How many people can this person reach directly?
Eigenvector	How well is this person connected to other well-connected people?
Betweenness	How likely is this person likely to be the most direct route between two people in the network?
Closeness	How fast can this person reach everyone in the network?

Transitivity and Clustering

- How predictive is the fact that A is friends with B and C of whether B and C are also friends?



The path *xyz* is **closed** if the third edge from *z* to *x* is present.

Clustering coefficient:

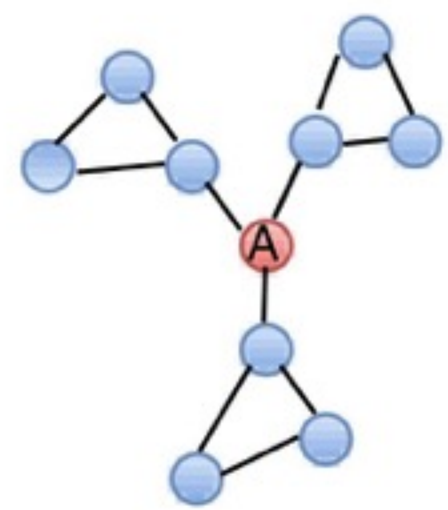
$$C = \frac{(\text{number of triangles}) \times 3}{(\text{number of connected triples})}$$

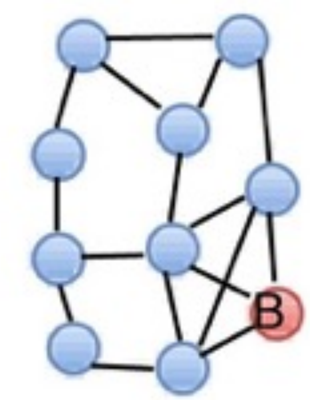
Local Clustering

Local clustering for node *i*:

$$C_i = \frac{\text{(number of pairs of neighbors of } i \text{ that are connected)}}{\text{(number of pairs of neighbors of } i\text{)}}$$

- Similar to betweenness centrality
- Can be used to probe for **structural holes**
- Watts-Strogatz “Average clustering” coefficient:



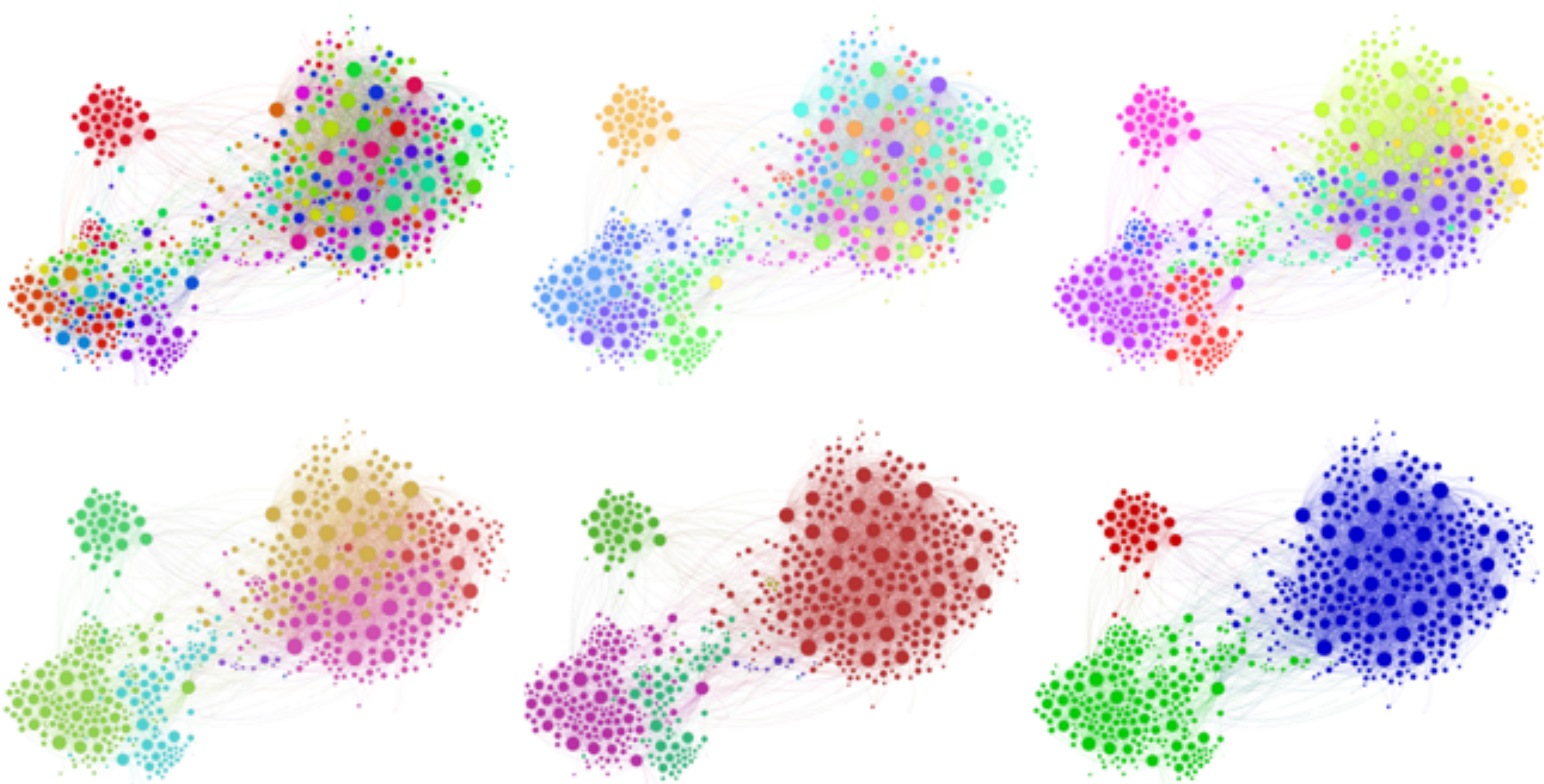


$$C_{WS} = \frac{1}{n}\sum_{i=1}^{n} C_i$$

Community Detection

- Separating the network into groups of nodes that are highly connected within groups and sparsely connected between groups.

- Several algorithms exist, each with their own pros and cons.



Interaction Models on Networks

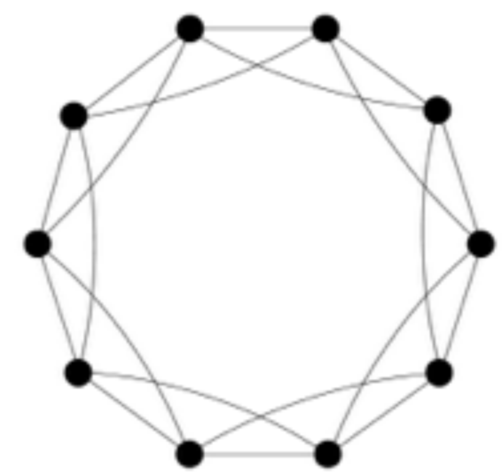
- Epidemics
- Diffusion of innovations or information
- Evolutionary games
- Economic transactions
- Food webs

Models of Network Architectures

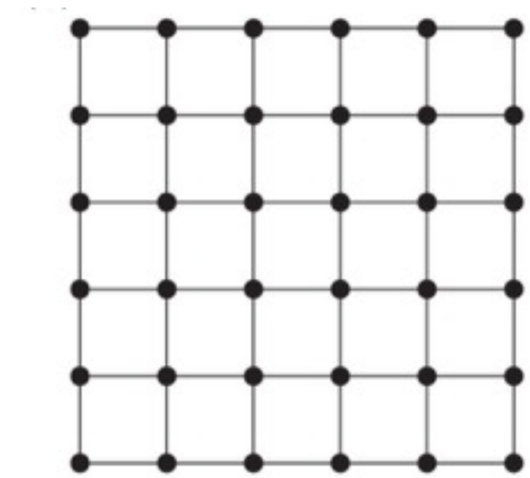
- Regular lattices
- Random networks
- Small-world networks
- Scale-free networks

Lattices

- Characterized by regular structure
- Easy to model computationally
- Sometimes possible to solve analytically
- Questionable realism

