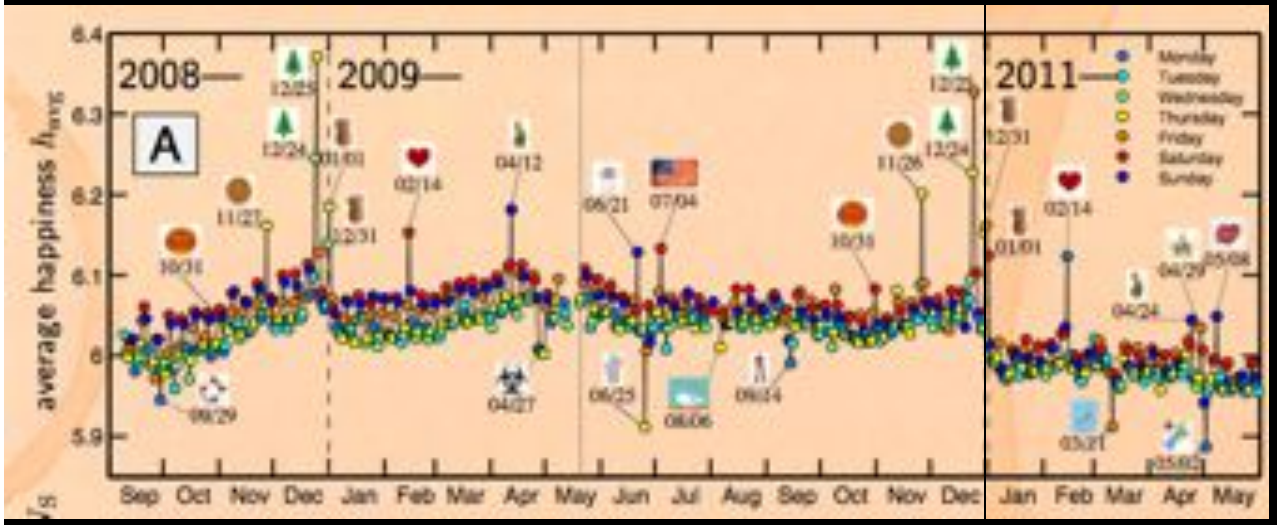
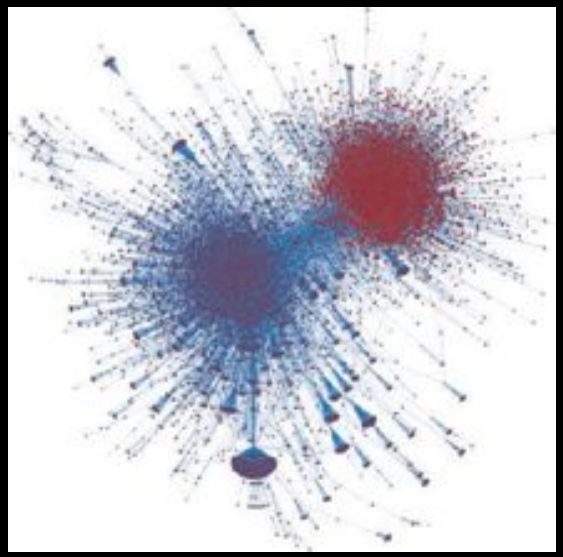
Social influence and drift in collective behavior

Alex Bentley
Archaeology and Anthropology, Bristol University

Data firehose: crowd wisdom, WEIRD?



Peter Dodds, U. Vermont



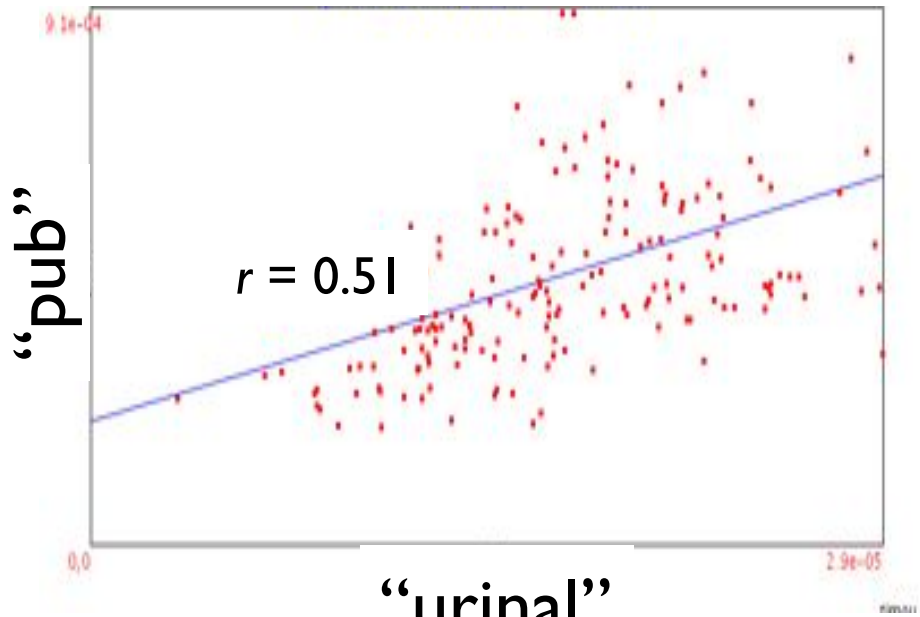
Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures

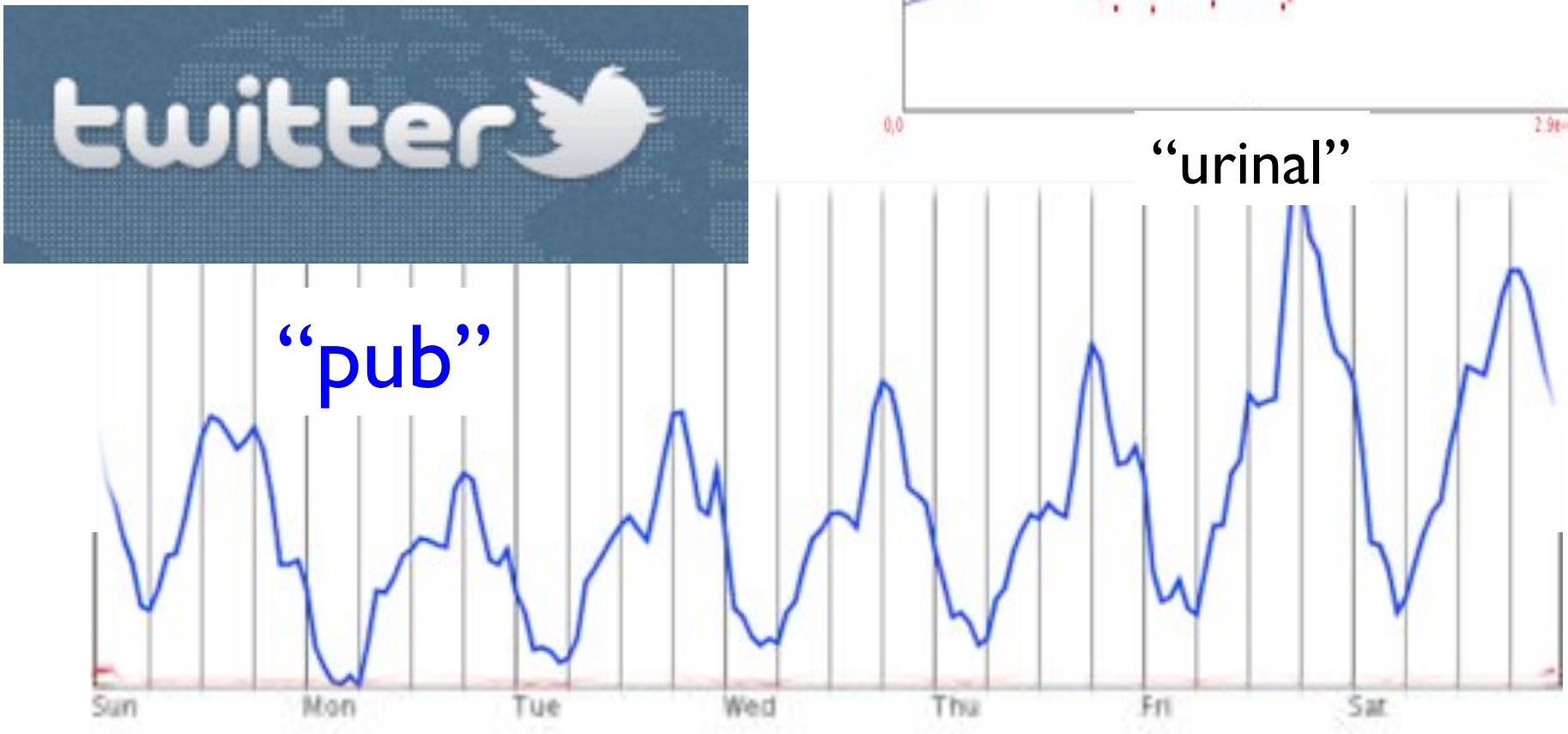
Scott A. Golder* and Michael W. Macy

Science, Oct 2011

We identified individual-level diurnal and seasonal mood rhythms in cultures across the globe, using data from millions of public Twitter messages. We found that individuals awaken in a good mood that deteriorates as the day progresses—which is consistent with the effects of sleep and circadian rhythm—and that seasonal change in baseline positive affect varies with change in daylength. People are happier on weekends, but the morning peak in positive affect is delayed by 2 hours, which suggests that people awaken later on weekends.

El Farol revisited

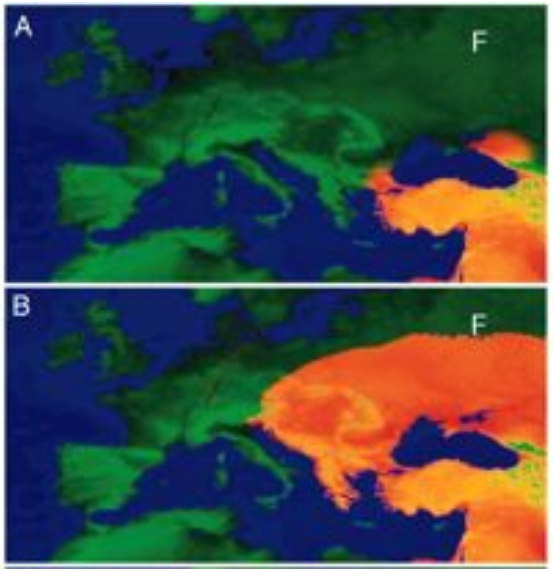




http://timeu.se (Scott Golder, Cornell U.)

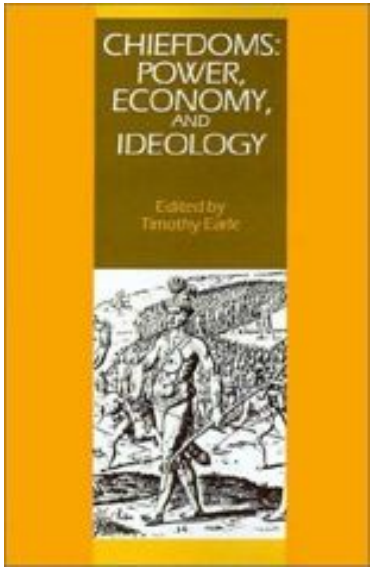
Anthropology and Archaeology

Prehistoric dispersals



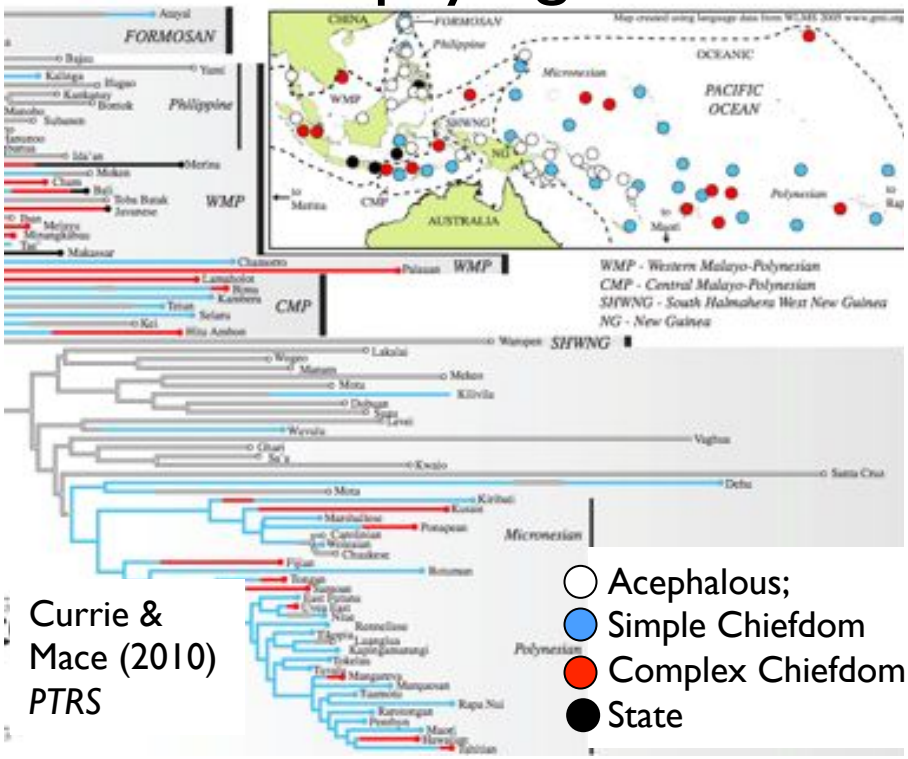
Ackland et al. (2005) *PNAS*

Hierarchies



Timothy Earle
NU Anthropology

Cultural phylogentics

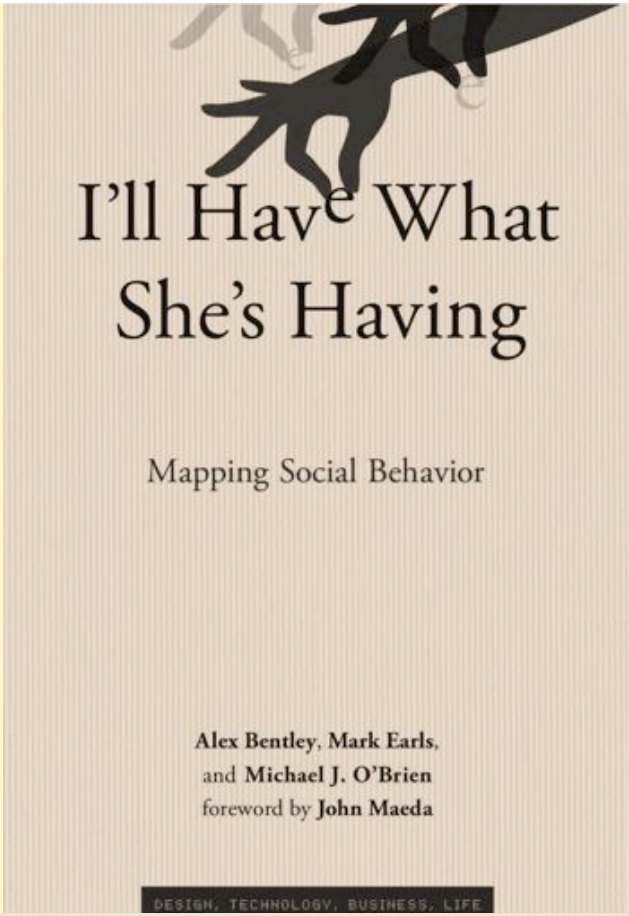


Currie & Mace (2010) *PTRS*

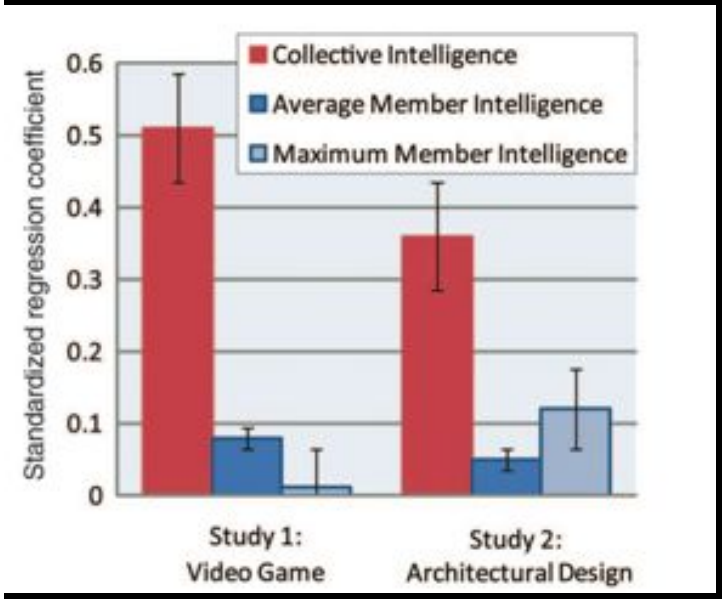
- Acephalous;
- Simple Chiefdom
- Complex Chiefdom
- State

- Cross-cultural economic experiments
- Cognition & social learning
- Evolution of cooperation & wealth

- 'Distributed mind' & cumulative culture
- Gene-culture evolution
- Prehistoric disease & networks

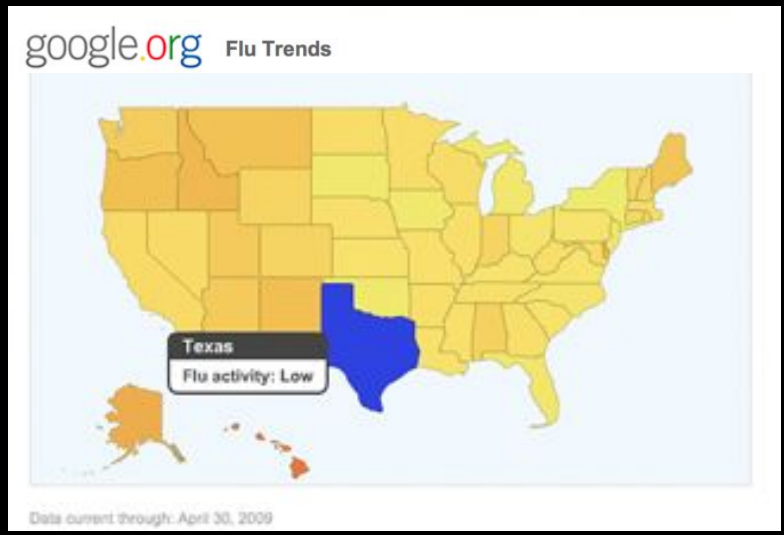


Independent, informed

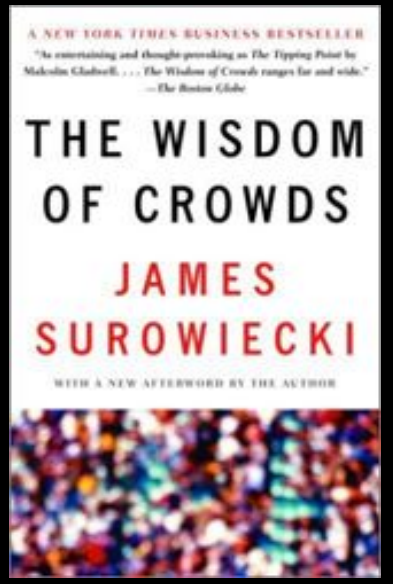


Evidence for a Collective Intelligence Factor in the Performance of Human Groups *Science* Oct 2010

Anita Williams Woolley,[1]* Christopher F. Chabris,[2,3] Alex Pentland,[3,4] Nada Hashmi,[3,5] Thomas W. Malone[3,5]

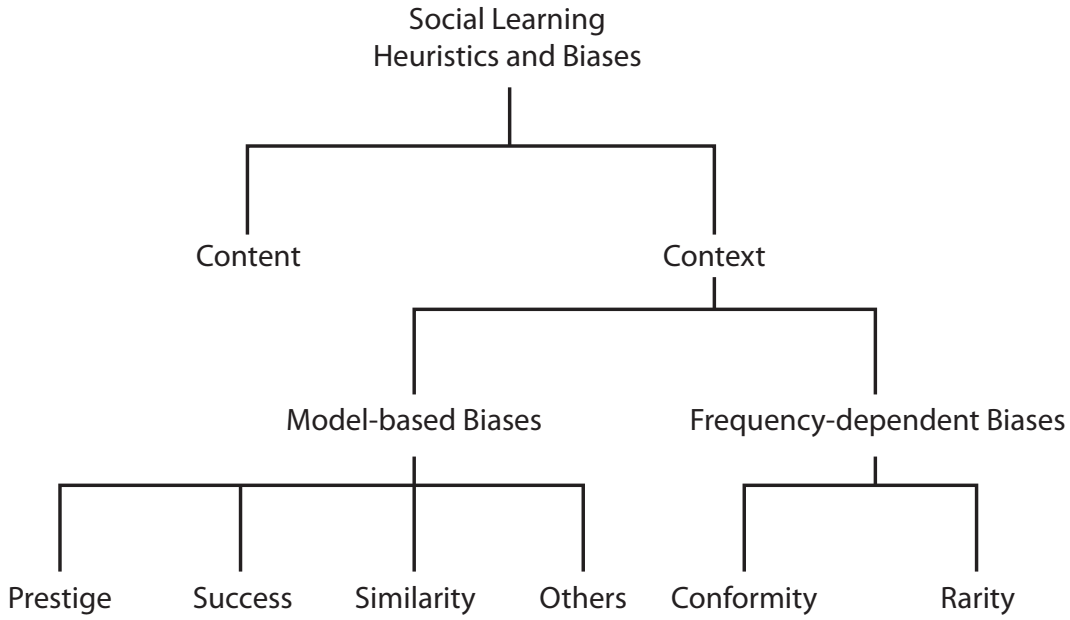


"Ask a group of diverse, independent people….the errors cancel themselves out… you're left with the information."