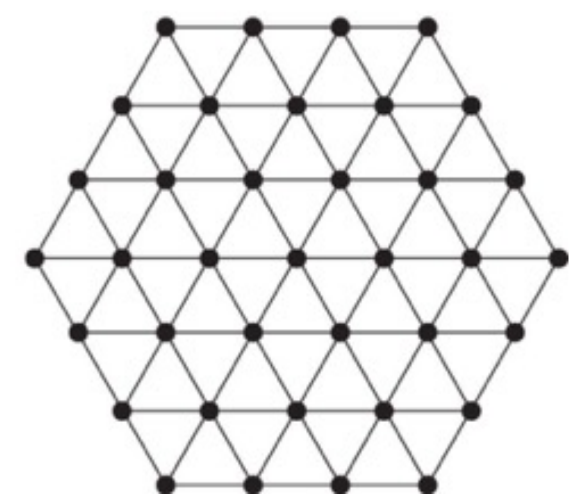


Ring lattice



Square lattice

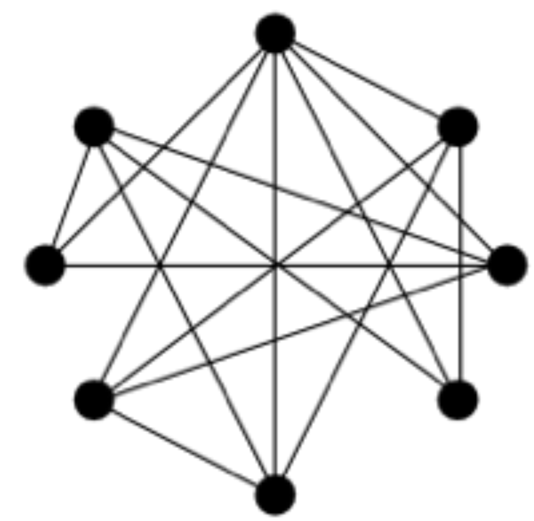
Triangular lattice

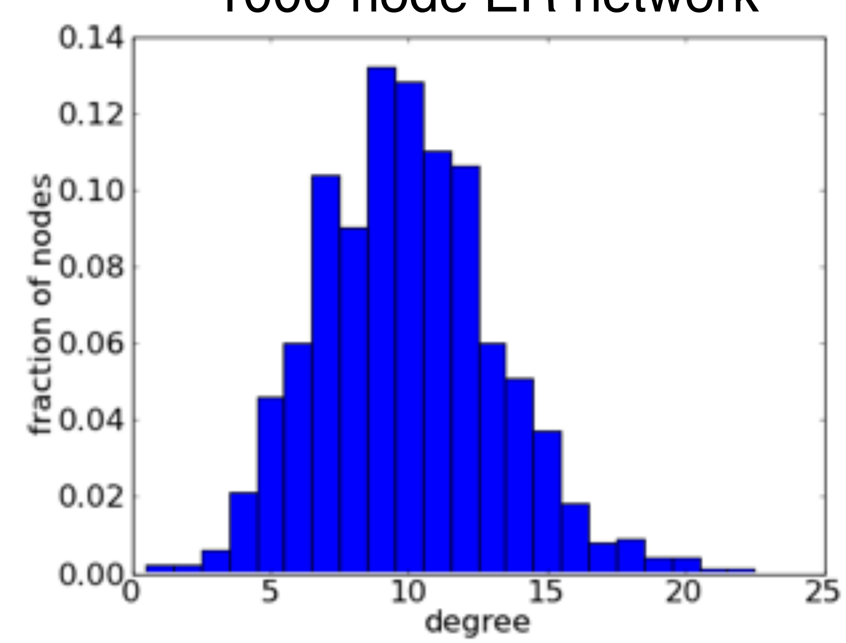


Hexagonal lattice

Random Networks

- Introduced by Erdös & Renyi (1959)
- Minimal assumption for a connected population
- Multiple network formation algorithms exist. Example: *N* nodes are specified, and each possible edge is added with a fixed probability
- Average degree is predictable, but degree varies between nodes
- Probably not realistic for many systems

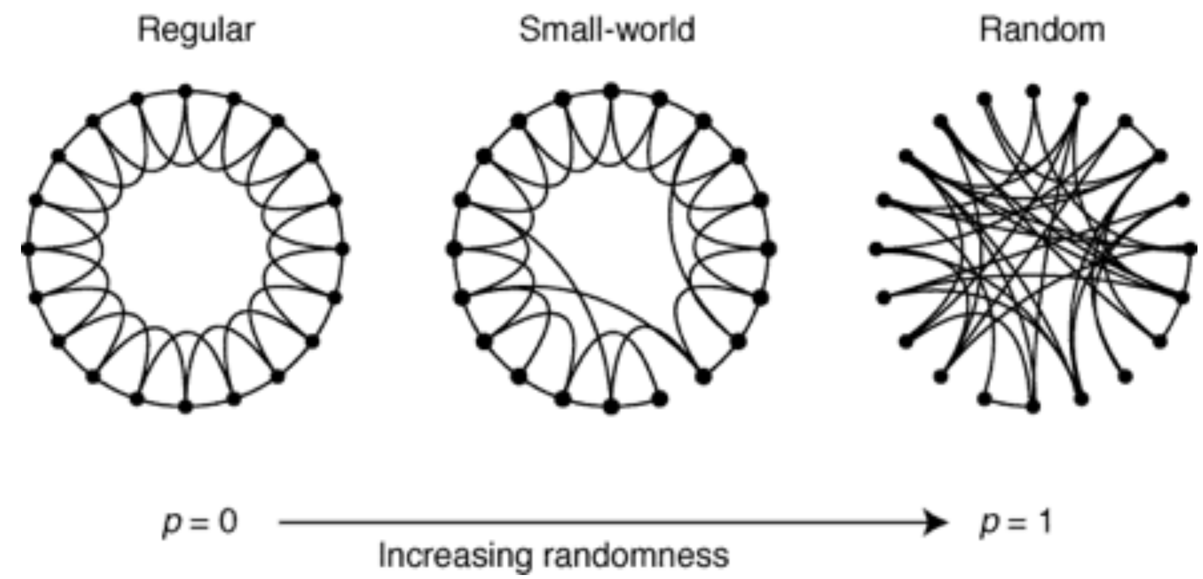


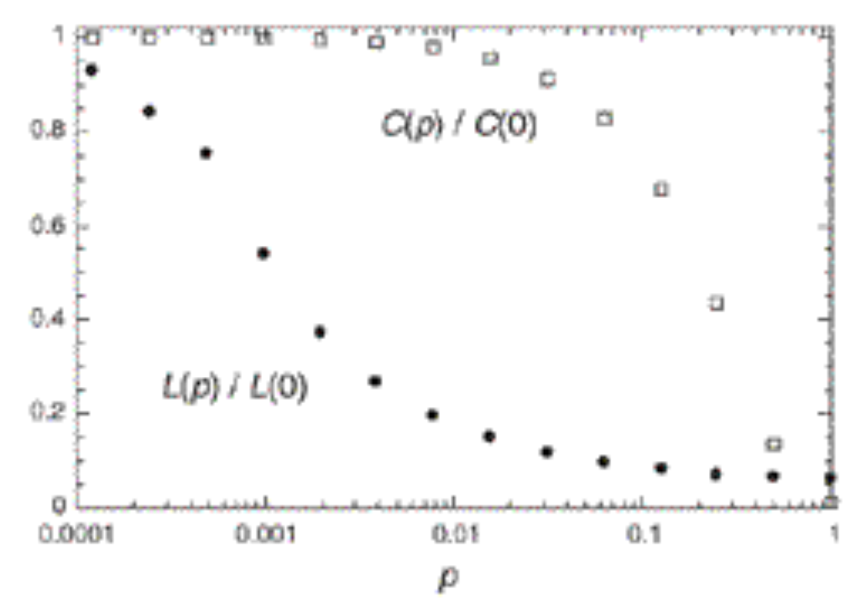


Small-world networks

- Introduced by Watts and Strogatz (1998)
- Characterized by high clustering (like lattices) and short path lengths (like random networks)
- Many real world networks share this property:
  - Film actors (IMDB)
  - Power grid nodes and high-voltage transmission lines in Western US
  - Neural network of *C. elegans*
- Fat-tailed degree distribution: overabundance of hubs