Social learning

Informed in different ways



O'Brien & Bentley (2011)

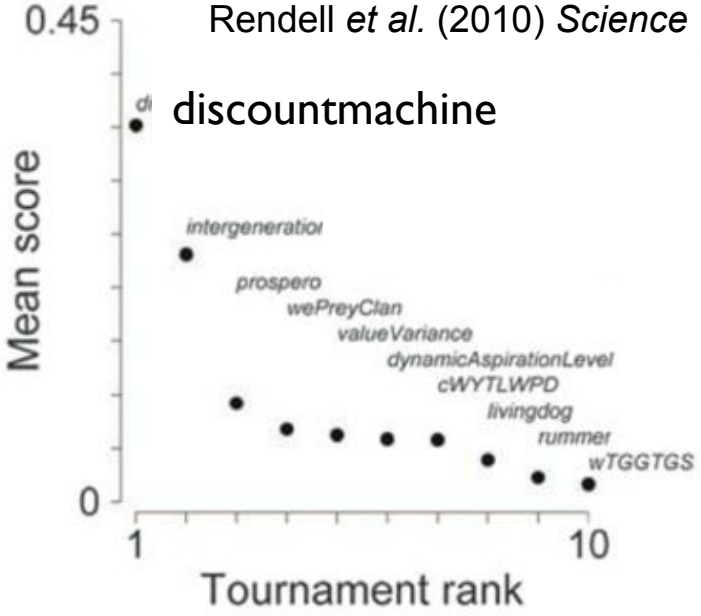
Rivera, Soderstrom, Uzzi (2010) *Ann. Rev. Sociology*
Laland (2004) *Learning and Behavior*

Hard-wired to imitate

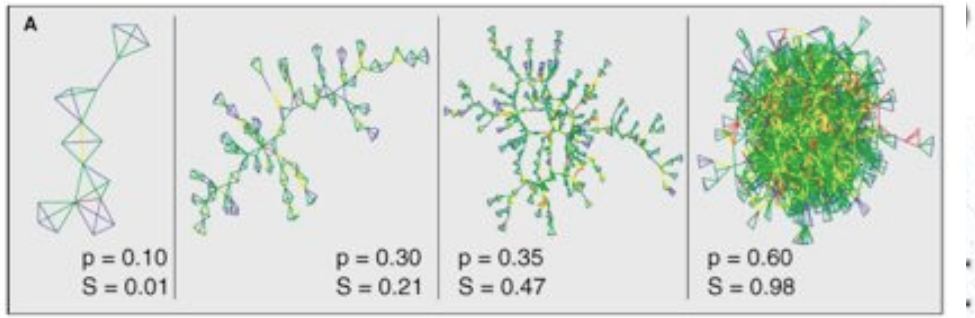
Iacoboni *et al.* (1999) *Science*
Gergely *et al.* (2004) *Nature*

Why Copy Others? Insights from the Social Learning Strategies Tournament

Rendell *et al.* (2010) *Science*

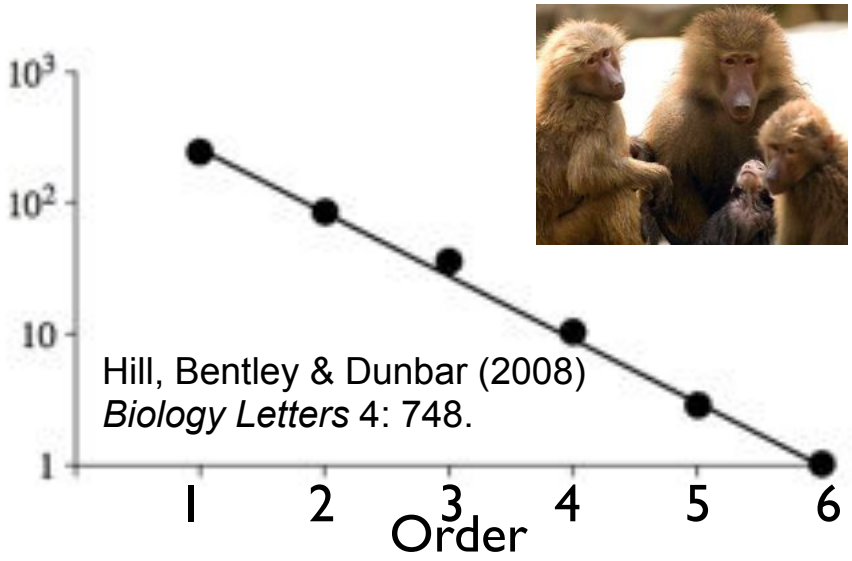


Informed copying: Social networks are ordered

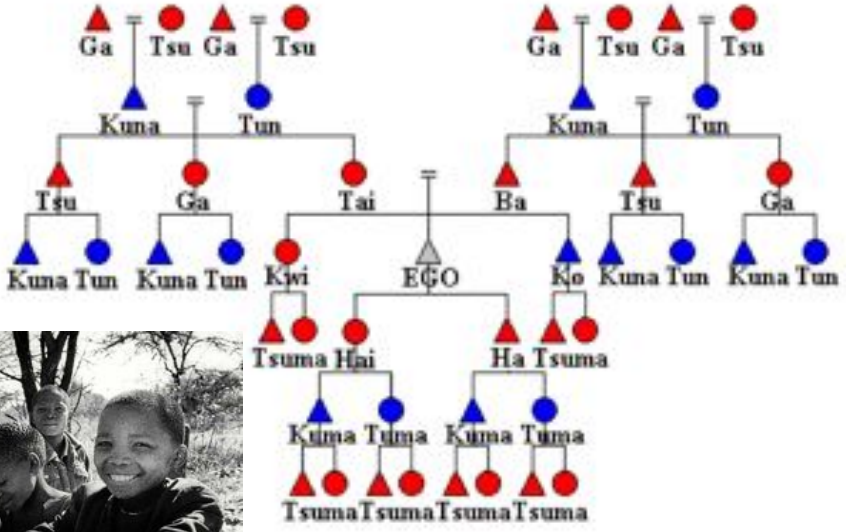


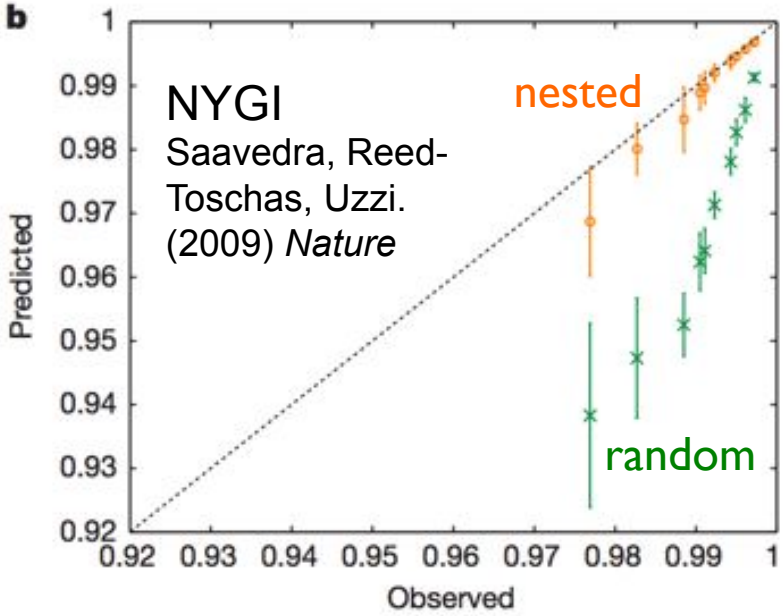
Guimera, Uzzi, Spiro, Amaral (2005) *Science* 308:697

Also Watts, Dodds, Newman (2002). *Science*



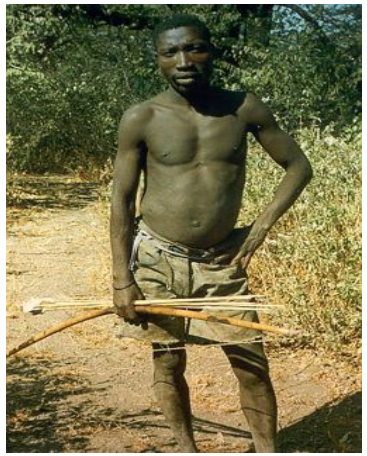
Hill, Bentley & Dunbar (2008)
Biology Letters 4: 748.



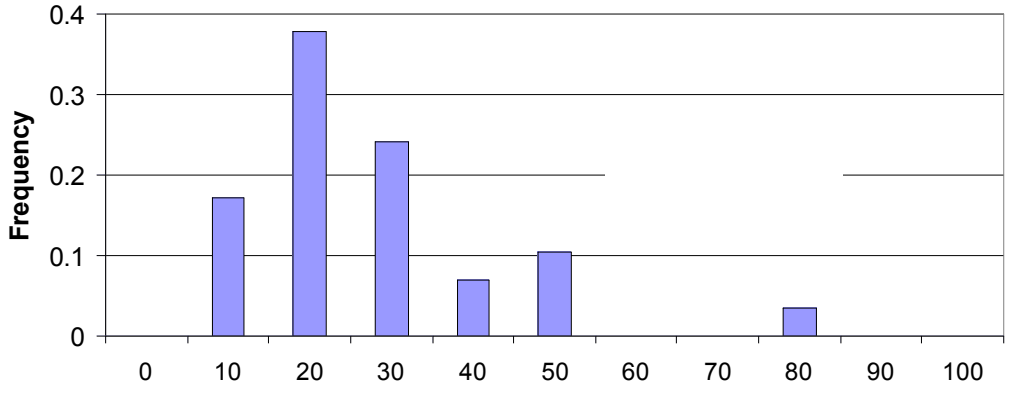


Social, well-informed: adaptive

Hadza hunter-gatherers (Tanzania)

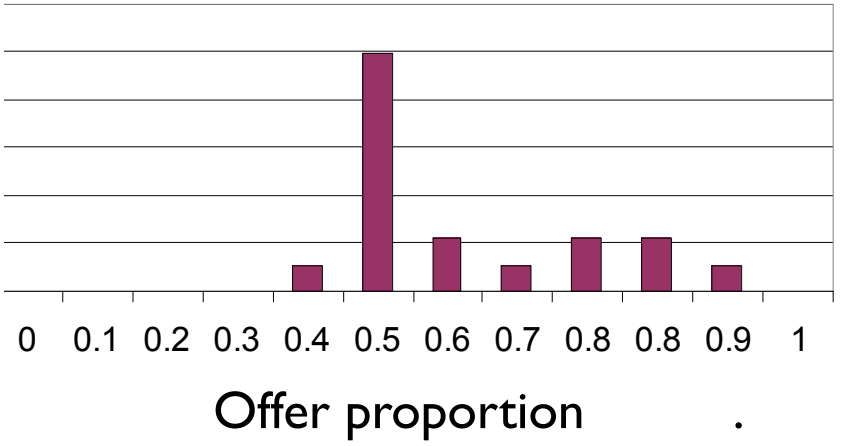


Ultimatum game

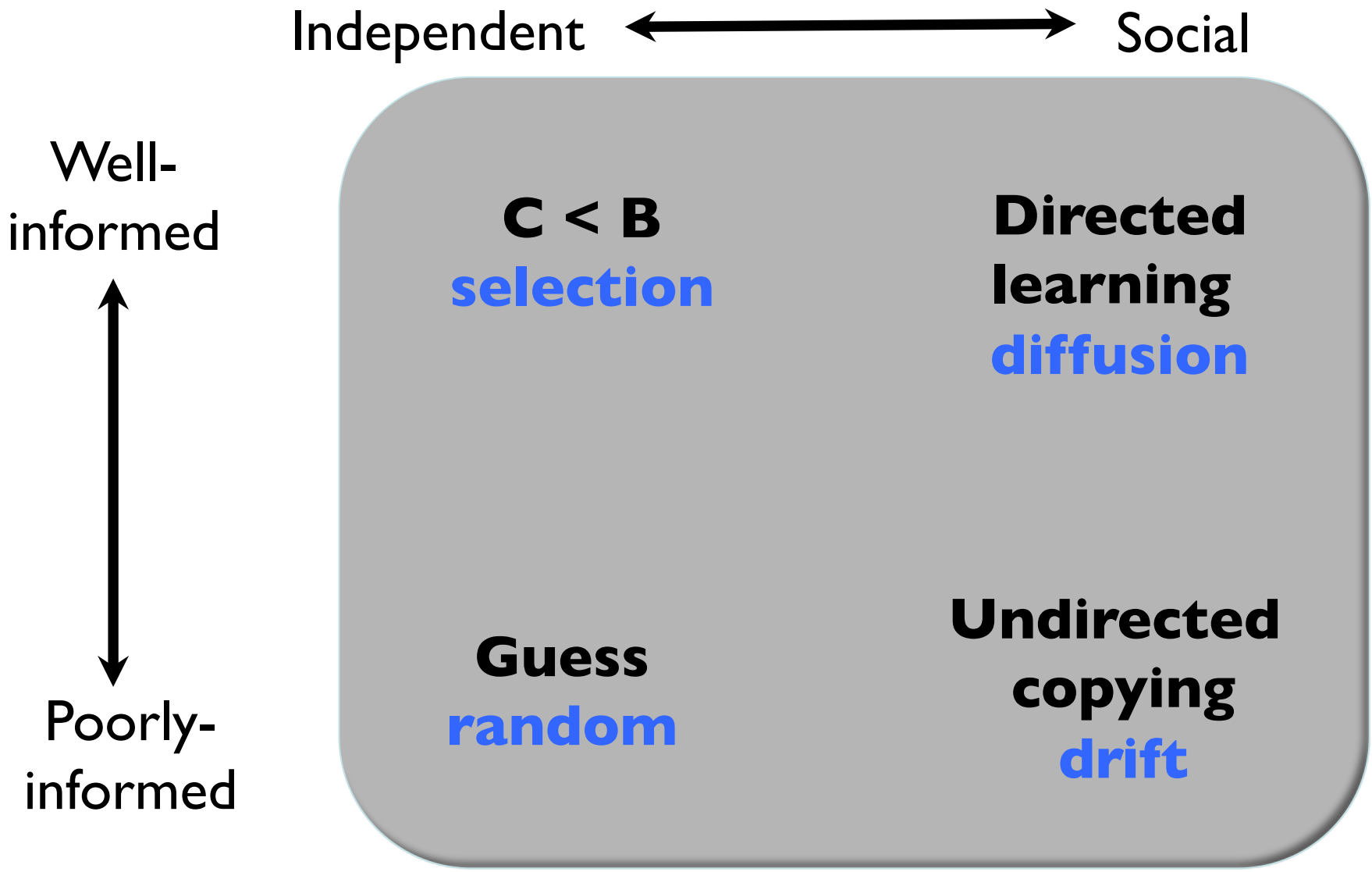


Lamalera whale hunters (Indonesia)

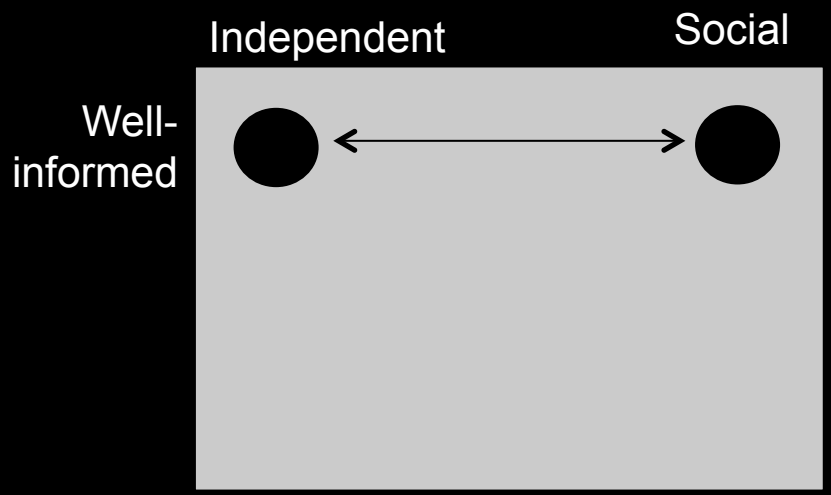
Ultimatum game



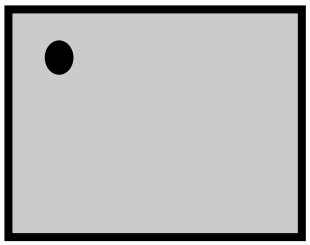
Henrich et al. (2006) *Behavioral and Brain Sciences*


Independent
Social
Well-informed
Poorly-informed
C < B
selection
Directed learning
diffusion
Guess
random
Undirected copying
drift

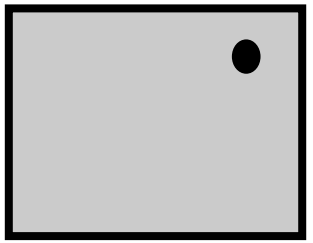
North



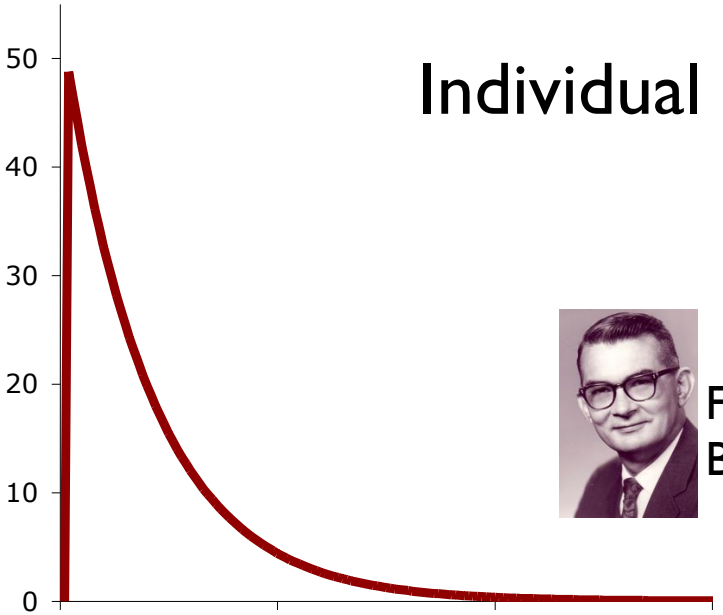
North: classic diffusion model



$$f(t) = \left(\mu + qF_t\right)\left(1 - F_t\right)$$

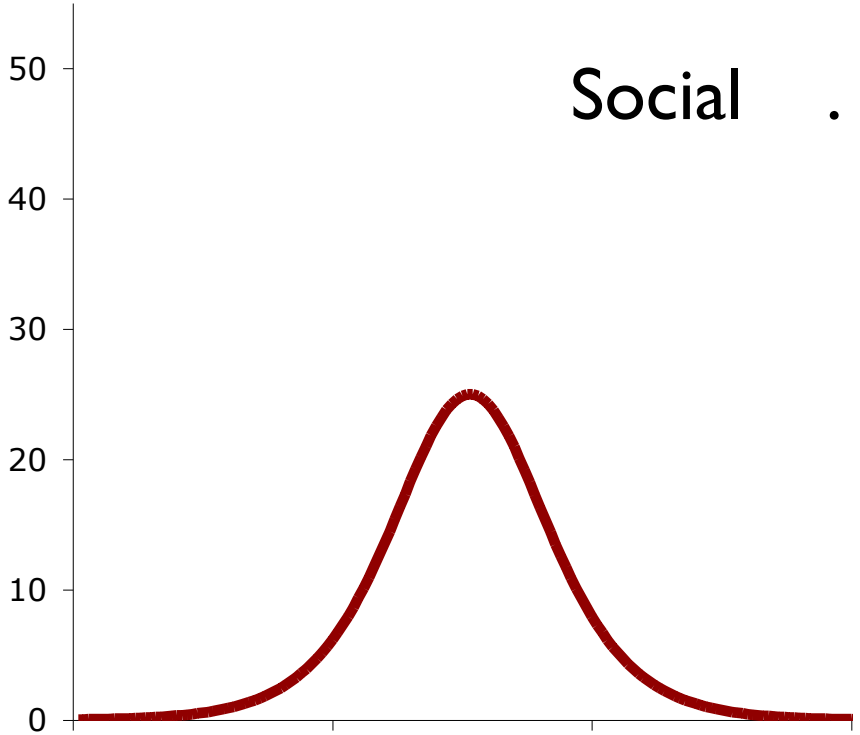


Individual

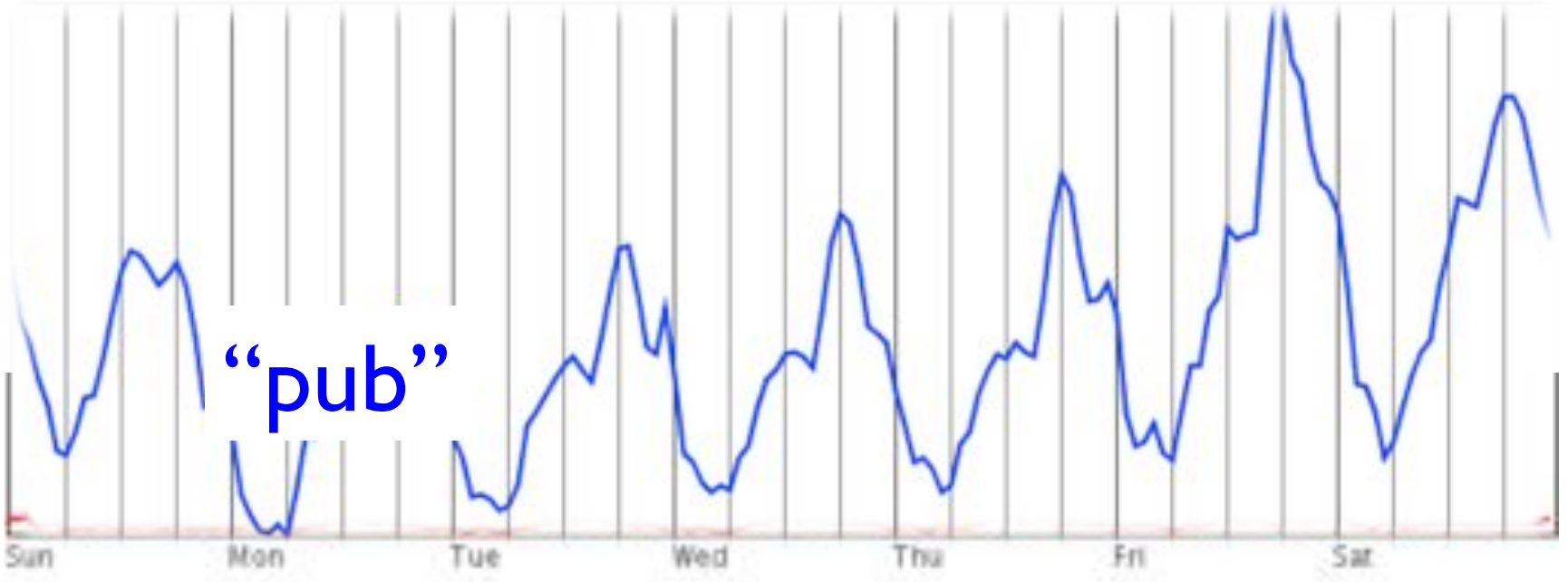


Frank Bass

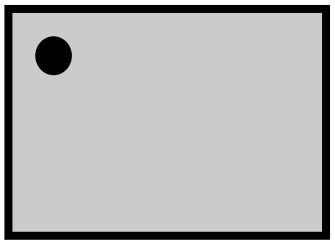
Social .



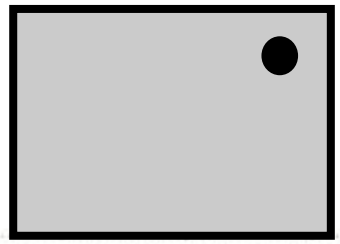
‘Bass’ curves not always social



http://timeu.se (Scott Golder, Cornell U.)