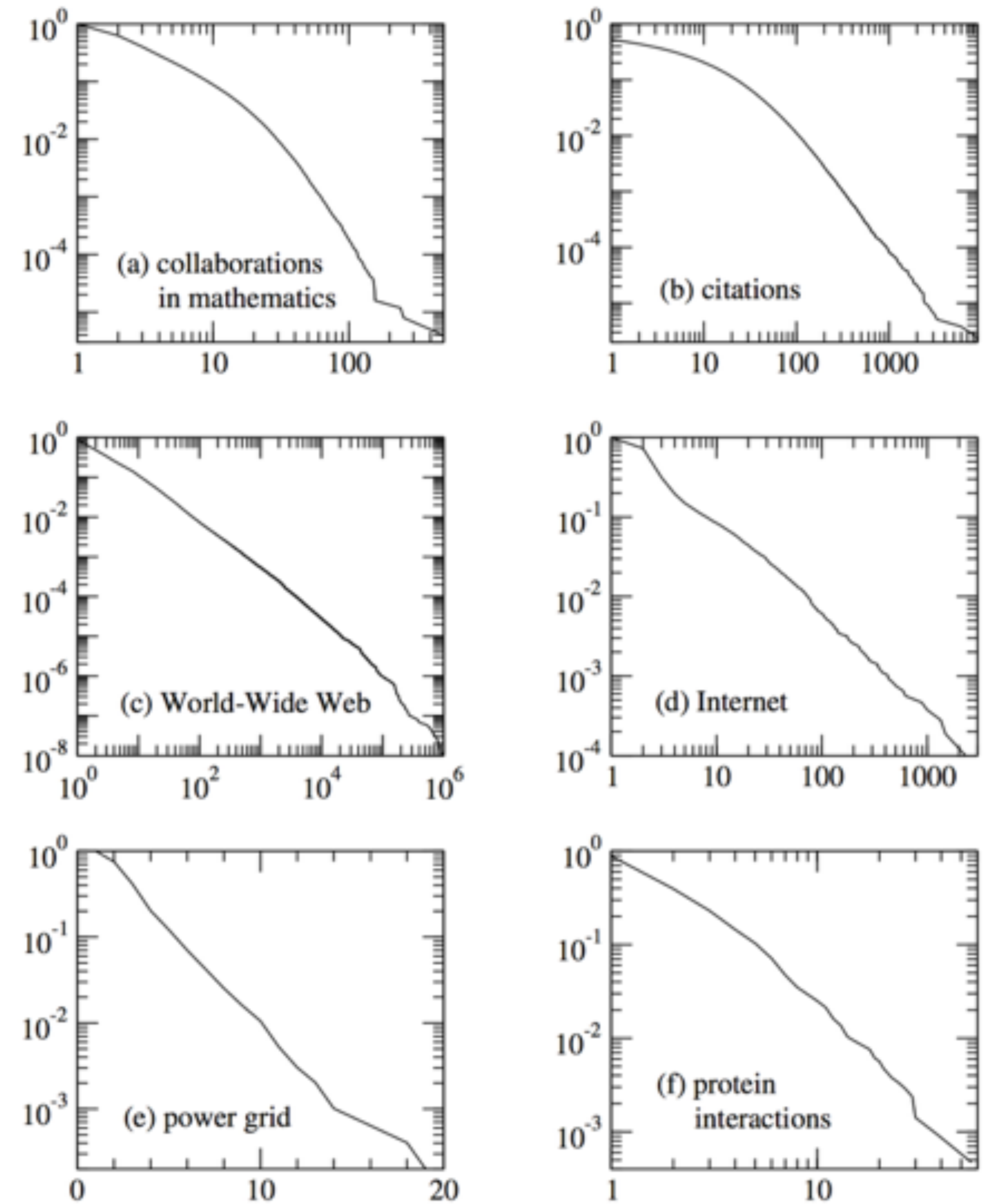
p = probability of rewiring edge





Scale-free networks

- Scale-free: Parts of the network exhibit similar features as the whole network
- Many real-world networks exhibit **power-law degree distributions**
- Few high-degree nodes (hubs), many low-degree nodes



Newman (2003) *SIAM*

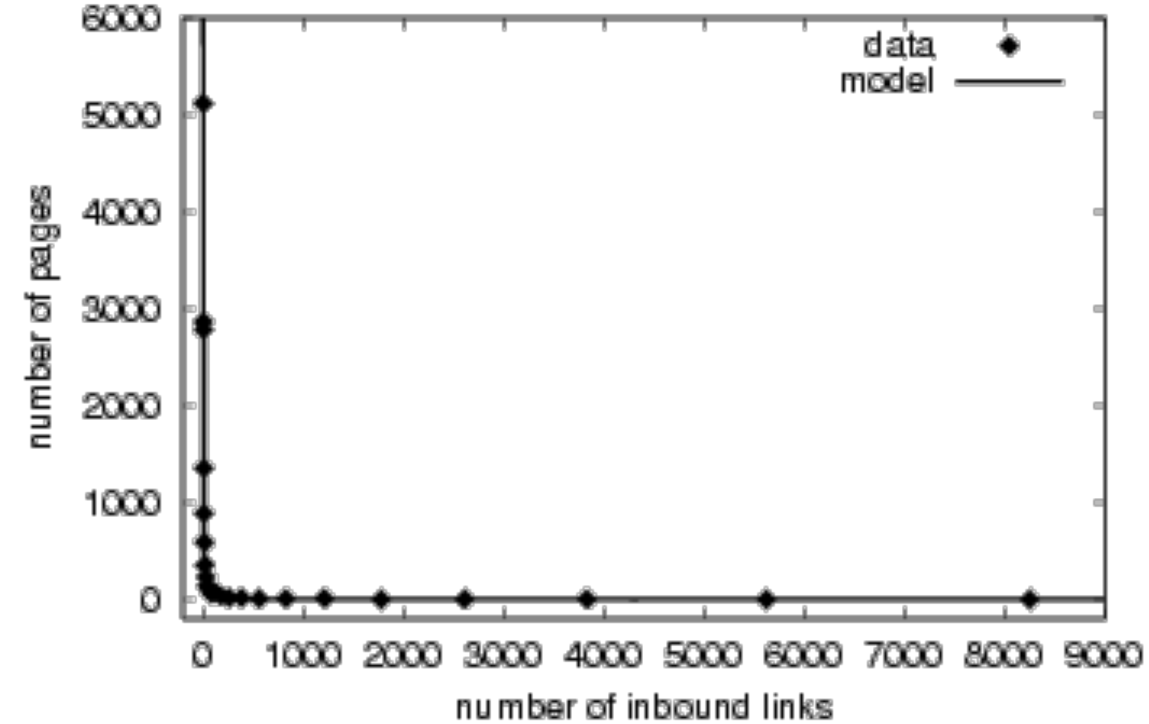
Why do power laws exhibit as straight lines on log-log plots?