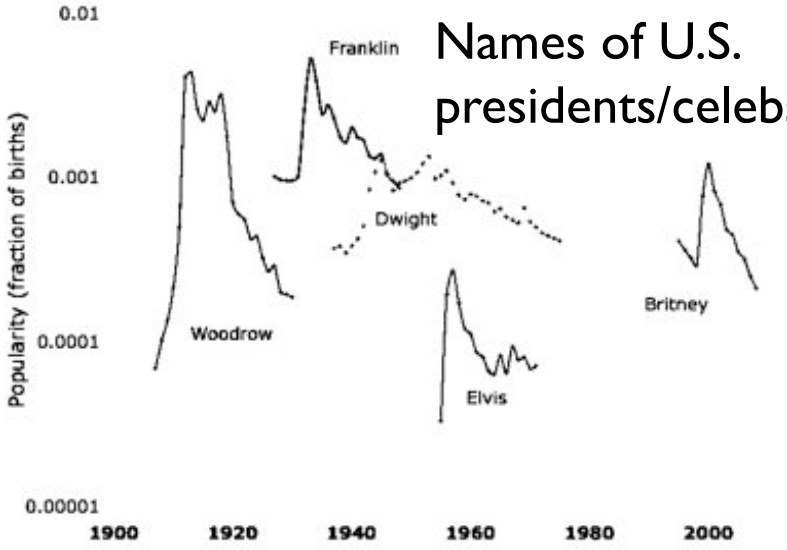


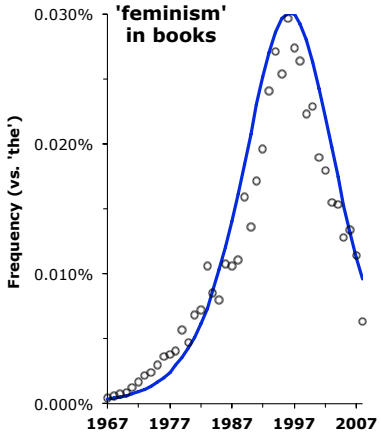
Independent, informed

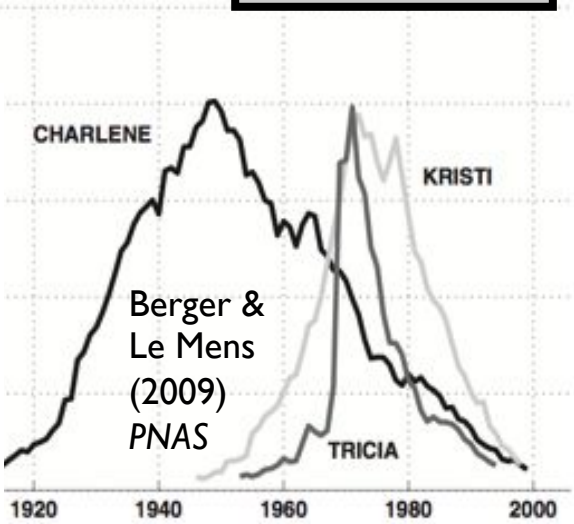


Social, informed

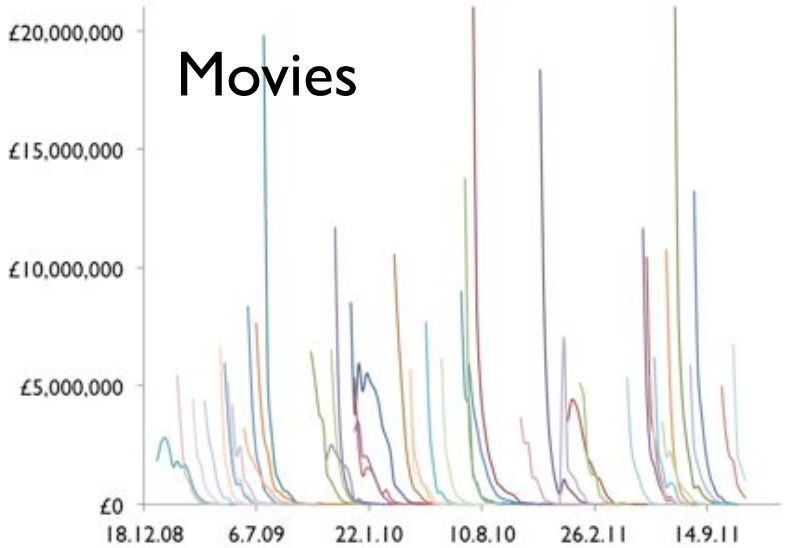
Names of U.S. presidents/celebs



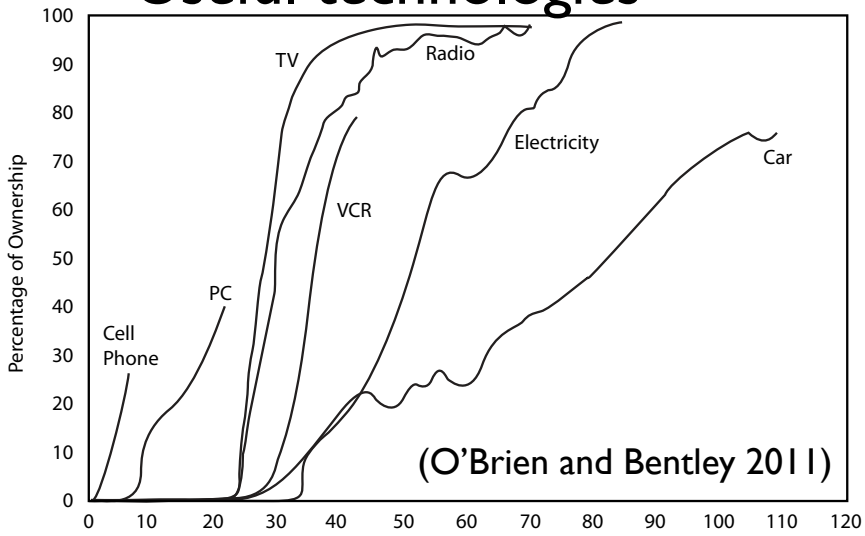




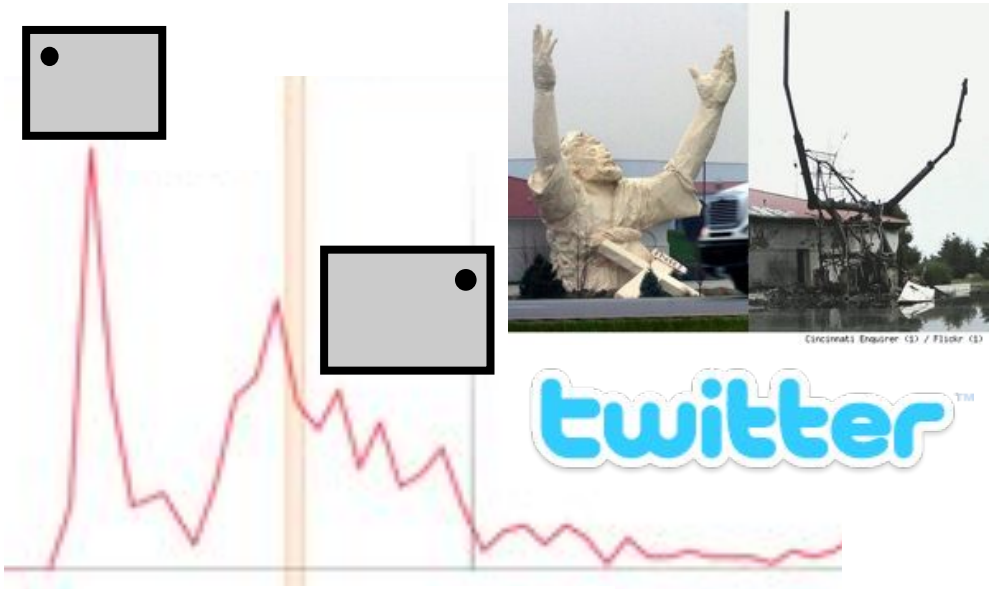
Movies

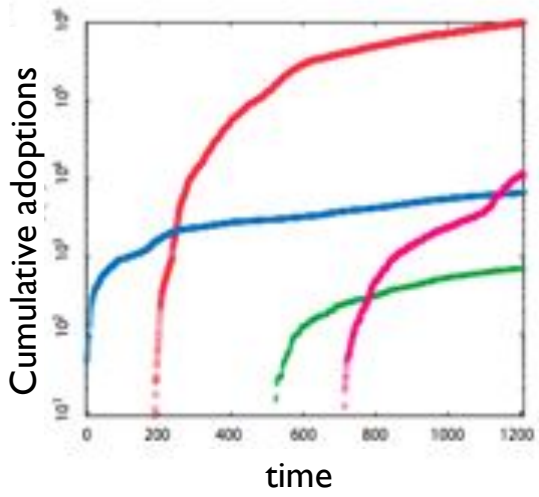


Useful technologies



Transitions

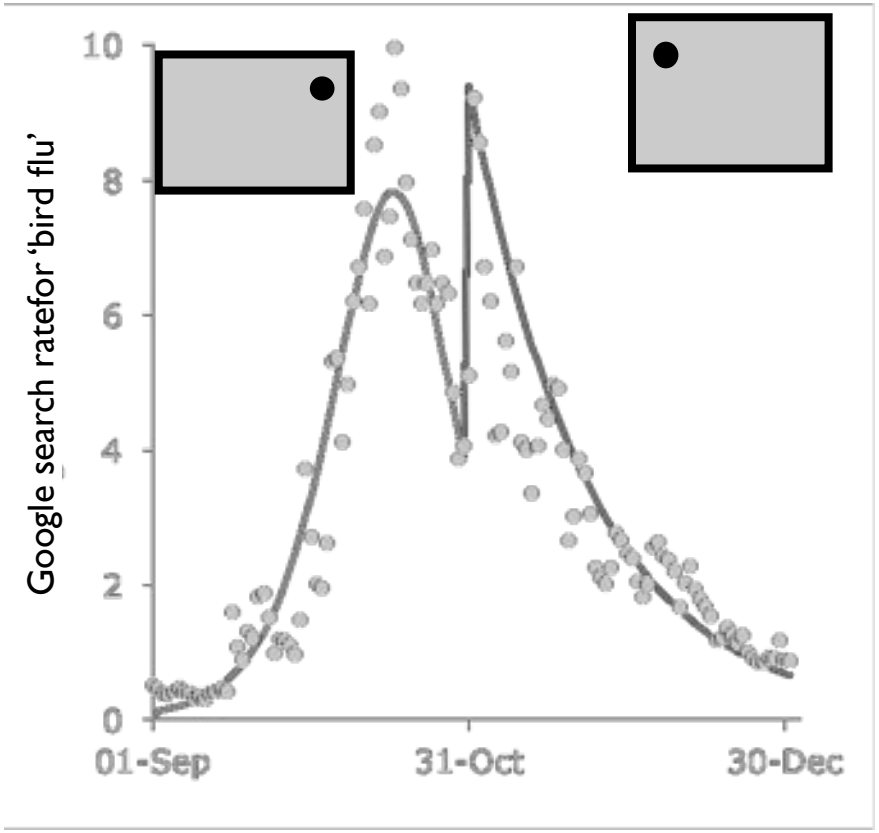




Facebook apps

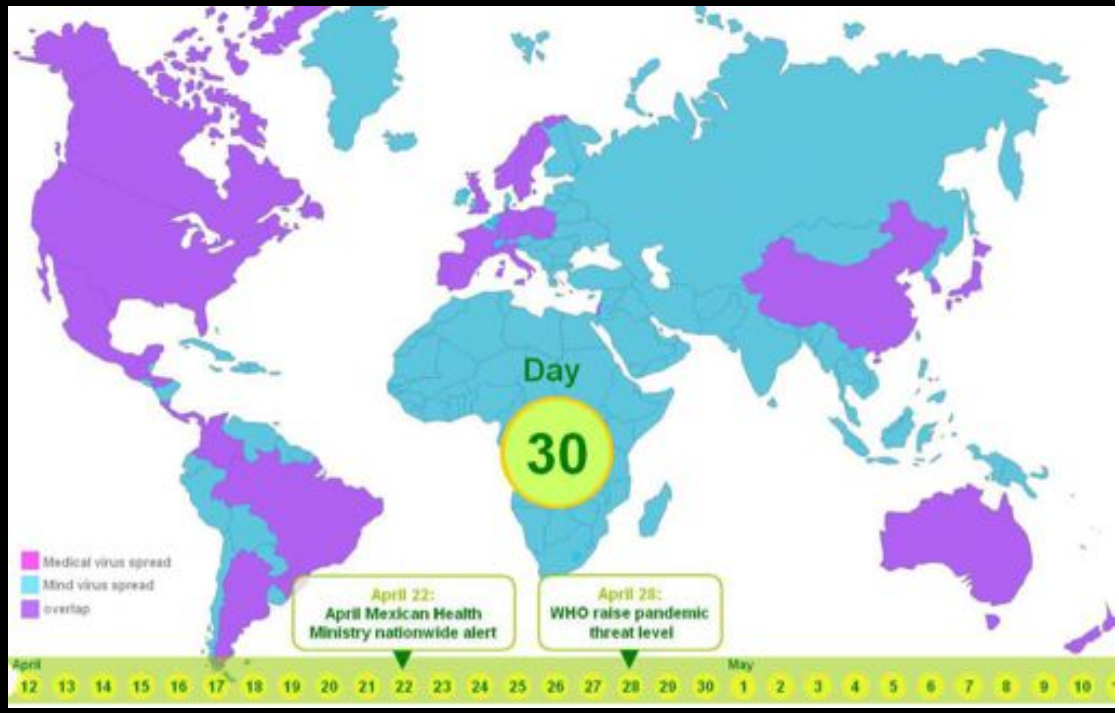
Onella and Reed-Tsochas 2010 PNAS

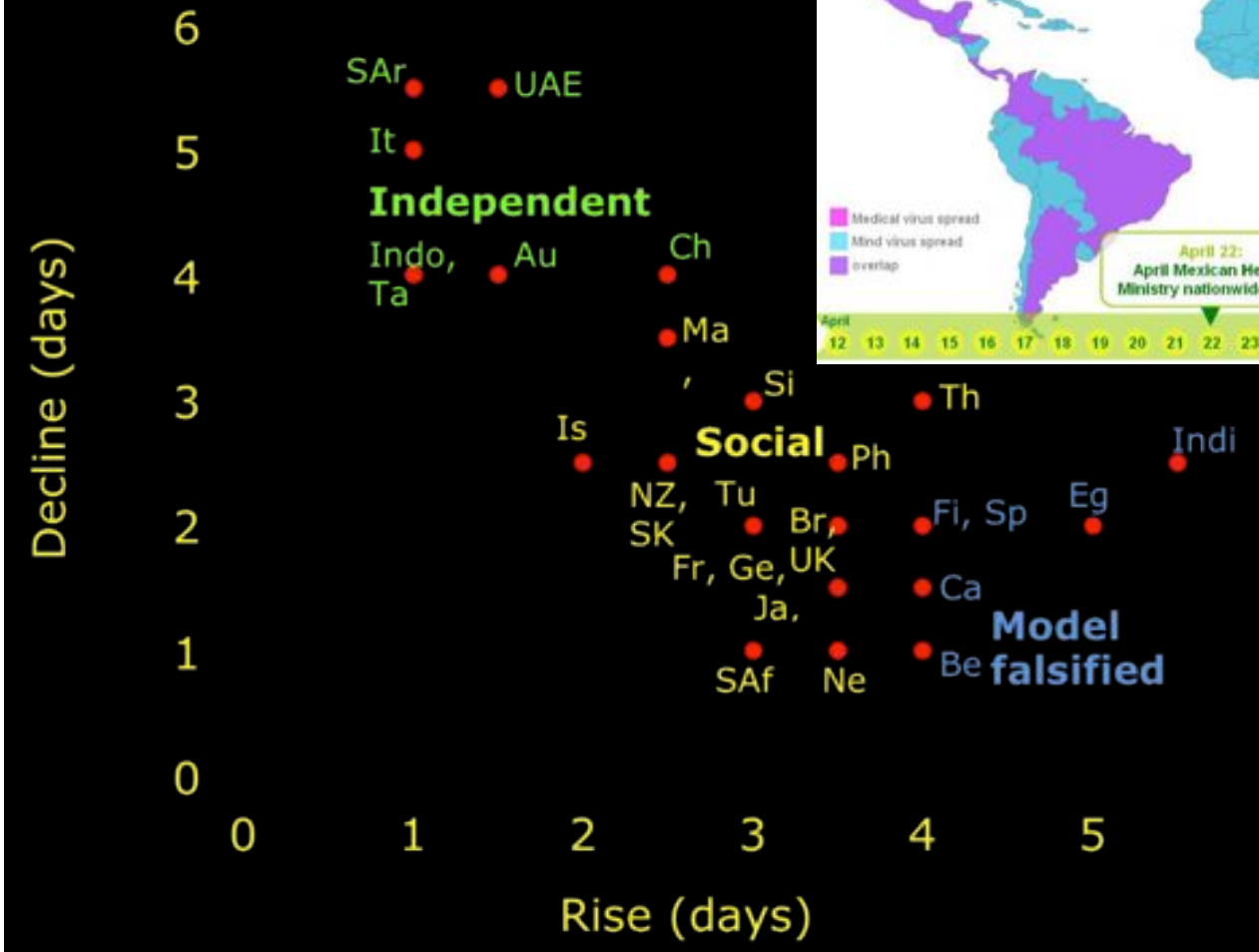
‘Bird flu’ 2005



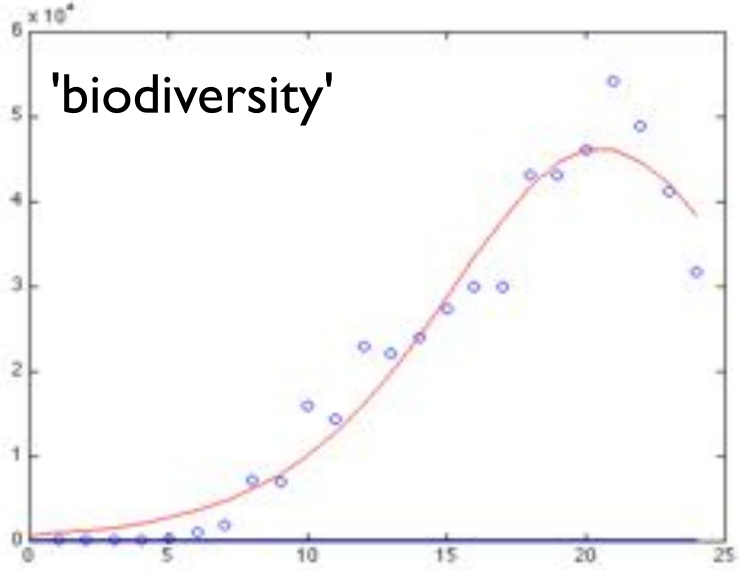
Bentley & Ormerod (2010) *Soc. Sci. & Med.* 71: 482-485.
Bentley Ormerod (2009) *PNAS 106: E*109

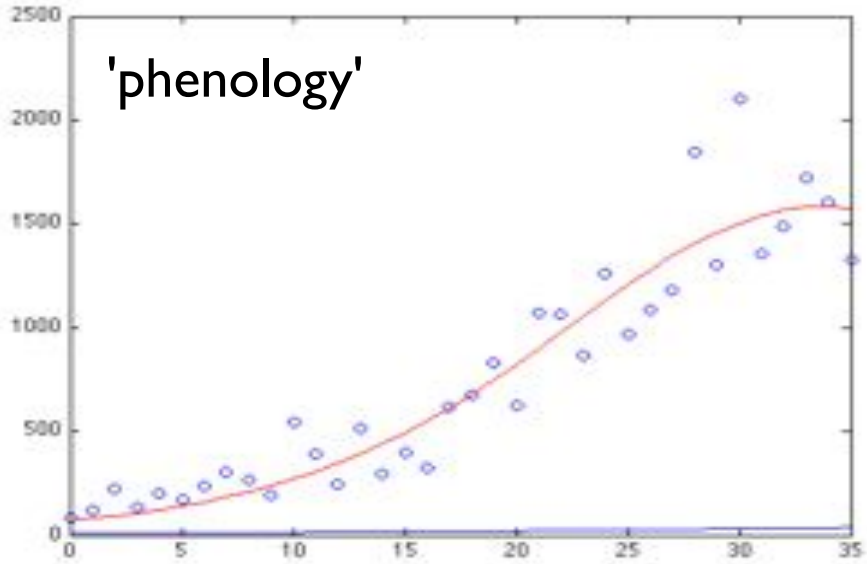
"Swine flu" April-May 2009

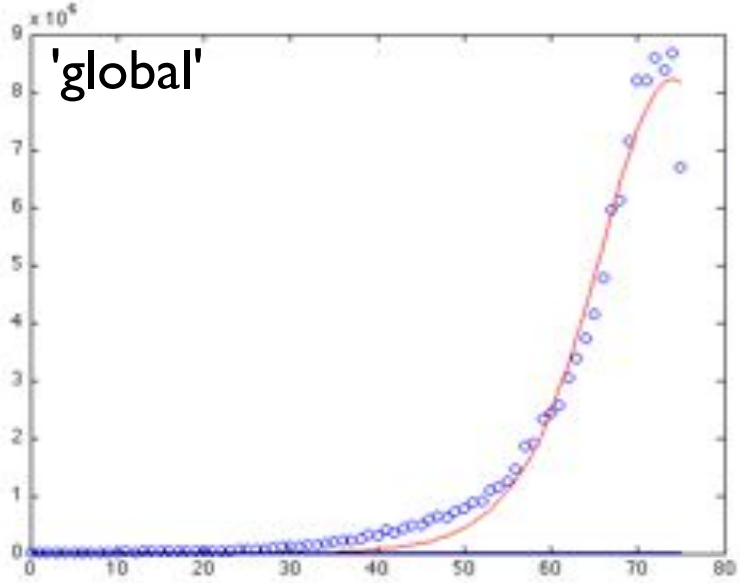




Some nice fits to bass model

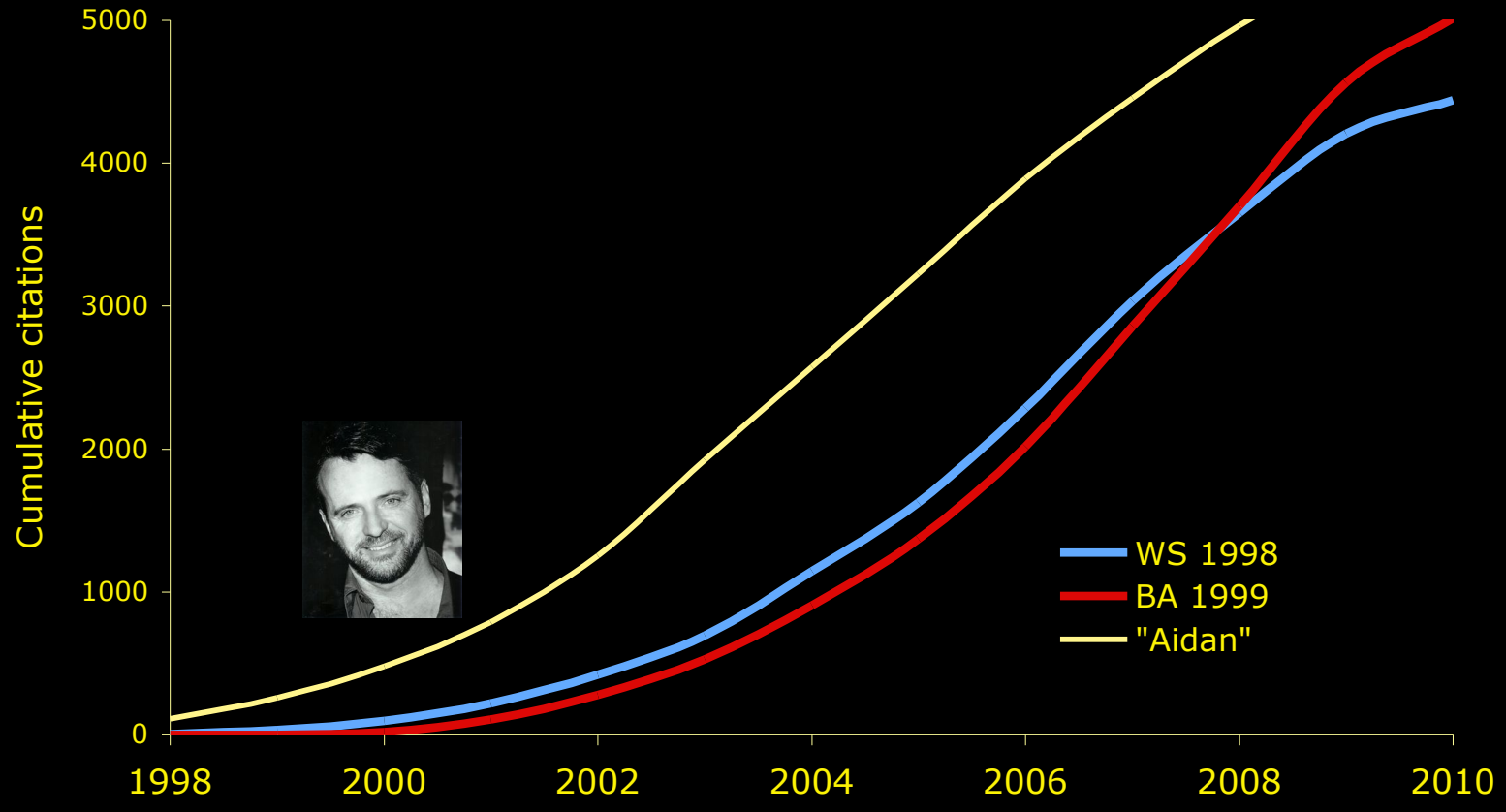






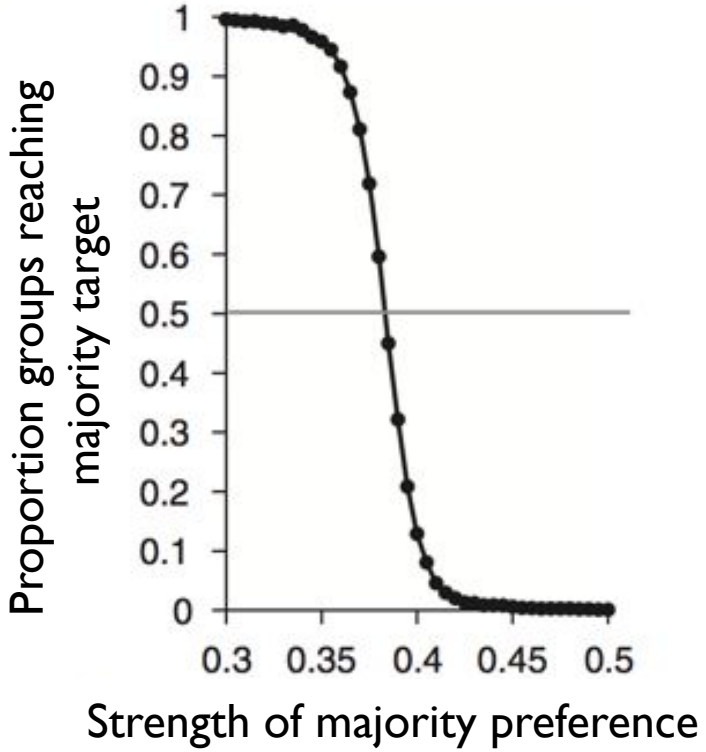
$$N(t)\frac{1-e^{-(\mu+q)t}}{1+\frac{q}{\mu}e^{-(\mu+q)t}}$$

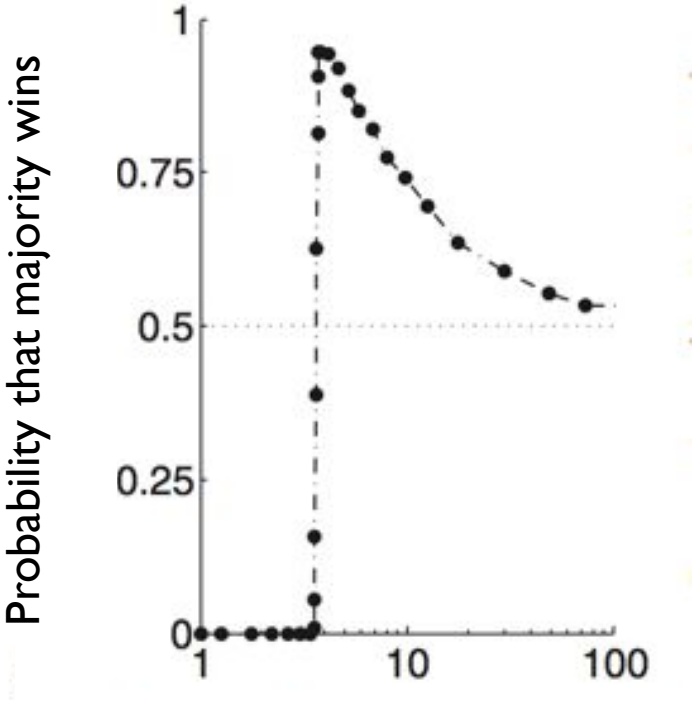
Post-hoc ‘predictability’



Citations of WS 1998 = 6% of “Aidan” 2 yrs before (r^2 = 0.998)
Same for BA 1999 (r^2 = 0.984)

How well informed?





Couzin et al. (2011). *Science* 334:1578-1580.

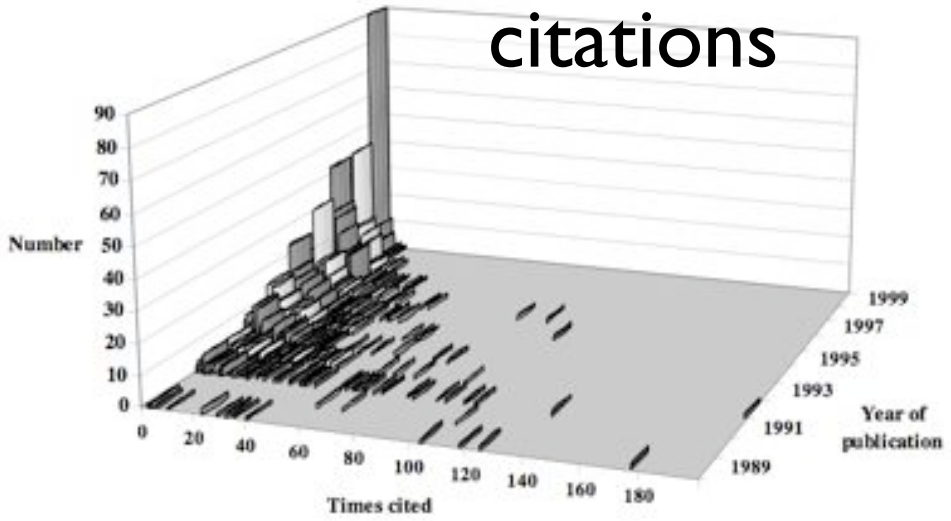
‘Mid-Atlantic’

‘noisy’ multiplicative growth

$$n_t = (1 + g_t) n_{t-1}$$

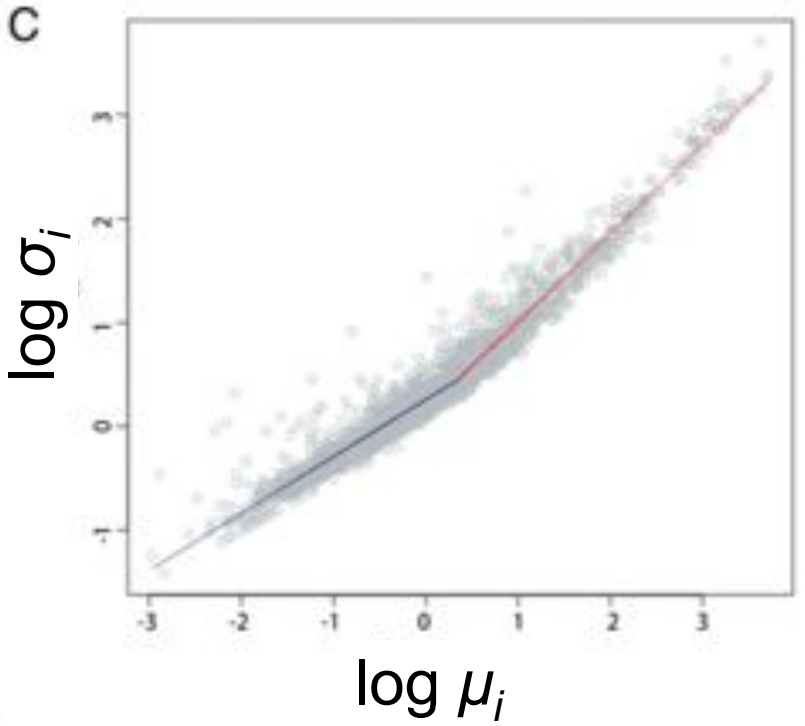
g_t normal with mean $\mu \pm \sigma$

$$P(n_s) = \frac{1}{n_s \sqrt{t 2\pi\sigma^2}} \exp\left(-\frac{(\ln n_s - g_0 t)^2}{2\sigma^2 t}\right)$$



Bentley and Maschner (2000) *Fractals*

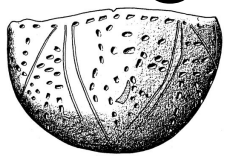
Stringer et al. (2010) *J. Am. Soc. Inf. Sci. Tech*



Onnela & Reed-Tsochas (2010) *PNAS*

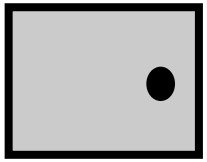
Also Adamic & Huberman 2000; Wu and Huberman (2007) *PNAS*

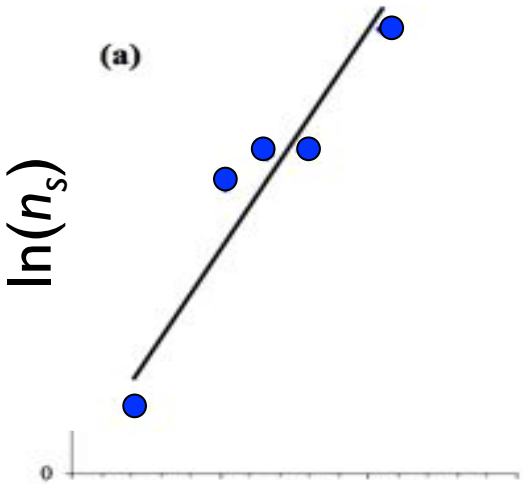
Neolithic pottery designs 5300-4900 B.C.