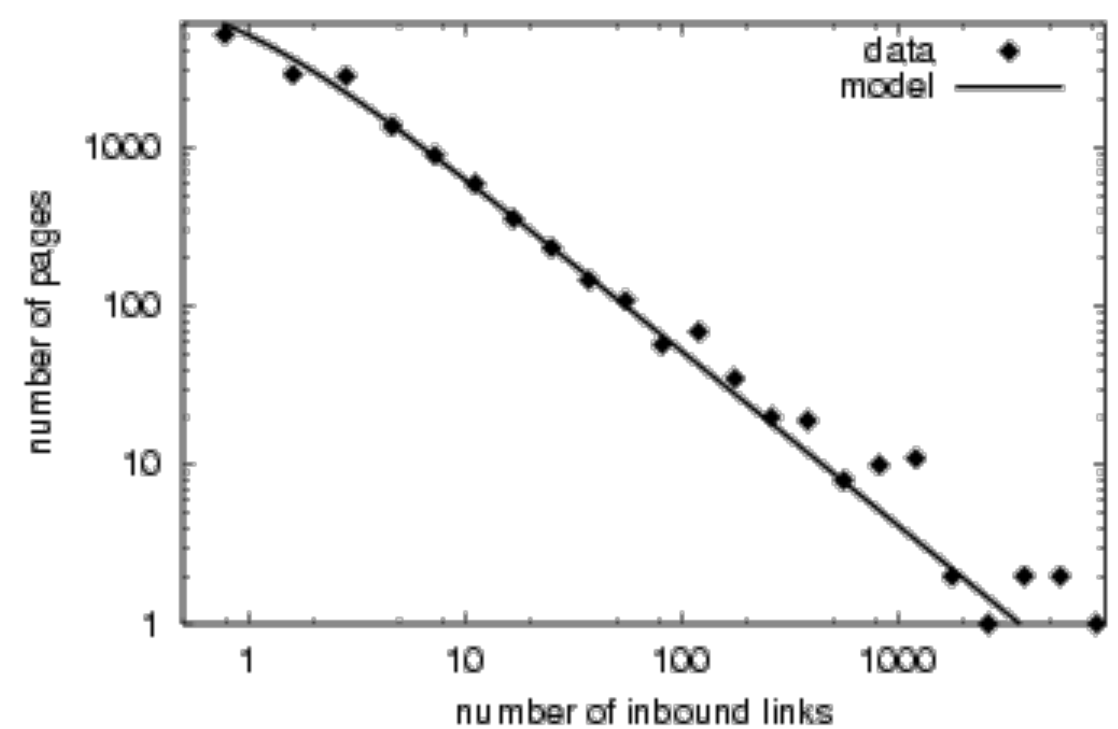
$$y = ax^{-k}$$

$$\log y = \log(ax^{-k})$$

$$= \log a + \log(x^{-k})$$

$$= \log a - k \log x$$

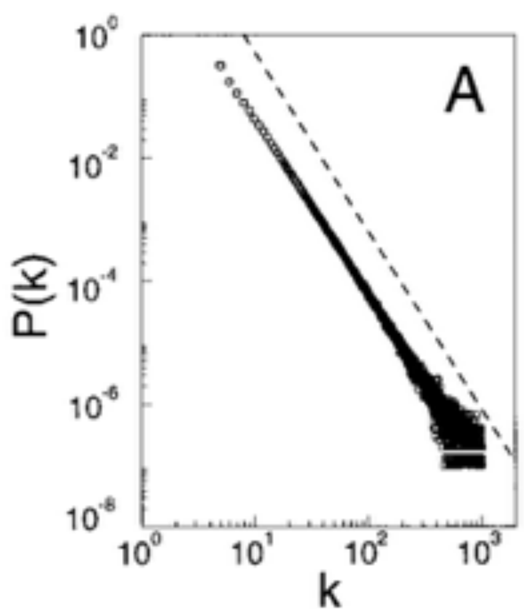


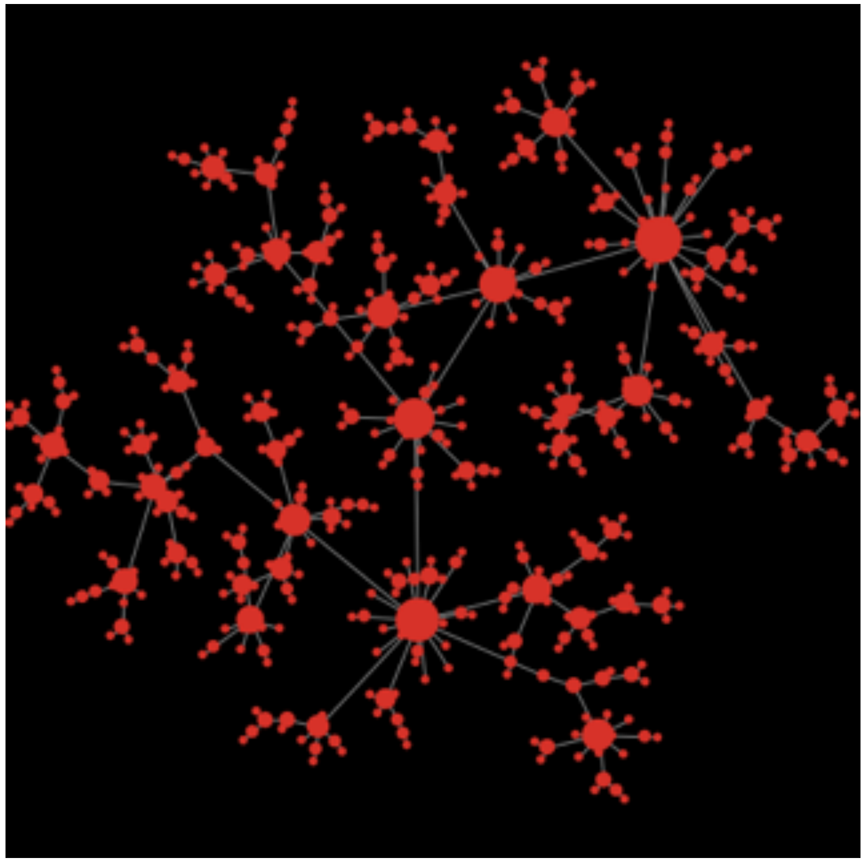


Pennock et al. (2002) *PNAS*

Preferential Attachment Algorithm

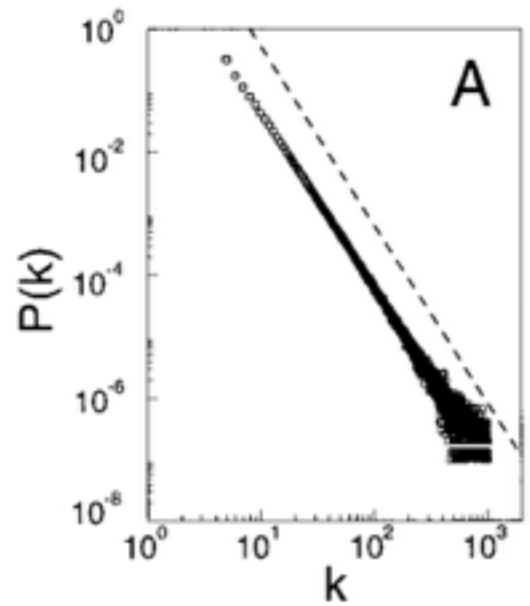
- Barabási & Albert (1999)
- Nodes are added sequentially
- Connectivity is not uniformly random, but preferential
- **Model**
  - Start with m_0 nodes
  - Each time step, add a new node with m edges, that link to m existing nodes with a probability proportionate to the current degree of those nodes (relative to all other nodes)
  - "The rich get richer"