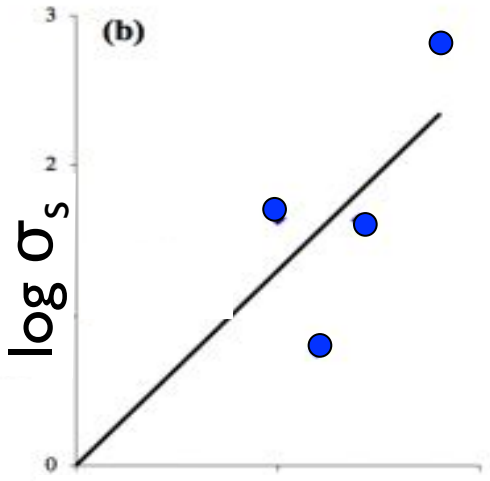


$$P(n_s) \cong \frac{1}{n_s\sqrt{t}} \exp\left[\frac{\left(\ln n_s - 0.05 n_h t\right)^2}{0.5t}\right]$$

n_s = popularity of design s n_h = houses (population)

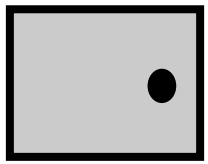






Bentley and Shennan (2003) *American Antiquity*
also Hamilton and Buchanan (2009) J. Anth Arch.

Limited lifespans



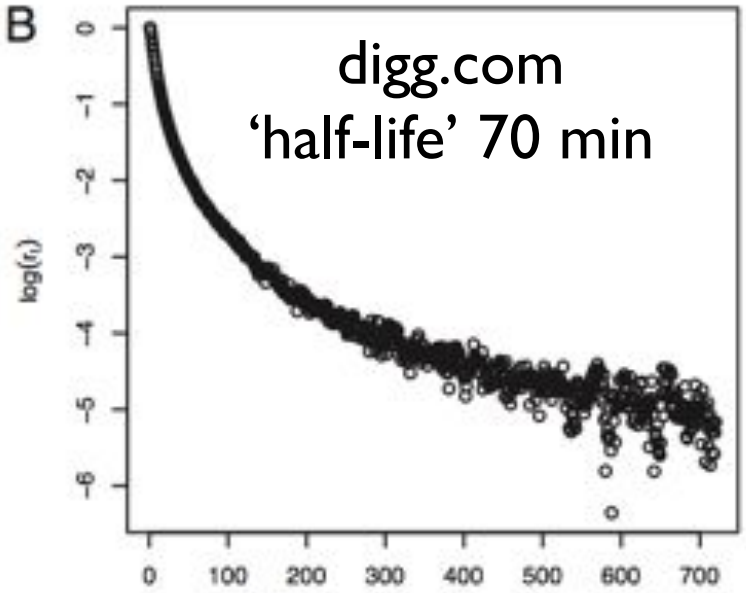
$$N_t = (1 + r_t g_t) N_{t-1}$$

$$r(t) \sim \exp[-0.4 t^{0.4}]$$

$$p(n|\mu,\sigma,\gamma) = \begin{cases} \int_{-\infty}^{\log_{10}(\gamma+1)} \frac{dq}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q-\mu)^2}{2\sigma^2}\right) & n = 0 \\ \int_{\log_{10}(n+\gamma)}^{\log_{10}(n+\gamma+1)} \frac{dq}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q-\mu)^2}{2\sigma^2}\right) & n \geq 1 \end{cases}$$

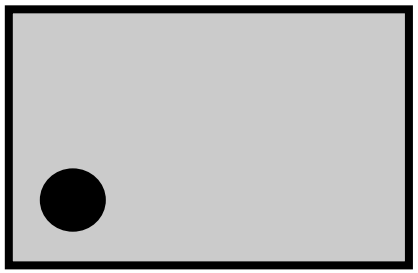
We surmise that the number of citations a paper receives is the result of a latent variable, q, and, thus, any paper with a value of q in some range will receive n citations (Burrell, 2001).

Stringer, Sales-Pardo & Amaral (2010) *J. Am. Soc. Inf. Sci. Tech*



Wu and Huberman (2007) *PNAS*

Southwest



$$\Pr(x = k) = \binom{k + r - 1}{k} p^k (1 - p)^r$$

Probability for *k* units bought among k+r total sales

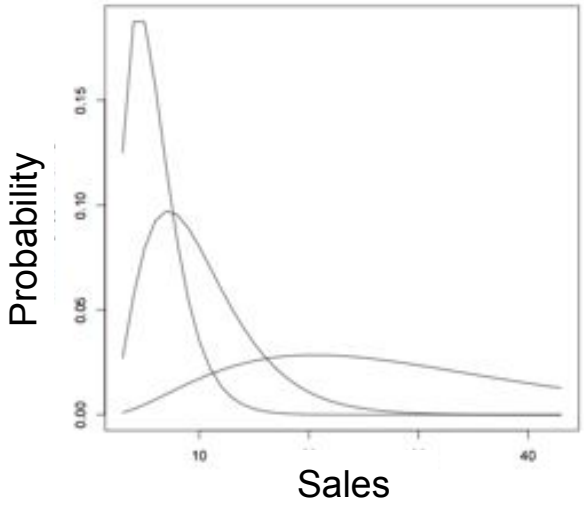
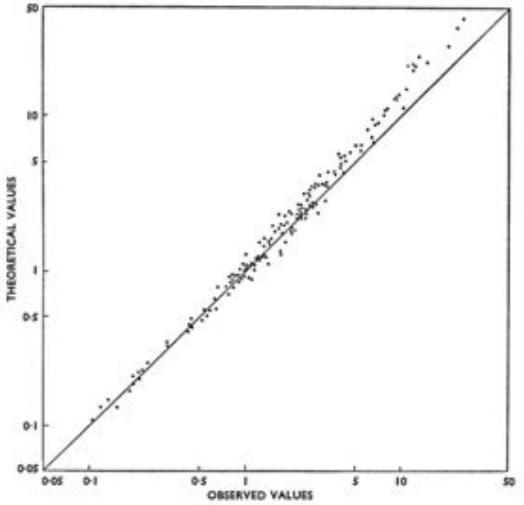


TABLE I

A typical example of the fit of a negative binomial distribution (26-weekly data for a 2000-household sample)

Number of Units Bought	Frequencies		Number of Units Bought	Frequencies	
	Observed	Theoretical		Observed	Theoretical
0	1612	1612	14	–	1·8
1	164	156·9	15	–	1·5
2	71	74·0	16	–	1·2
3	47	44·2	17	2	0·9
4	28	29·2	18	–	0·8
5	17	20·3	19	–	0·6
6	12	14·7	20	1	0·5
7	12	10·8	21	–	0·4
8	5	8·2	22	2	0·3
9	7	6·2	23	–	0·3
10	6	4·8	24	–	0·2
11	3	3·8	25	1	0·2
12	3	2·9	26	2	0·1
13	5	2·3	27+	–	0·9

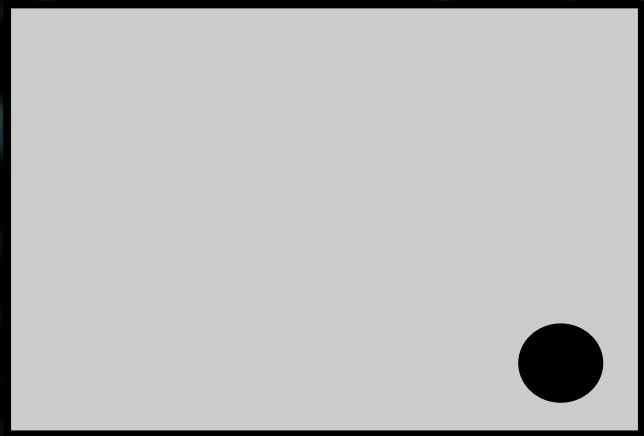
$m = 0{\cdot}636$, $p_0 = 0{\cdot}806$, $k = 0{\cdot}115$, $a = 5{\cdot}53$
Standard Deviations: Root-mean-square 2·12, $\sqrt{[m(1+a)]} = 2{\cdot}04$



Ehrenberg (1959) *J. Roy. Stat. Soc C* 8: 26-41.

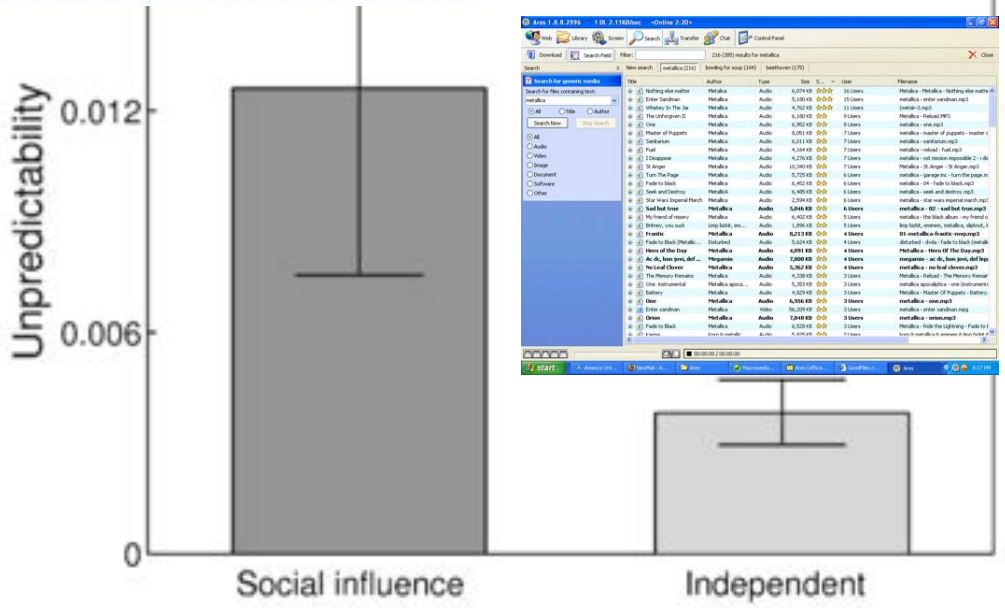
Social poorly-informed*

**Or only a small minority informed*
(Couzin *et al.* 2005, *Nature*)



Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market *Science*, 2006

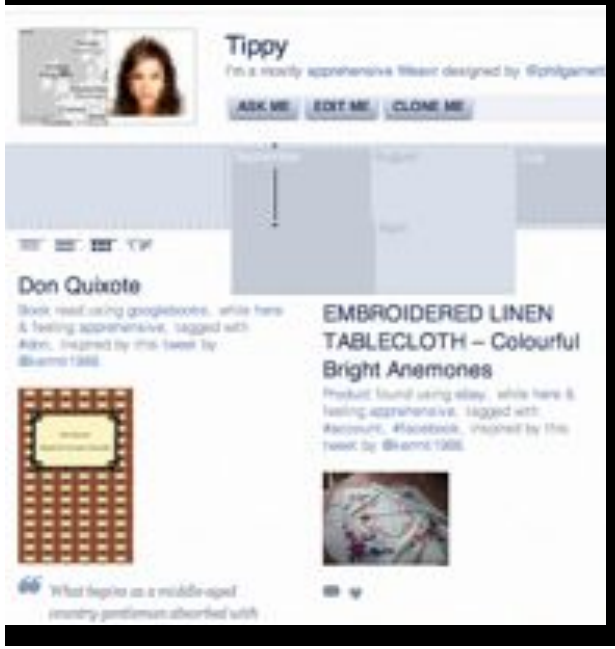
Matthew J. Salganik,[1,2*] Peter Sheridan Dodds,[2*] Duncan J. Watts[1,2,3*]

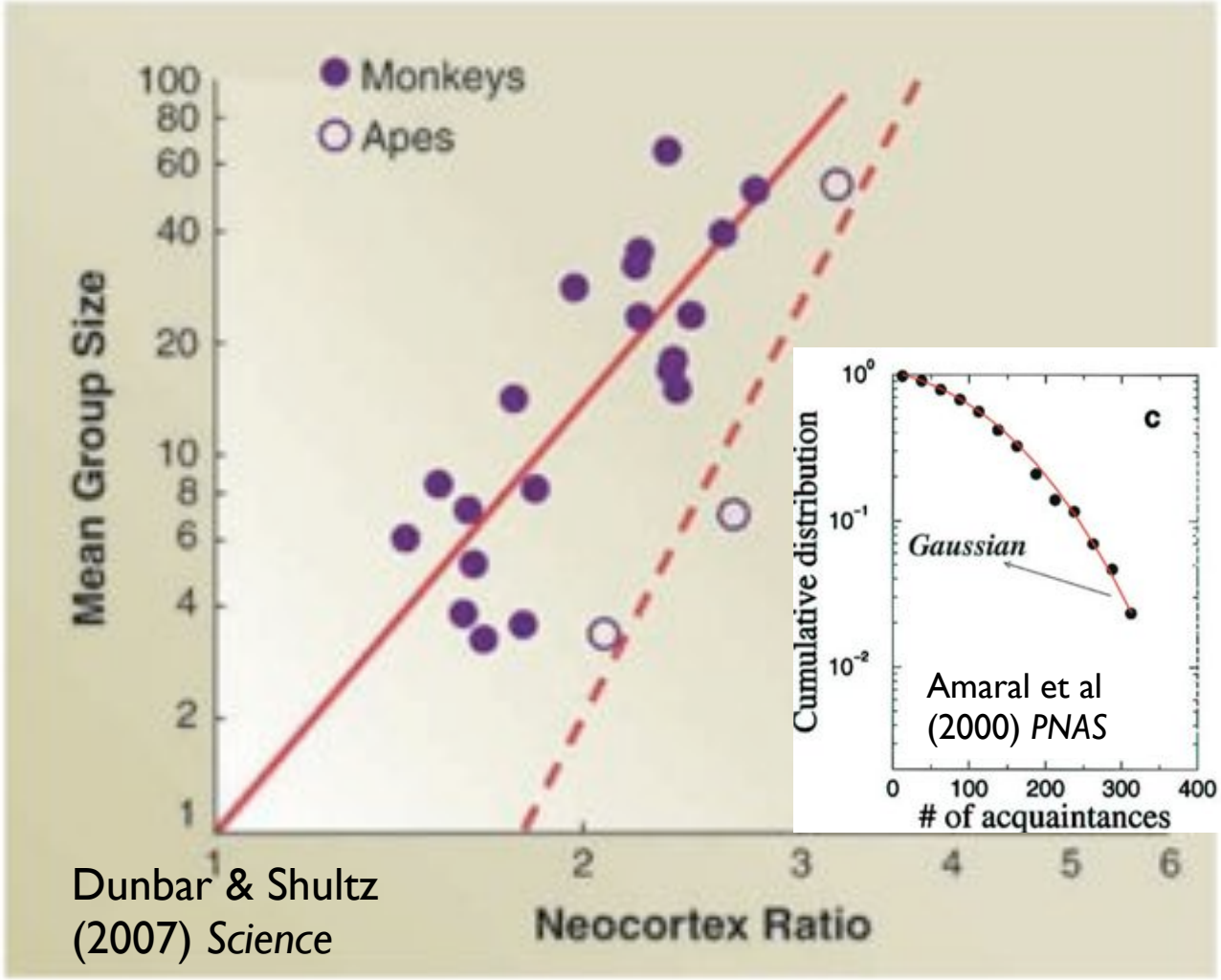


...modern writing at its worst … consists in gumming together long strips of words which have already been set into order by someone else – George Orwell

Too many people to copy

⬆ Then vs now ⬇





Dunbar & Shultz (2007) *Science*

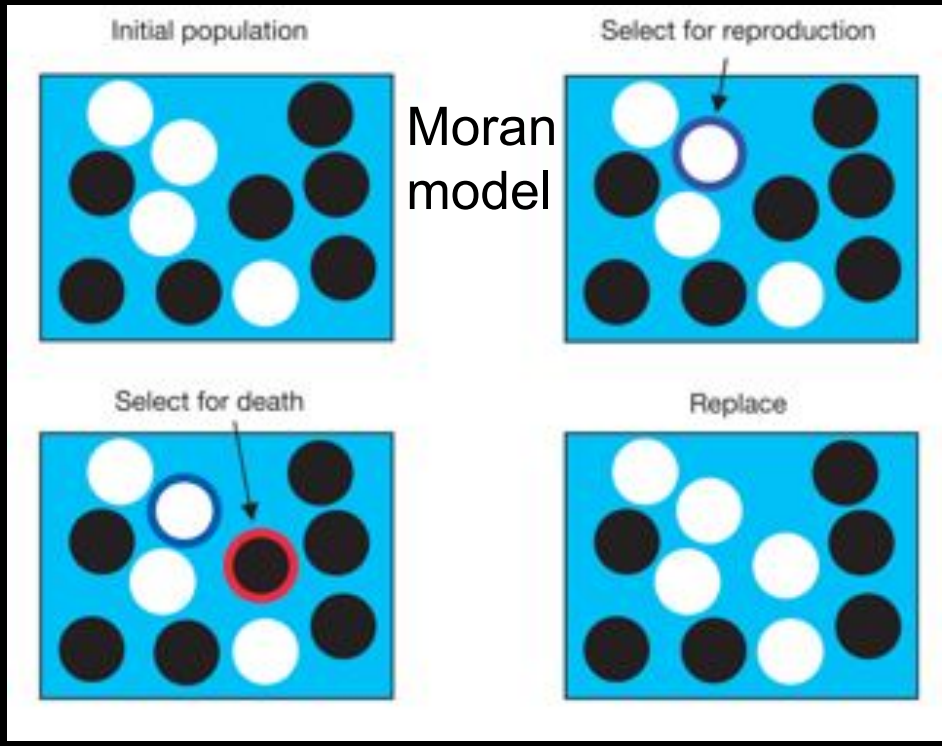
Amaral et al (2000) *PNAS*

Neutral copying models

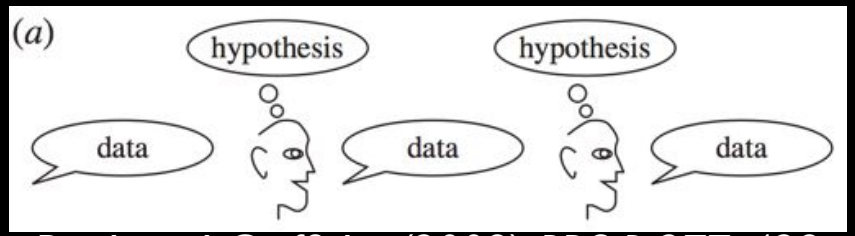
Hahn and Bentley (2003) *Biology Letters*
Bentley, Hahn, Shennan (2004). *PRS B*
Bentley, Ormerod Batty (2010) *Evol. Biol. & Sociobiol.*