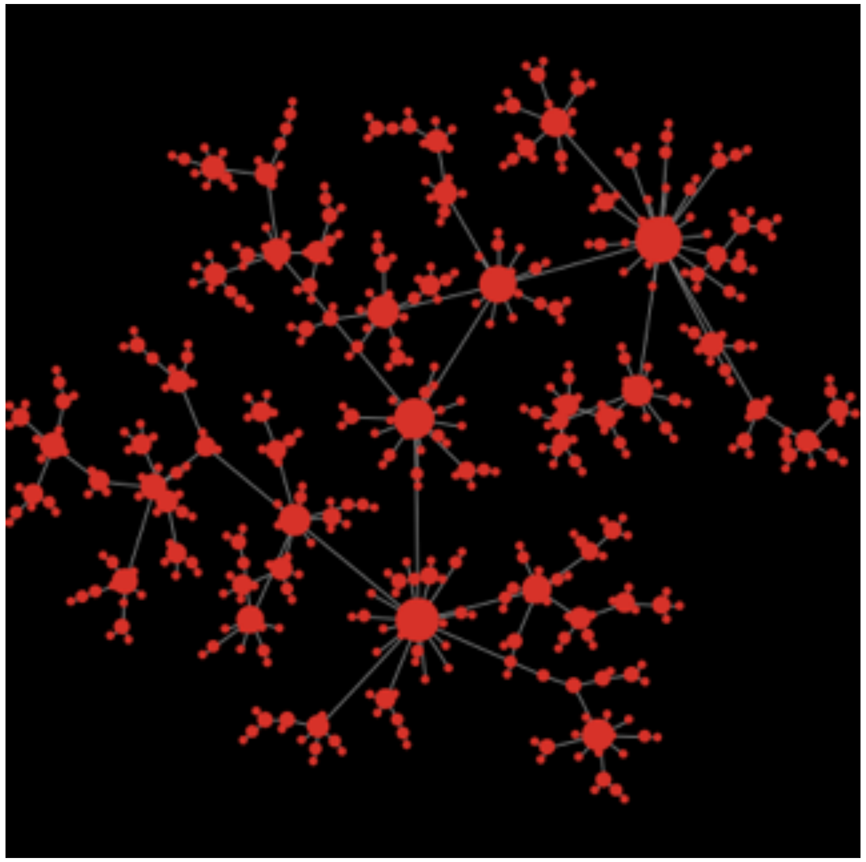




Preferential Attachment Algorithm

- Barabási & Albert (1999)
- Nodes are added sequentially
- Connectivity is not uniformly random, but preferential
- **Model**
  - Start with m_0 nodes
  - Each time step, add a new node with *m* edges, that link to *m* existing nodes with a probability proportionate to the current degree of those nodes (relative to all other nodes)
  - "The rich get richer"





```
to setup
  clear-all
  set-default-shape turtles "circle"
  ;; make the initial network of two turtles and an edge
  make-node nobody        ;; first node, unattached
  make-node turtle 0      ;; second node, attached to first node
  reset-ticks
end

;;;;;;;;;;;;;;;;;;;;;;;
;;; Main Procedures ;;;
;;;;;;;;;;;;;;;;;;;;;;;

to go
  ;; new edge is green, old edges are gray
  ask links [ set color gray ]
  make-node find-partner         ;; find partner & use it as attachment
                                 ;; point for new node
  tick
  if layout? [ layout ]
end

;; used for creating a new node
to make-node [old-node]
  create-turtles 1
  [
    set color red
    if old-node != nobody
      [ create-link-with old-node [ set color green ]
        ;; position the new node near its partner
        move-to old-node
        fd 8
      ]
  ]
end

;; This code is the heart of the "preferential attachment" mechanism, and acts like
to-report find-partner
  report [one-of both-ends] of one-of links
end
```

Why model?

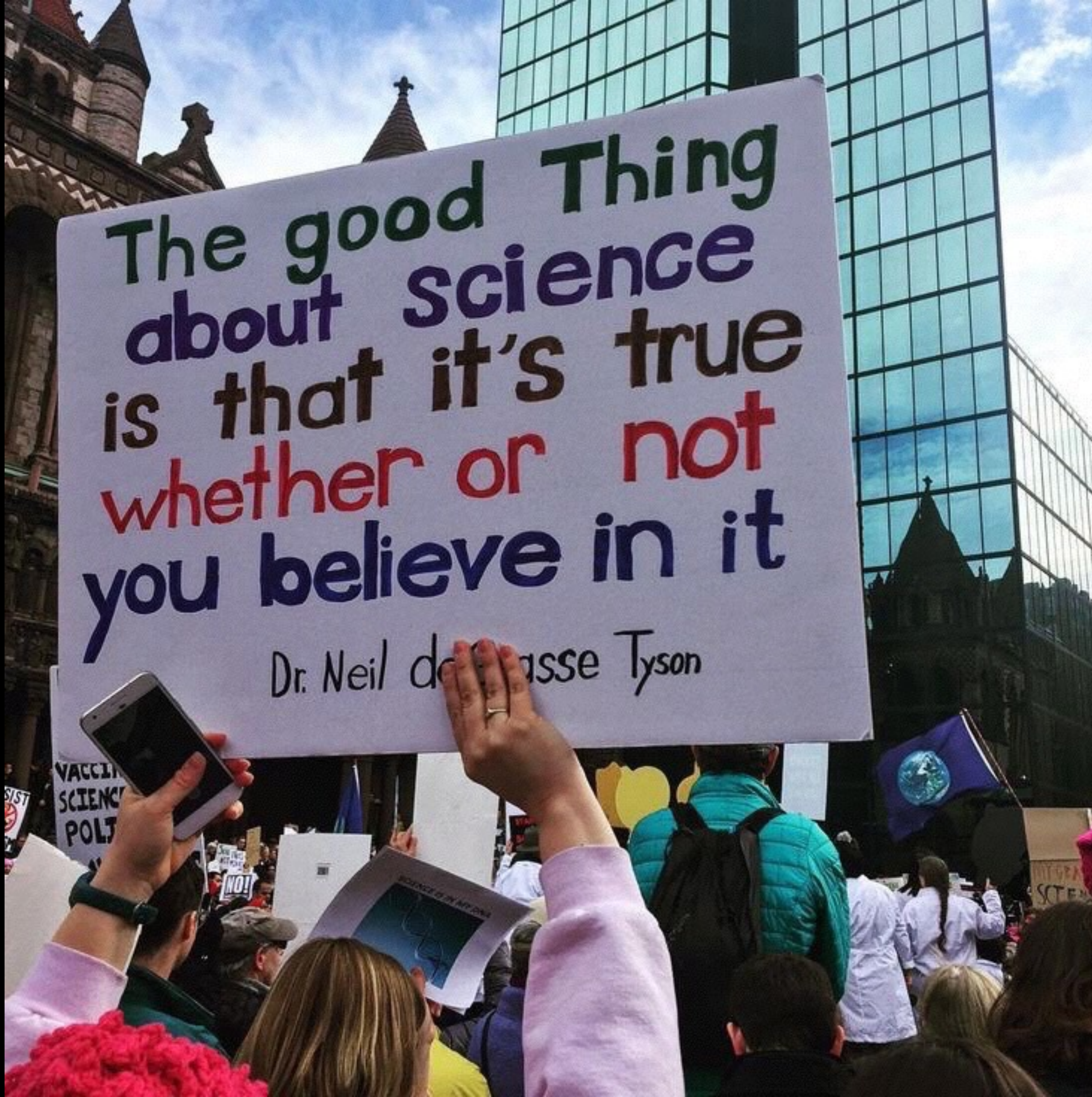
- Models formalize and scaffold theory development
- Good theory structures the interpretation of data
- Good theory leads to better hypothesis formation

Published in: *Theory & Psychology, 8* (2), 1998, 195–204. http://tap.sagepub.com/
© 1998 Sage Publications.

Surrogates for Theories

Gerd Gigerenzer
Max Planck Institute for Human Development

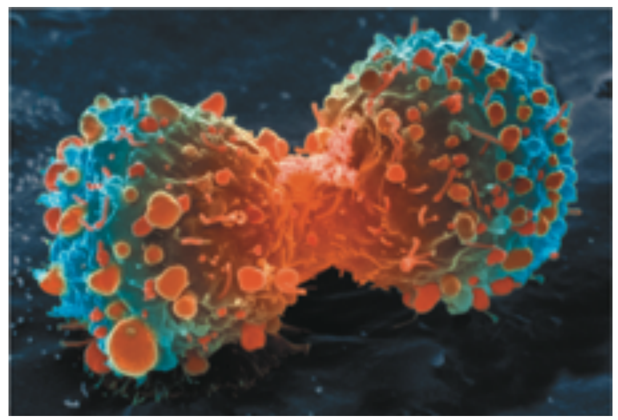
Several years ago, I spent a day and a night in a library reading through issues of the *Journal of Experimental Psychology* from the 1920s and 1930s. This was professionally a most depressing experience. Not because these articles were methodologically mediocre. On the contrary, many of them make today's research pale in comparison to their diversity of methods and statistics, their detailed reporting of single-case data rather than mere averages, and their careful selection of trained subjects. And many topics—such as the influence of the gender of the experimenter on the performance of the participants—were of interest then as now. What depressed me was that almost all of this work is forgotten; it does not seem to have left a trace in the collective memory of our profession. It struck me that most of it involved collecting data without substantive theory. Data without theory are like a baby without a parent: their life expectancy is low.


The good Thing
about science
is that it's true
whether or not
you believe in it
Dr. Neil d asse Tyson

Counterpoint:

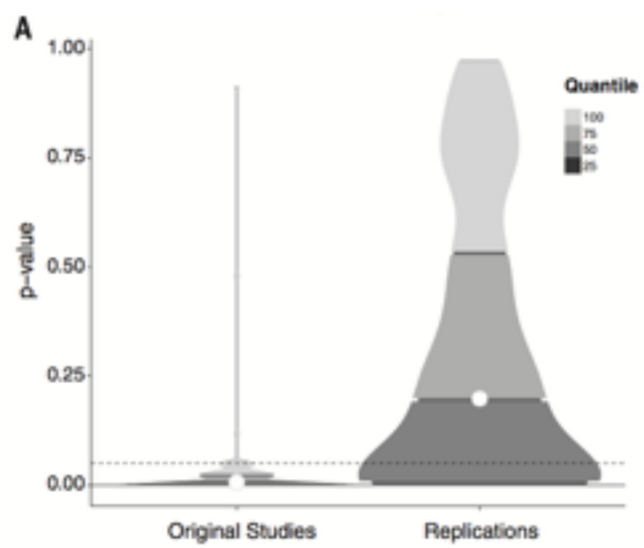
Oncology
47/53 ‘landmark’ studies did not replicate