Ijiri & Simon (1964). *Am Econ Rev* 54: 77

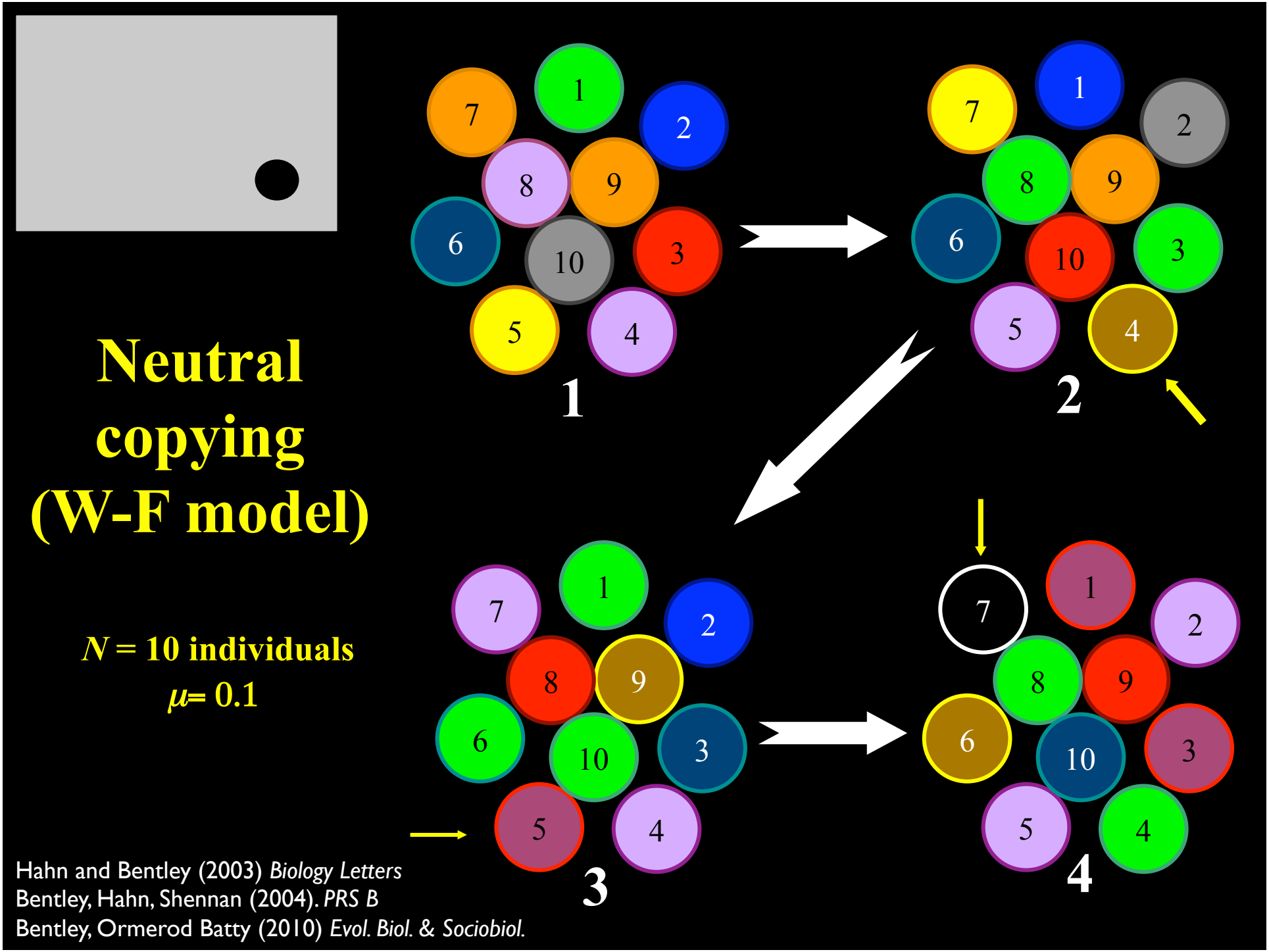
Kimura & Crow (1964) *Genetics*
Yule (1924) *F.R.S. Phil. Trans.*
Neiman F (1995). *Am Antiq.*



Lieberman, Hauert & Nowak (2005) *Nature*



Reali and Griffiths (2009) *PRS B* 277: 429


Neutral copying (W-F model)
N = 10 individuals
μ= 0.1
1
2
3
4
Hahn and Bentley (2003) Biology Letters
Bentley, Hahn, Shennan (2004). PRS B
Bentley, Ormerod Batty (2010) Evol. Biol. & Sociobiol.

Neutral copying

(1) Stochastic change $V = \frac{\nu(1-\nu)}{2N}$

(2) Variety of skewed distributions

(3) Diversity $x \approx 4N\mu$

(4) Turnover $z = y\sqrt{\mu}$

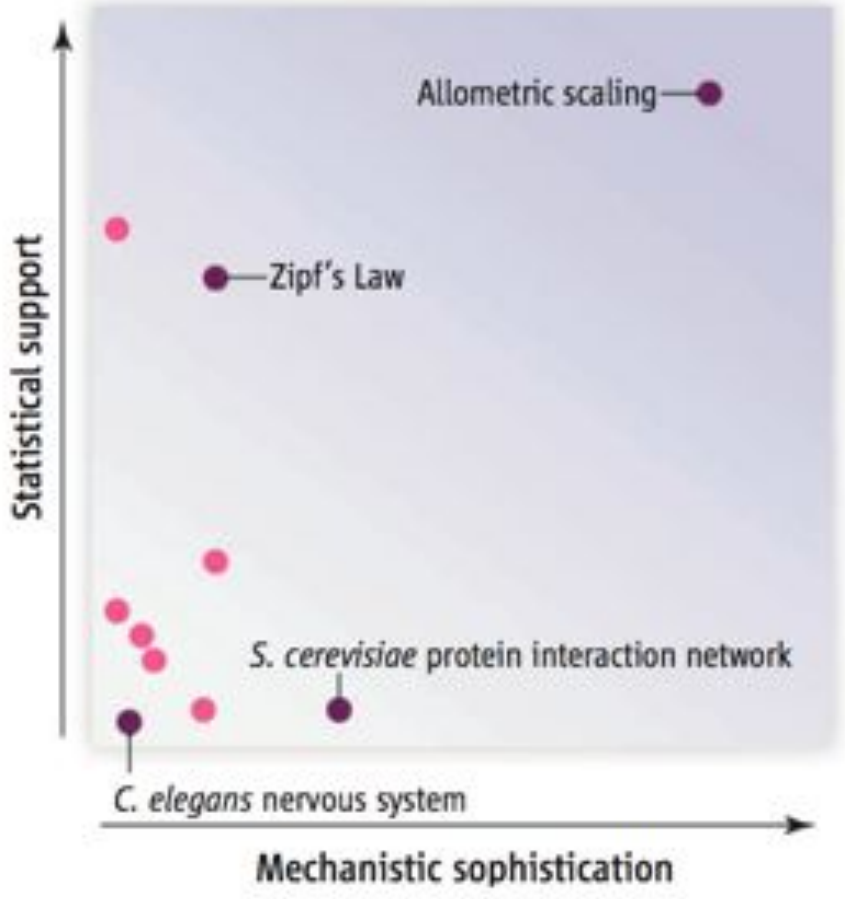
N = population size, μ = innovation fraction; ν = variant frequency;
x = no. different variants; z = turnover rate; V = variance in frequency over time

Hahn and Bentley (2003) *Biology Letters;* Bentley, Hahn, Shennan (2004) *PRS B*
Erikson et al. (2010) *Ramanujan Journal*

No power law claims - we compare (by Anderson-Darling test) averaged modeled distributions to data

Critique goes back a long way

Newman (2005) *Contemporary Physics* 46: 323-51.
Clauset et al. (2009) *SIAM Review* 51:661-703.
Frank (2009) J. Evol. Biol. 22: 1563–1585
Mitzenmacher (2008) *Internet Mathematics* 1: 226-51.
Laherrère & Sornette (1998) *European Physical Journal B* 2: 525-39.
Newman (1996) *Proc. R. Soc. Lond. B* 263: 1605-10.

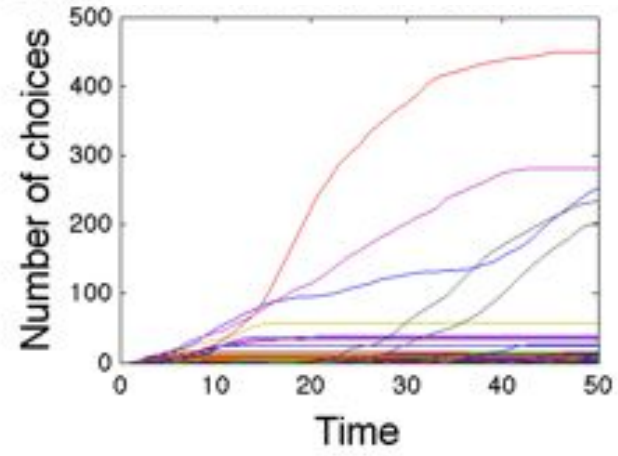


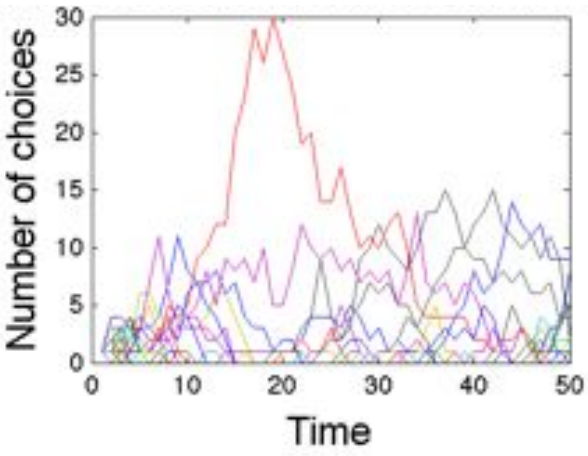
How good is your power law?

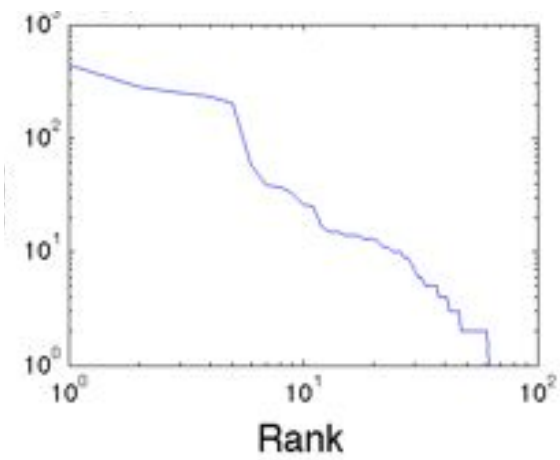
Stumpf, M.P.H. & M.A. Porter (2012) *Science* 335: 665-6

Bassel Tarbush
(Ox. Econ PhD prog)

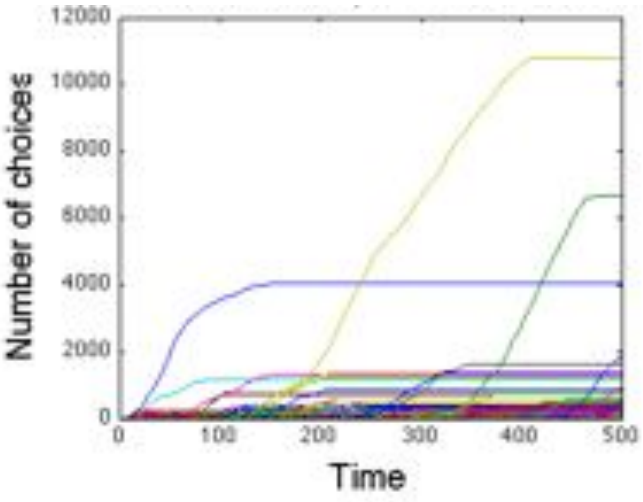
Neutral model N=40, μ = 2.5%

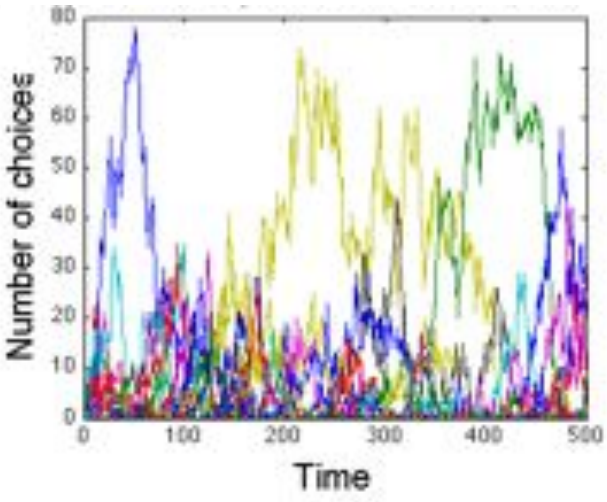


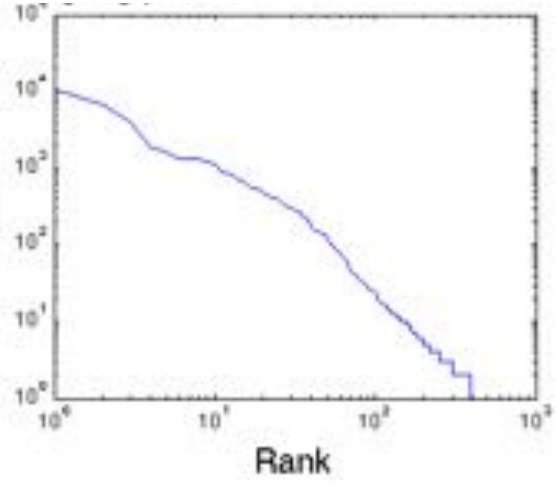




Neutral model N=500, μ = 1%



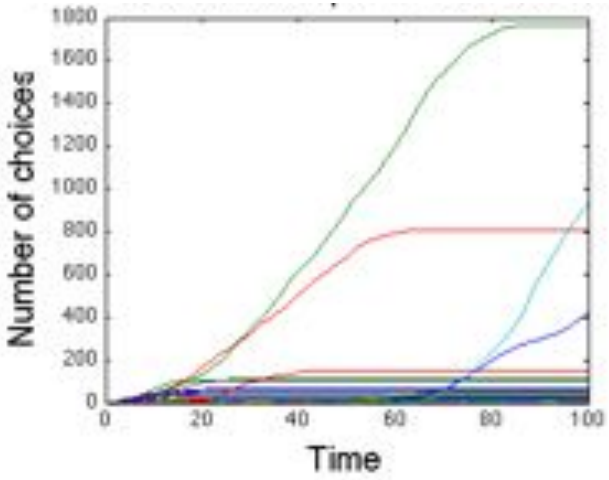


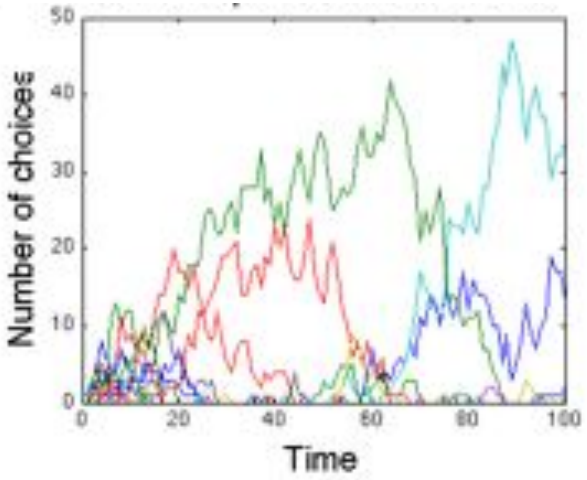


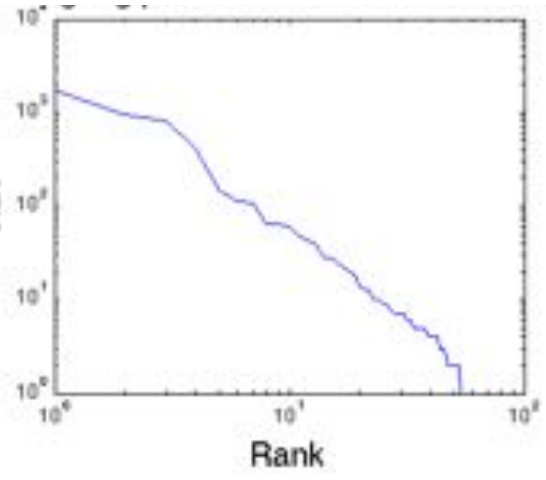
Neutral model N=50, μ = 1%

Bassel Tarbush

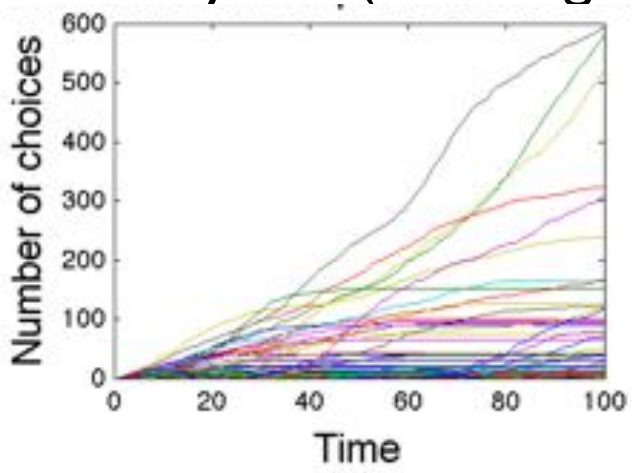
memory = 1

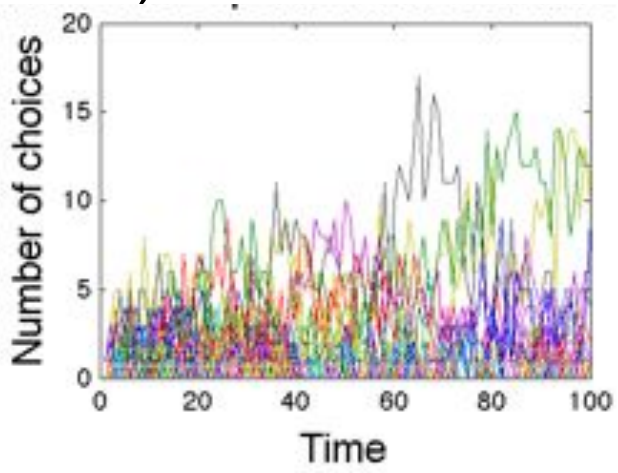


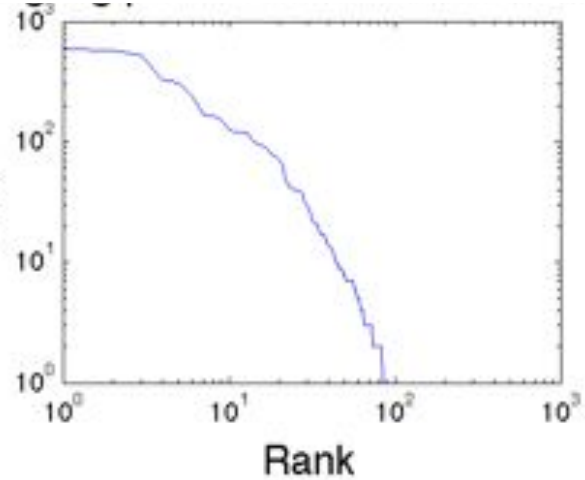




memory = 5 (more egalitarian)

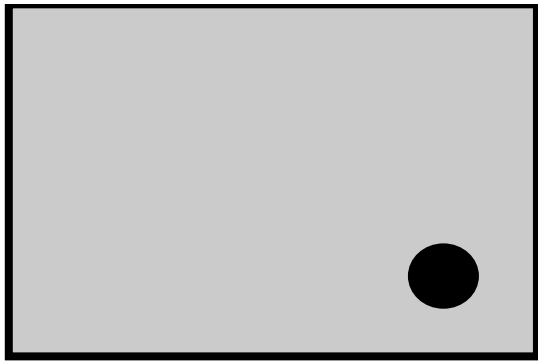






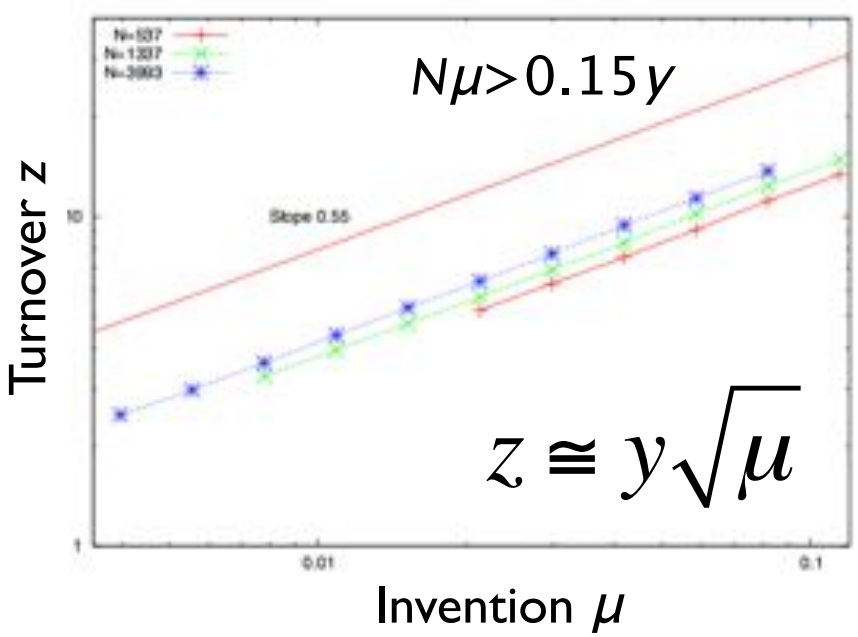
Turnover

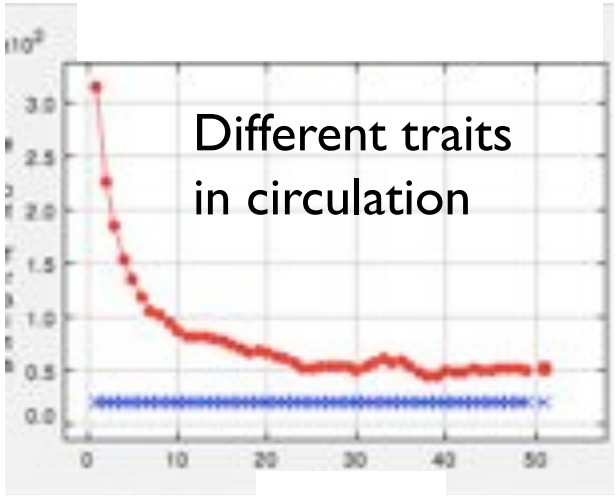
Conjecture by Bentley et al. (2007) *Evol. Hum Behav*.



Turnover Rate of Popularity Charts in Neutral Models

T.S. Evans[1,2] and A. Giometto[1]



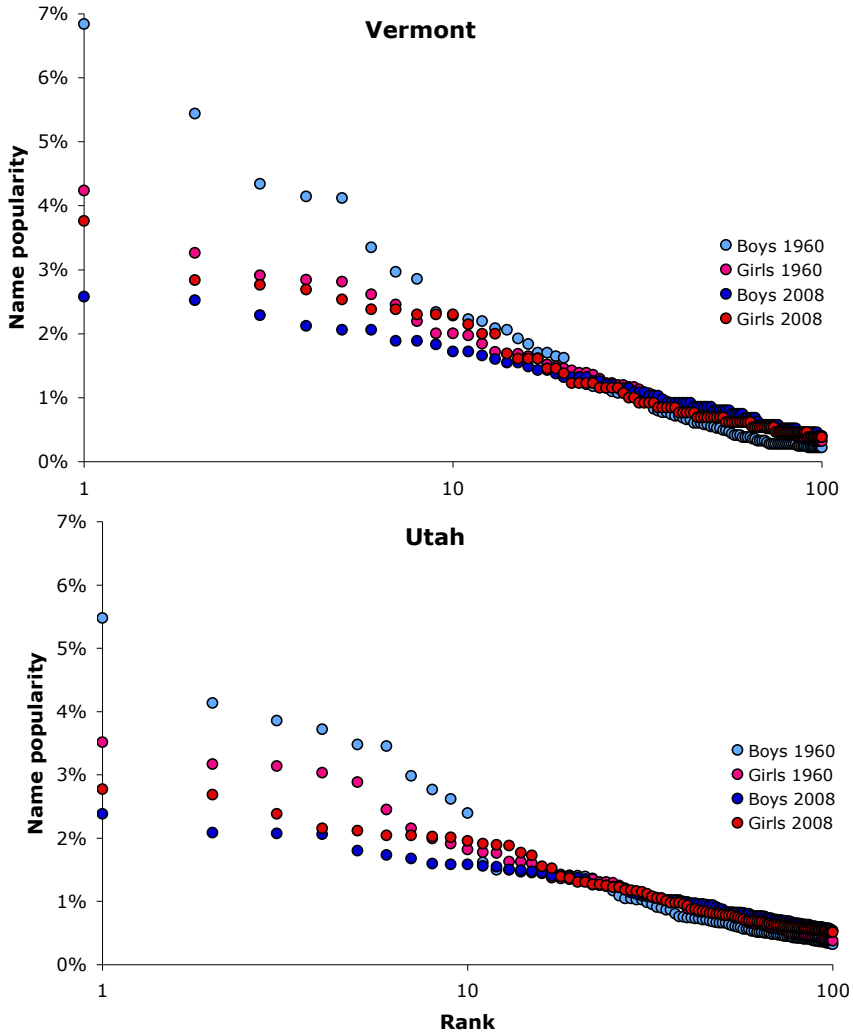


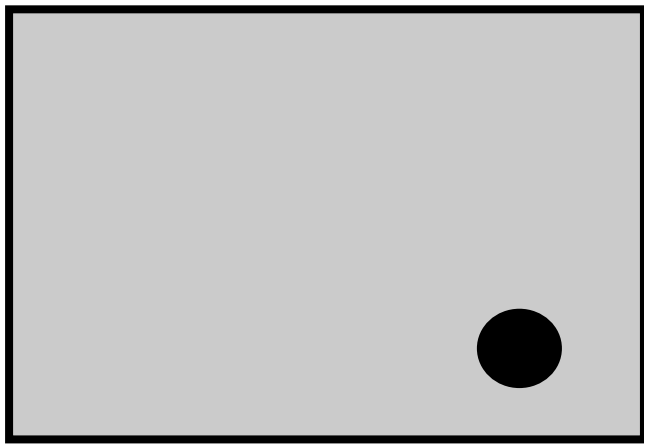
Bentley's conjecture on popularity toplist turnover under random copying

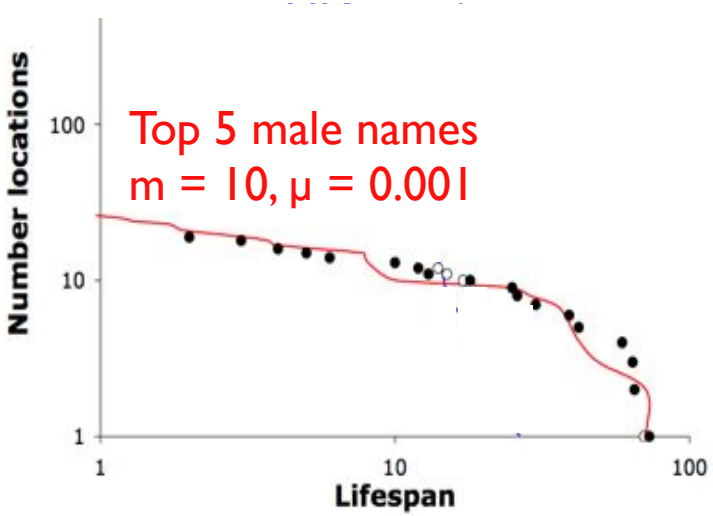
Ramanujan J (2010)

Kimmo Eriksson · Fredrik Jansson · Jonas Sjöstrand

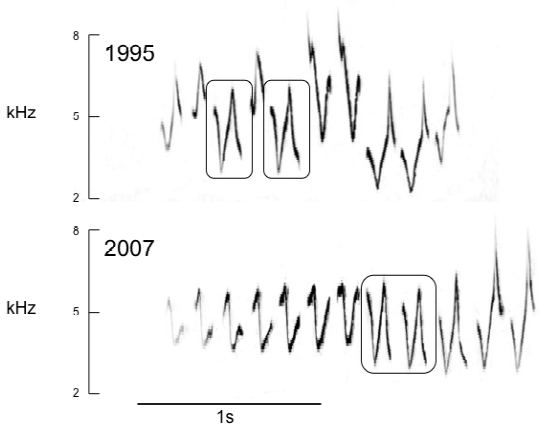
Distributions → Lifespans





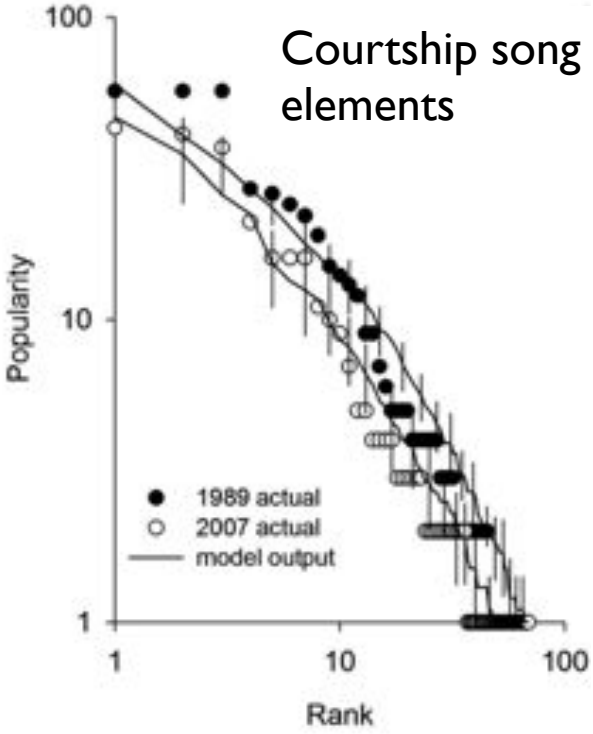


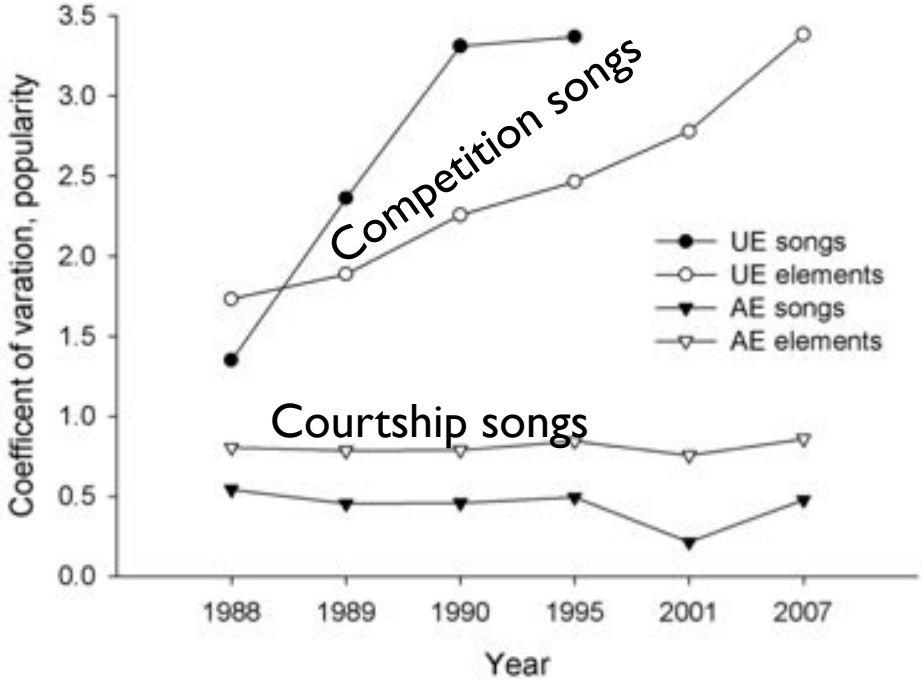
Bentley & Ormerod 2011. *Adv. Comp Syst.*
Bentley, Ormerod, Batty (2011) *Behav. Ecol. & Sociobiol.*



Bruce Byers, U. Mass-Amherst

Male competition songs: drift
Courtship songs: selection



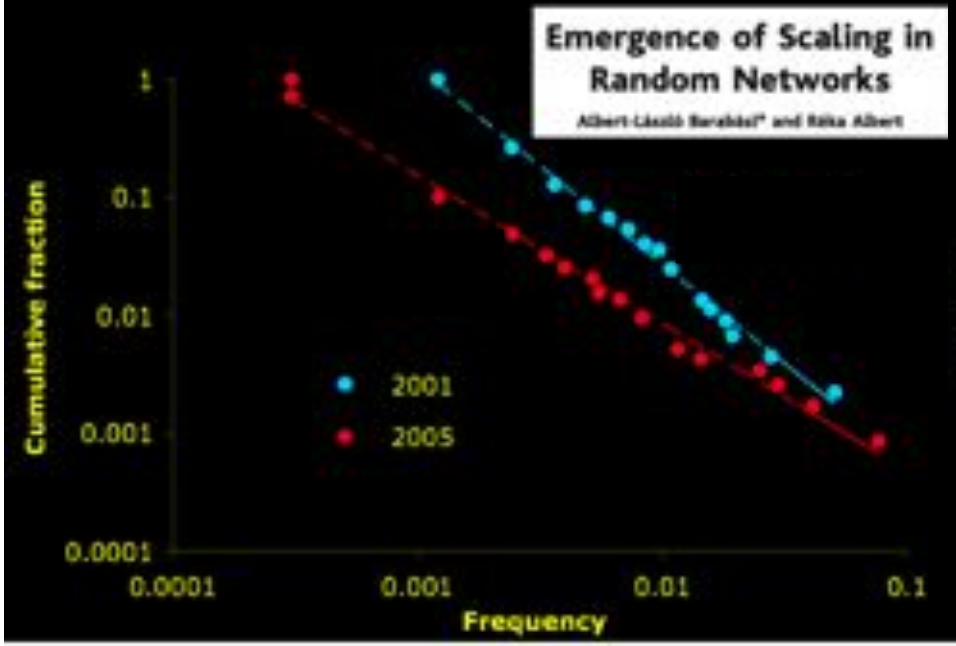


Independent Cultural Evolution of Two Song Traditions in the Chestnut-Sided Warbler

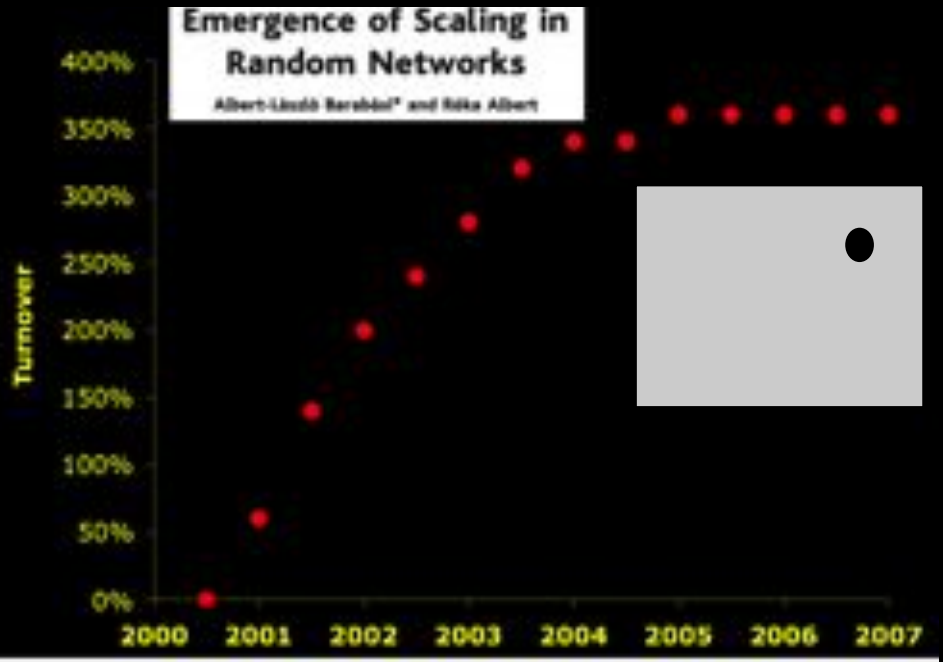
American Naturalist, 2010

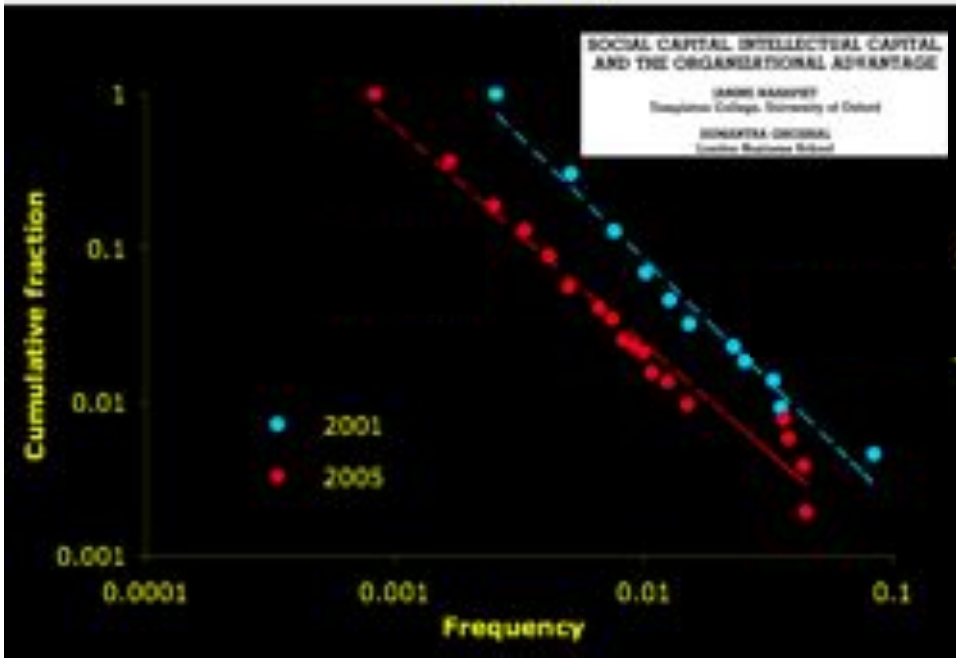
Bruce E. Byers,[1,*] Kara L. Belinsky,[2] and R. Alexander Bentley[3]