(Begley & Ellis 2012, *Nature*)



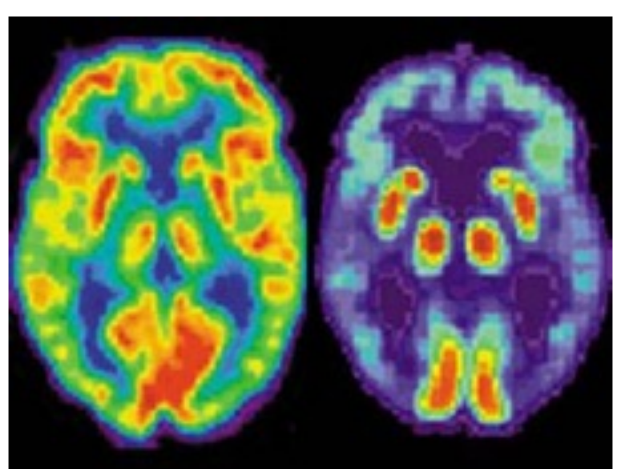
Psychology
61/100 studies in top journals failed to replicate ($p < .05$)

(Open Science Collaboration 2015, *Science*)

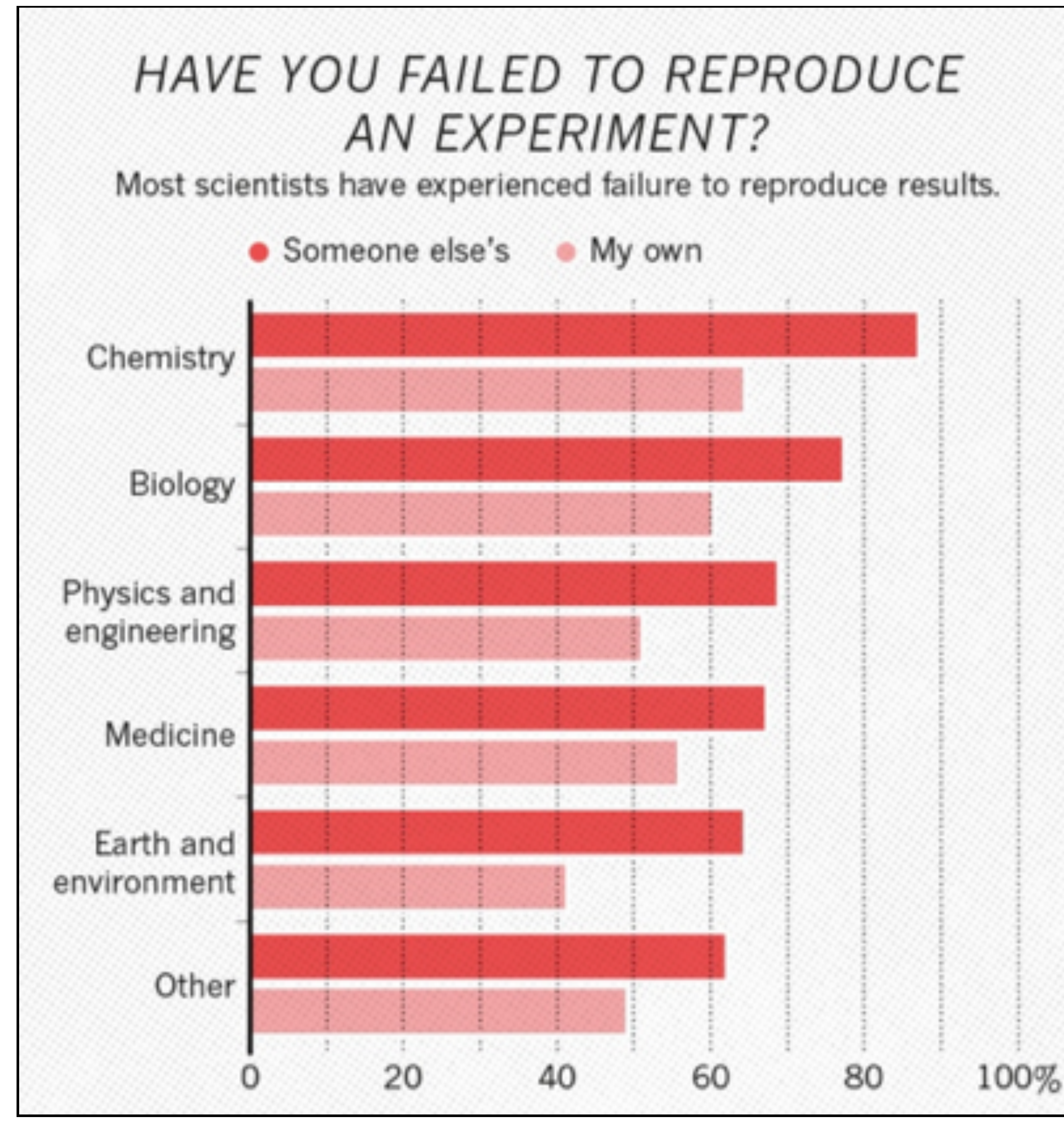


Neuroscience
Errors in popular statistical methods imply false positive rate of up to 70%

(Eklund et al. 2016, *PNAS*)

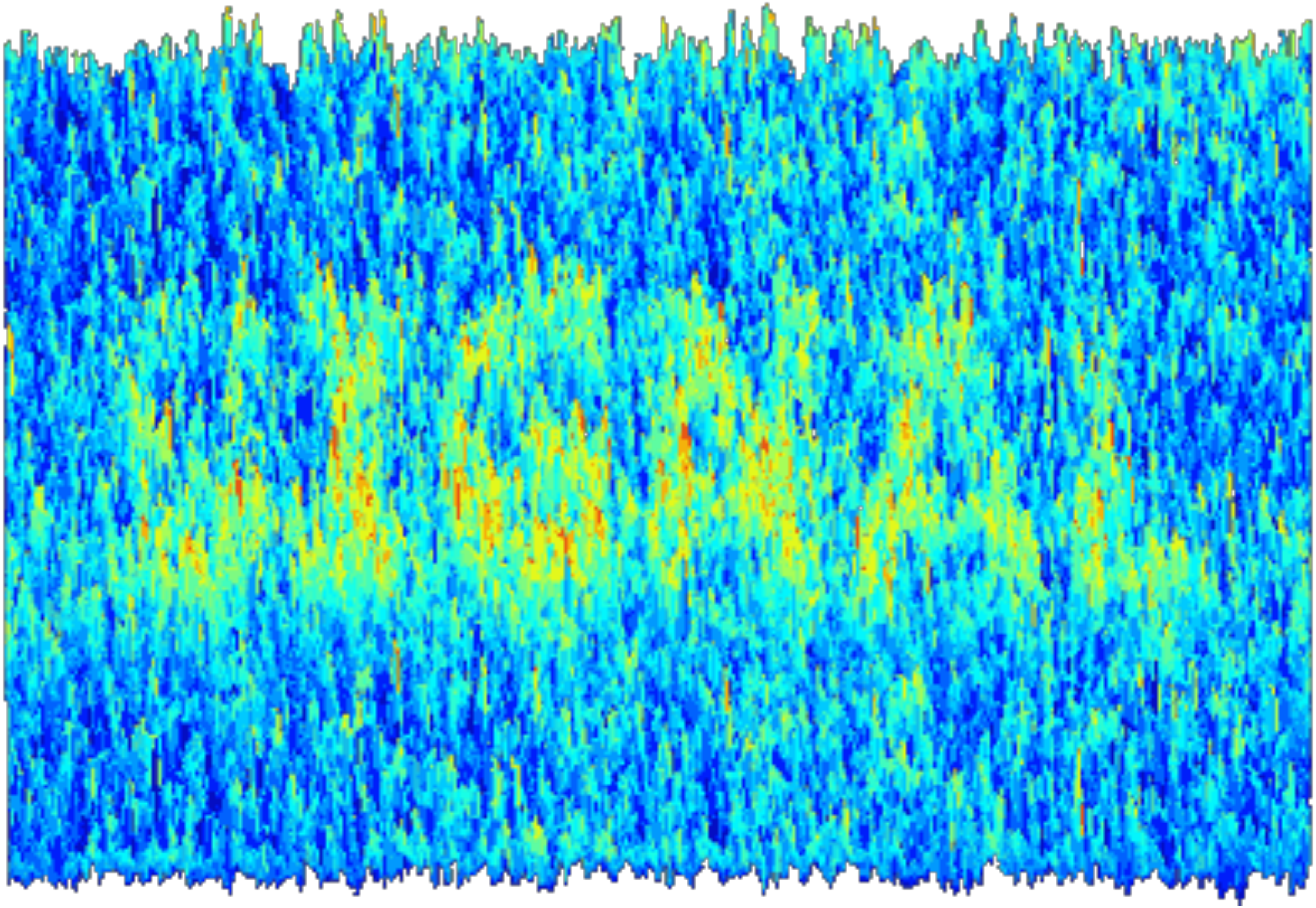


Most fields?



(Baker 2016, *Nature*)

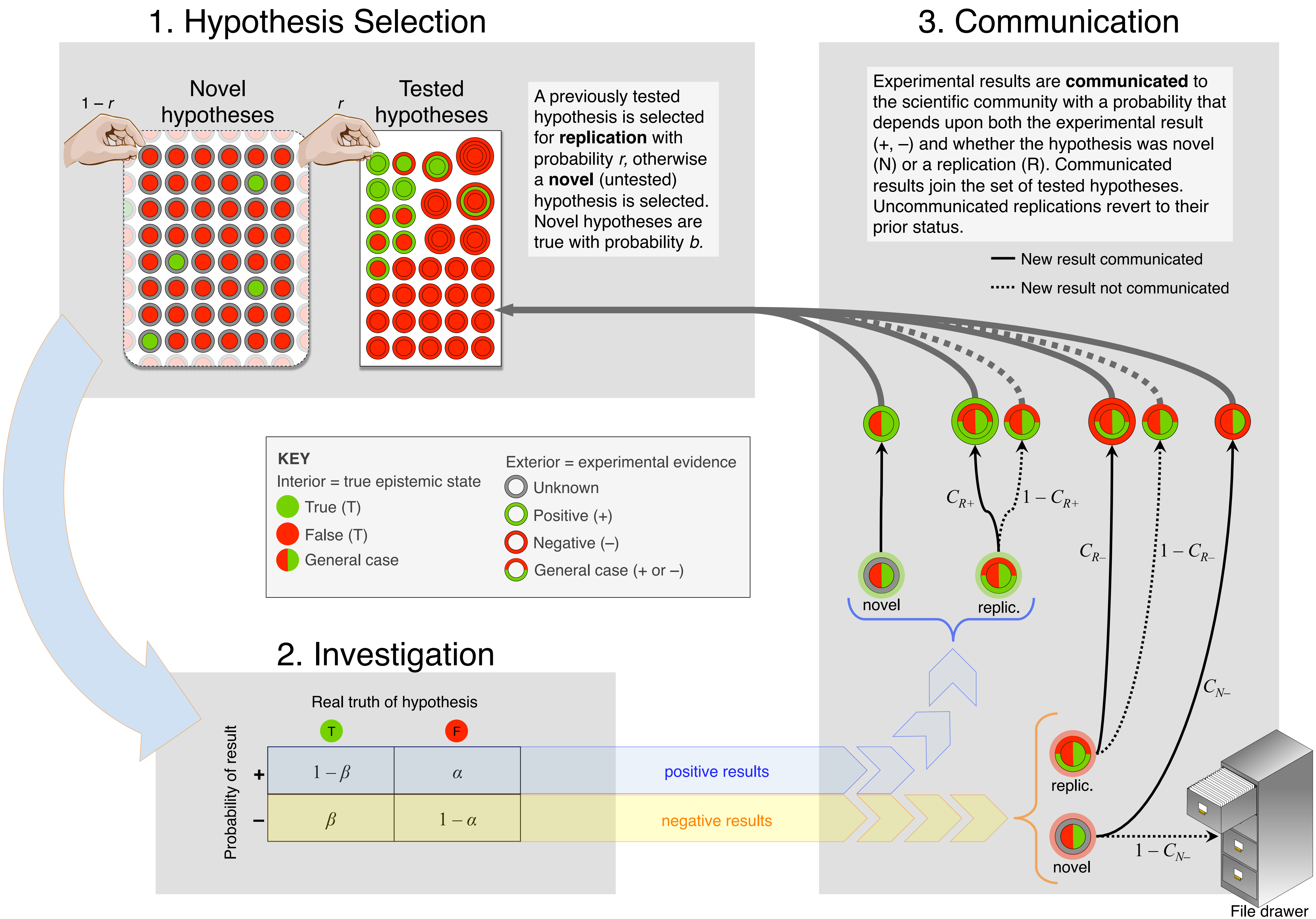
Science as Signal Detection for Facts



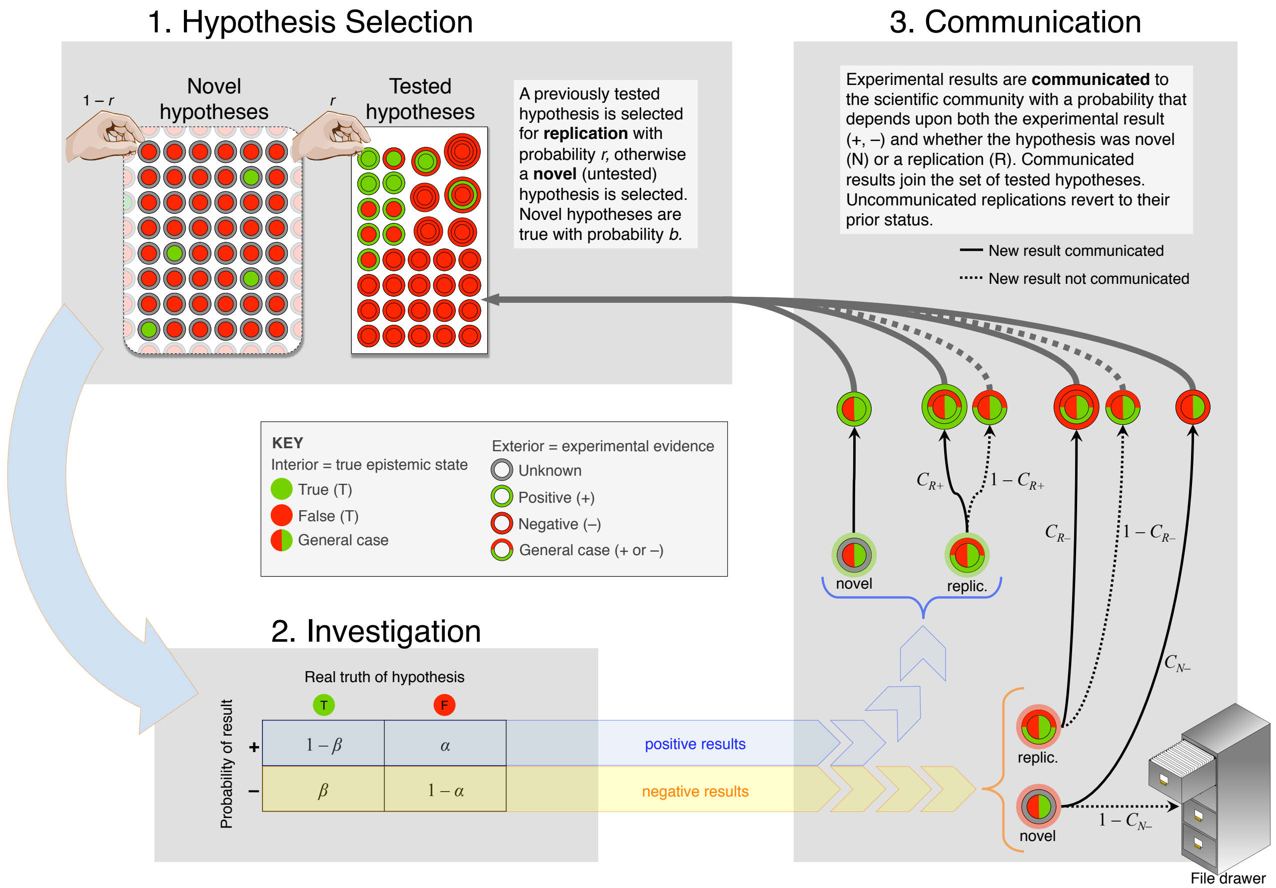
How do we find facts?

Real truth of hypothesis

Probability of result	T	F	
+	$1-\beta$	α	positive results
−	β	$1-\alpha$	negative results



McElreath R & Smaldino PE (2015) Replication, communication, and the population dynamics of scientific discovery. *PLOS ONE* 10(8):e0136088.



Recursions:

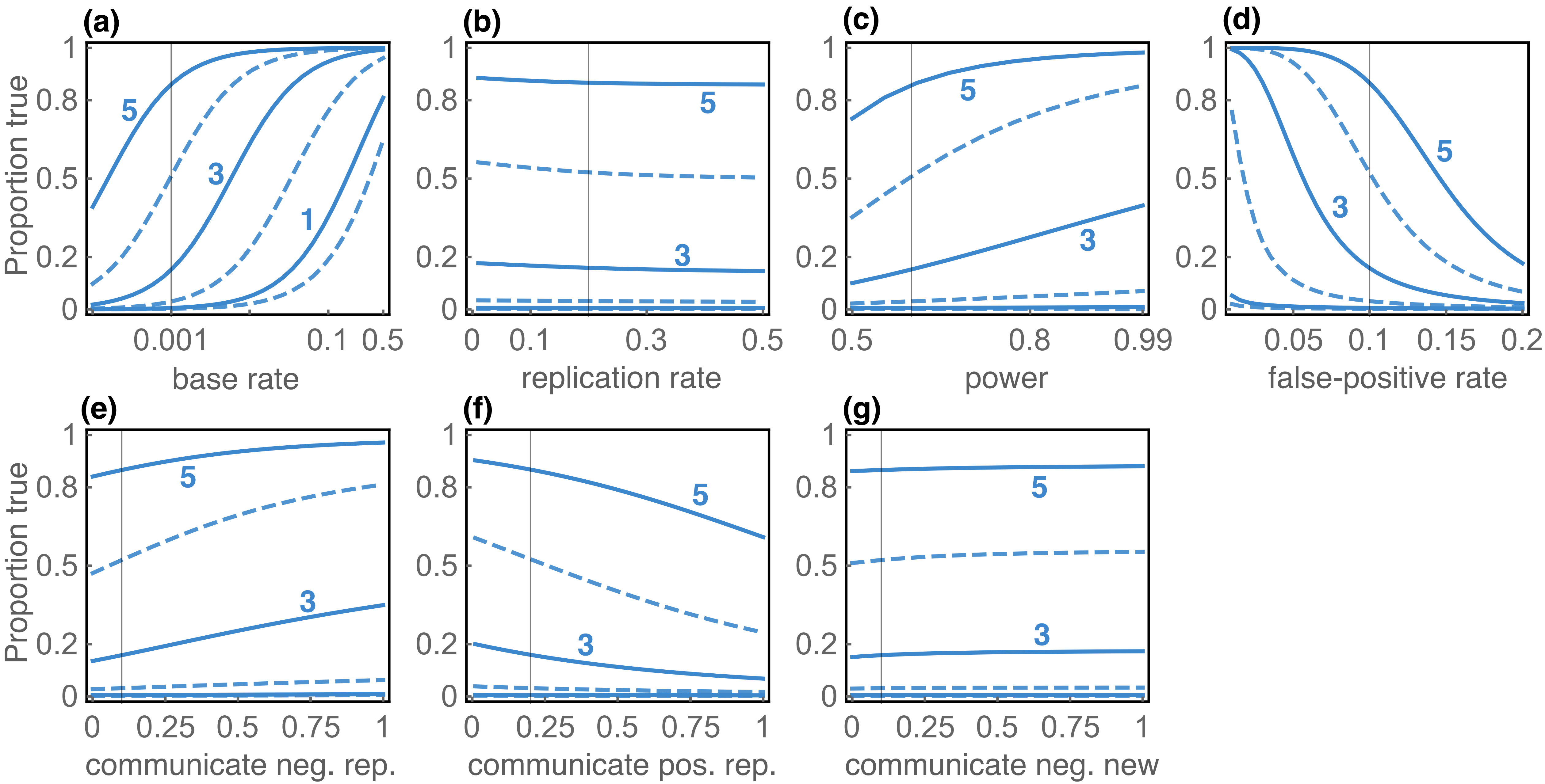
$$n'_{\mathrm{T},s} = n_{\mathrm{T},s} + anr\big(-f_{\mathrm{T},s}(c_{\mathrm{R}+}(1-\beta) + c_{\mathrm{R}-}\beta) + f_{\mathrm{T},s-1}(1-\beta)c_{\mathrm{R}+} + f_{\mathrm{T},s+1}\beta c_{\mathrm{R}-}\big)$$

Solutions:

$$\hat{p}_{\mathrm{T},s} = b(1-r)\sum_{m=1}^{\infty} r^{m-1}K\big(m, (m+s)/2\big)(1-\beta)^{\frac{1}{2}(m+s)}\beta^{\frac{1}{2}(m-s)}$$

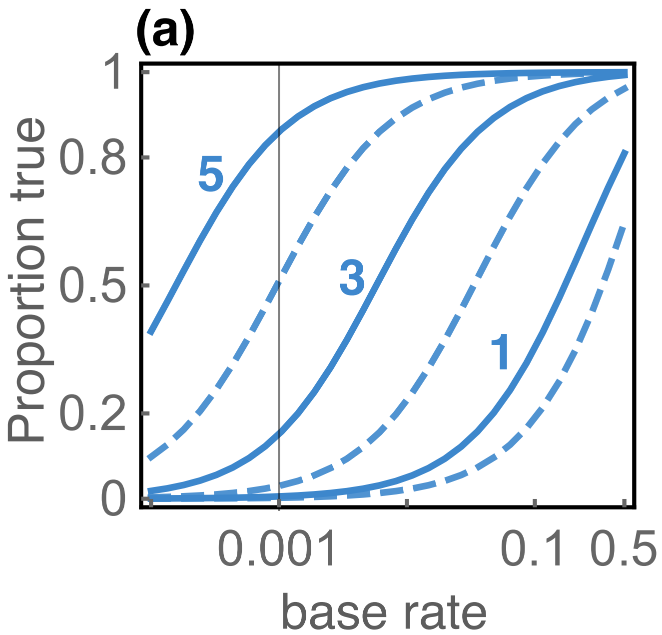
McElreath R & Smaldino PE (2015) Replication, communication, and the population dynamics of scientific discovery. *PLOS ONE* 10(8):e0136088.

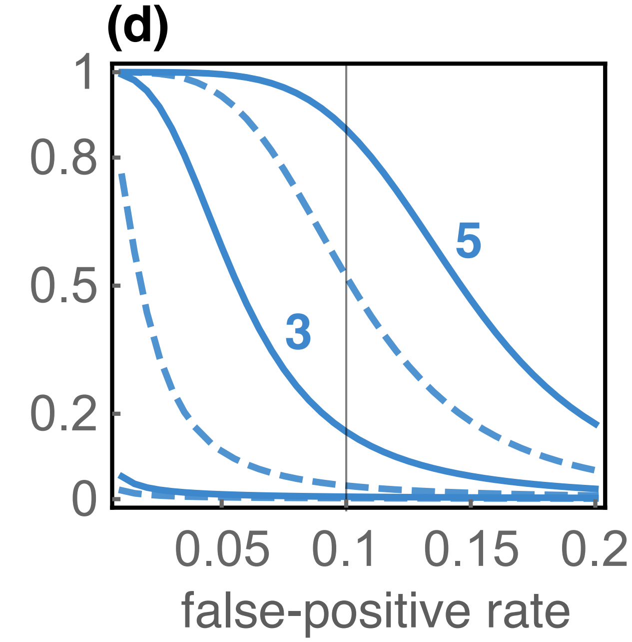
Proportion true hypotheses at different numbers of net positive findings



McElreath R & Smaldino PE (2015) Replication, communication, and the population dynamics of scientific discovery. *PLOS ONE* 10(8):e0136088.

Proportion true hypotheses at different numbers of net positive findings





Base rate and false-positive rate most important factors

McElreath R & Smaldino PE (2015) Replication, communication, and the population dynamics of scientific discovery. *PLOS ONE* 10(8):e0136088.

“Nothing in biology makes sense except in light of evolution”

–Theodosius Dobzhansky (1973)

“All social science research must do some violence to reality in order to reveal simple truths.”

–Lazer & Friedman (2007)

Turning your idea into a model

- Not a trivial problem
- Look for existing solutions
- Get creative
- Keep it simple (KISS)
- Solicit feedback
- Remember Hofstadter's Law