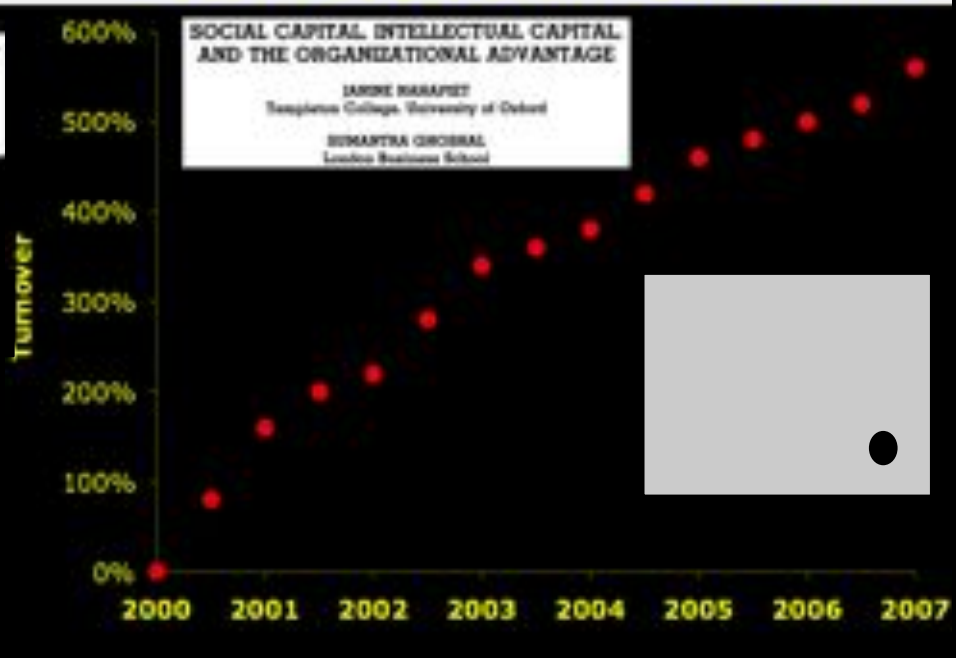
Keyword popularity



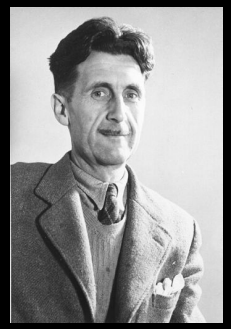
Turnover, top 5 keywords

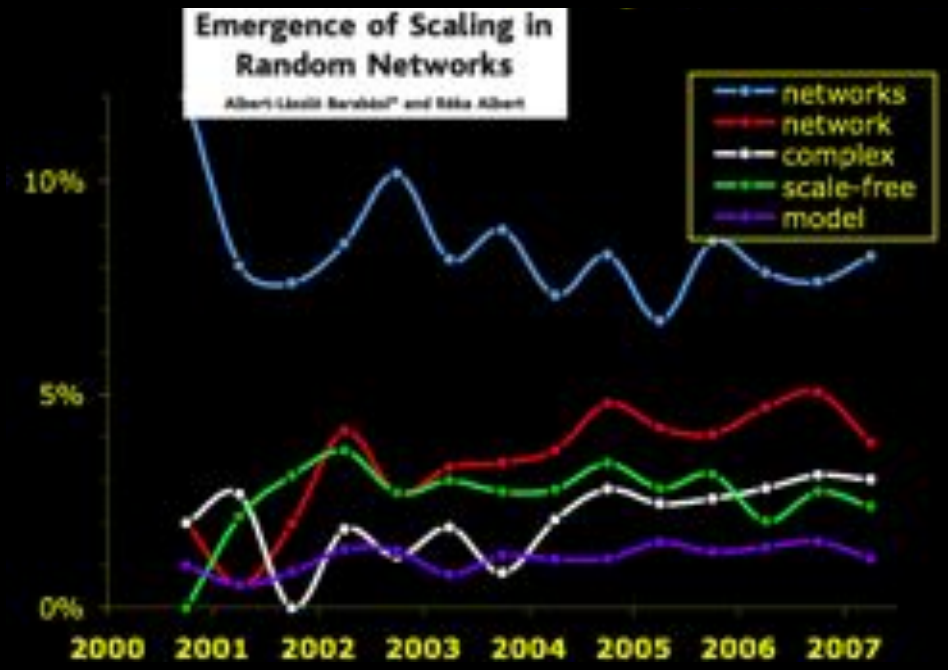




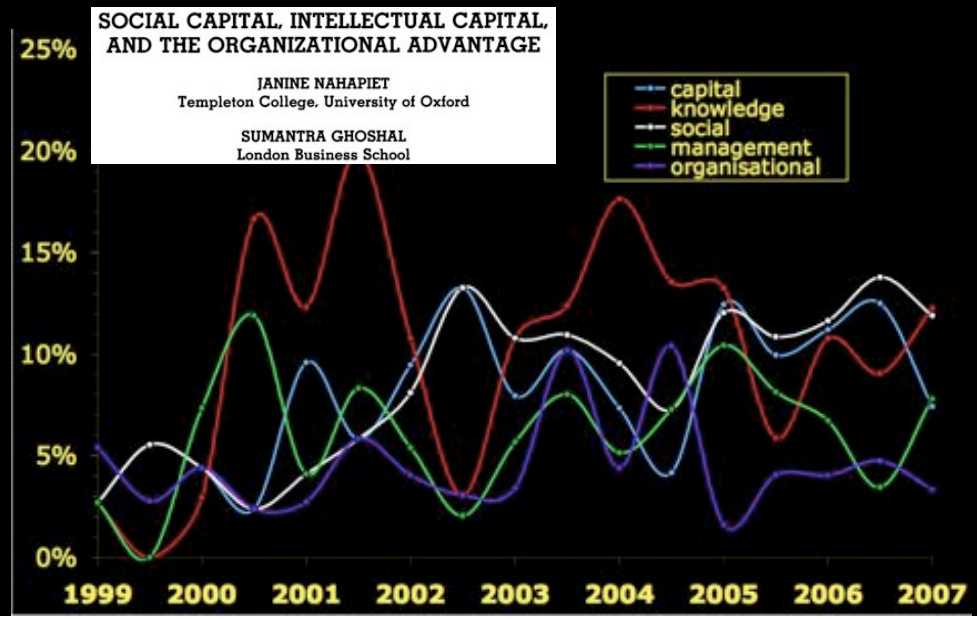


Selection



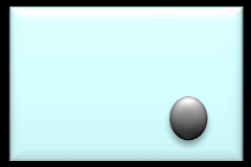


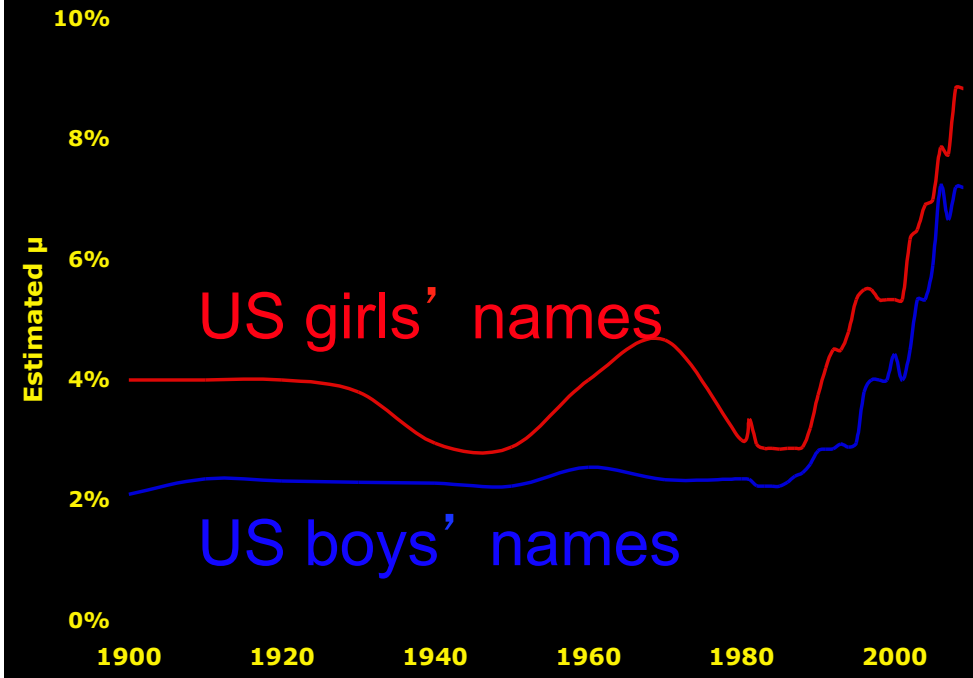
Drift

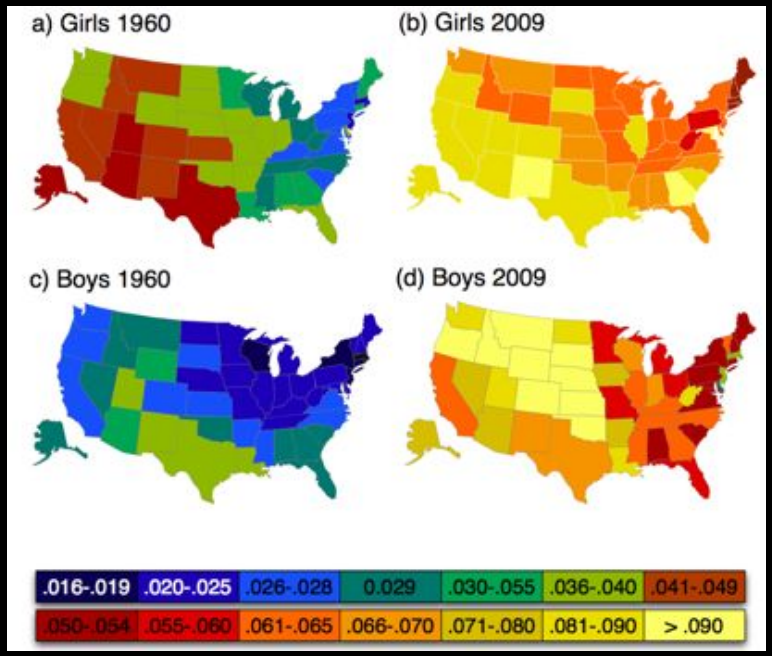


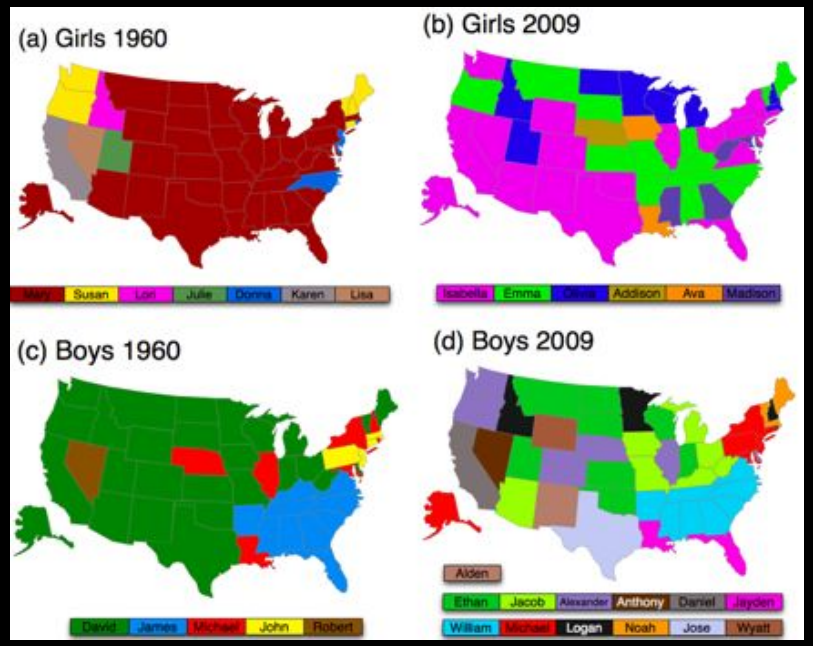
Bentley (2008) *PLoS ONE* 3(8): e3057

More novelty





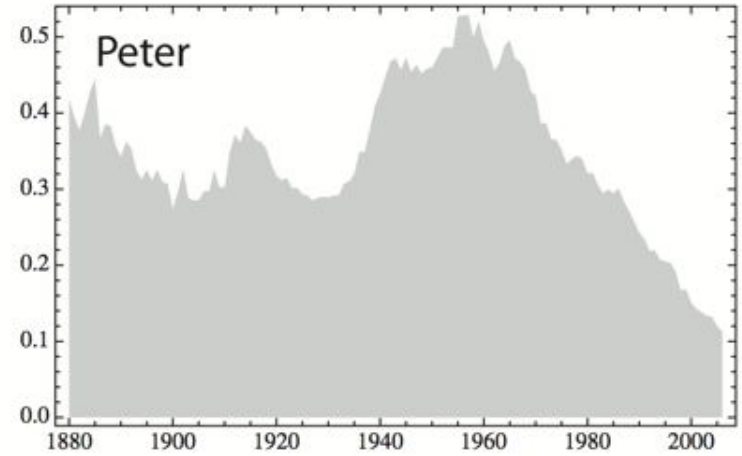


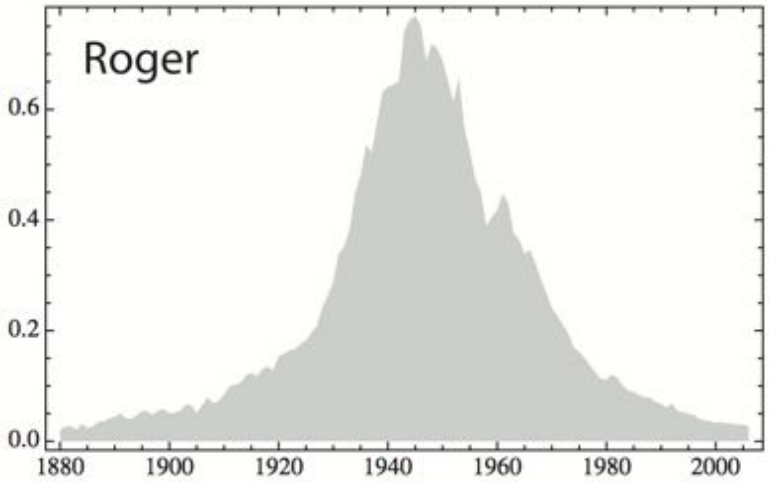


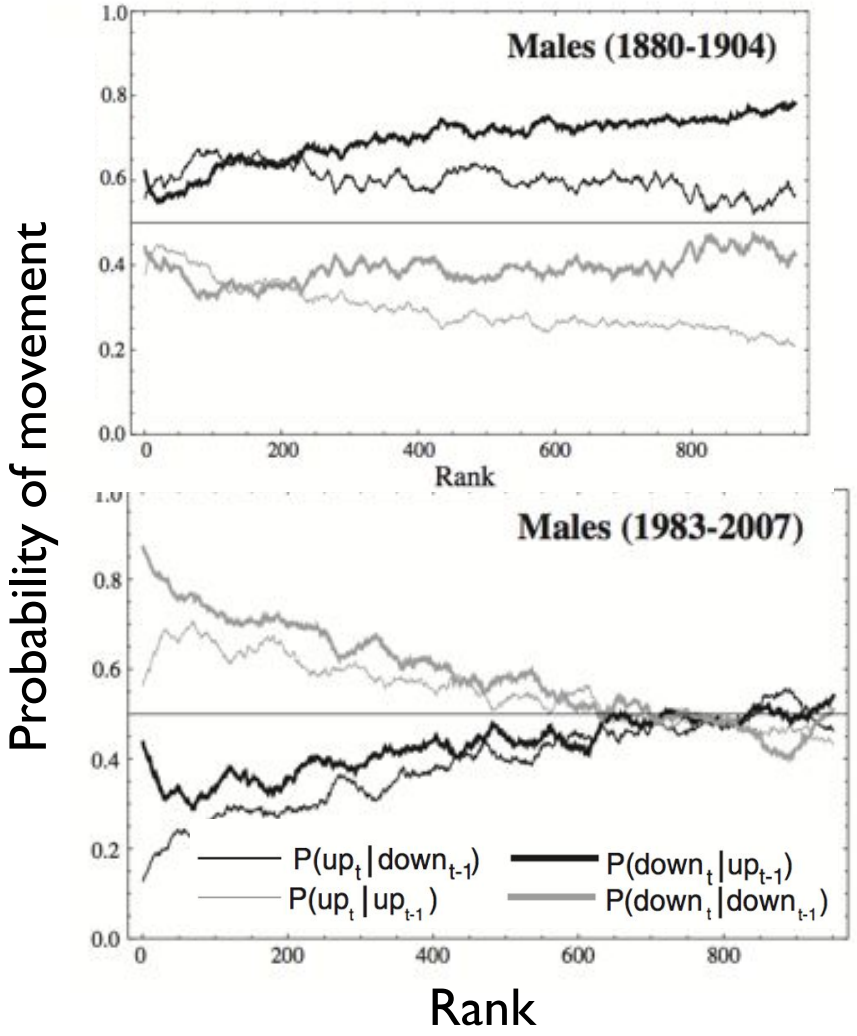
- Data vs. 500 model solutions for 125 parameters comb ($\mu < 10\%, m \leq 10$)
- Anderson-Darling for parameter combinations (~ 100), null rejected at $p < 0.05$
- Accepted the (typically several) combinations where $p > 0.2$

Bentley & Ormerod (2011) *Advances in Complex Systems*

What about diffusion curves?



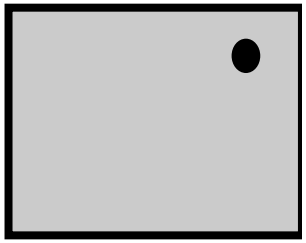




Gureckis and Goldstone (2009) *Topics Cog Sci*
Berger and Le Mens (2009)

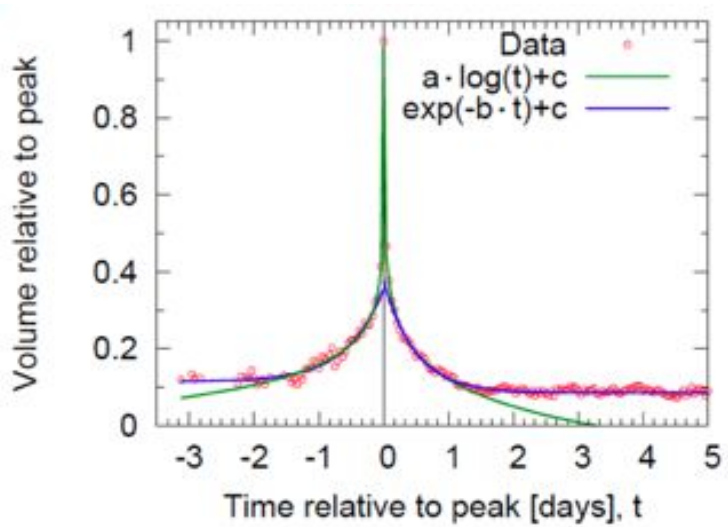
} Advocate trend *perception*

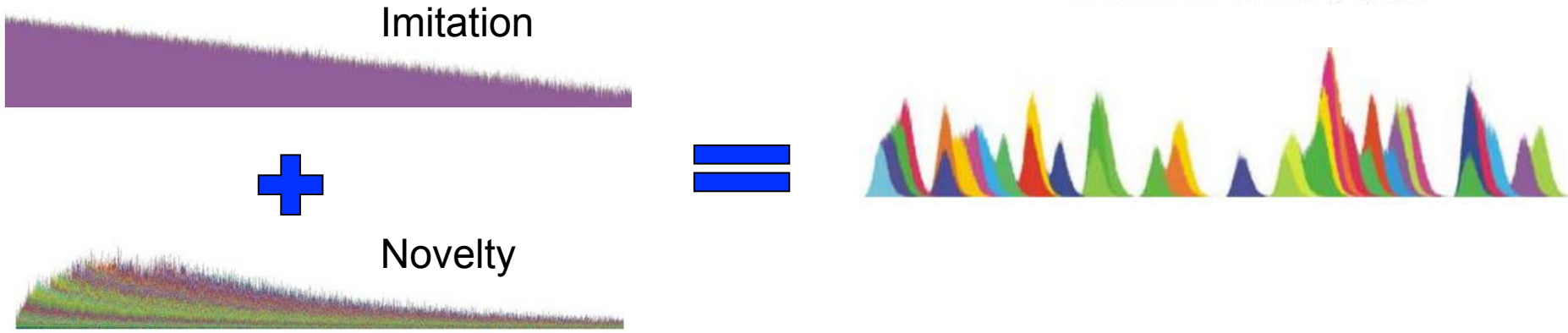
Imitation plus exp. decay



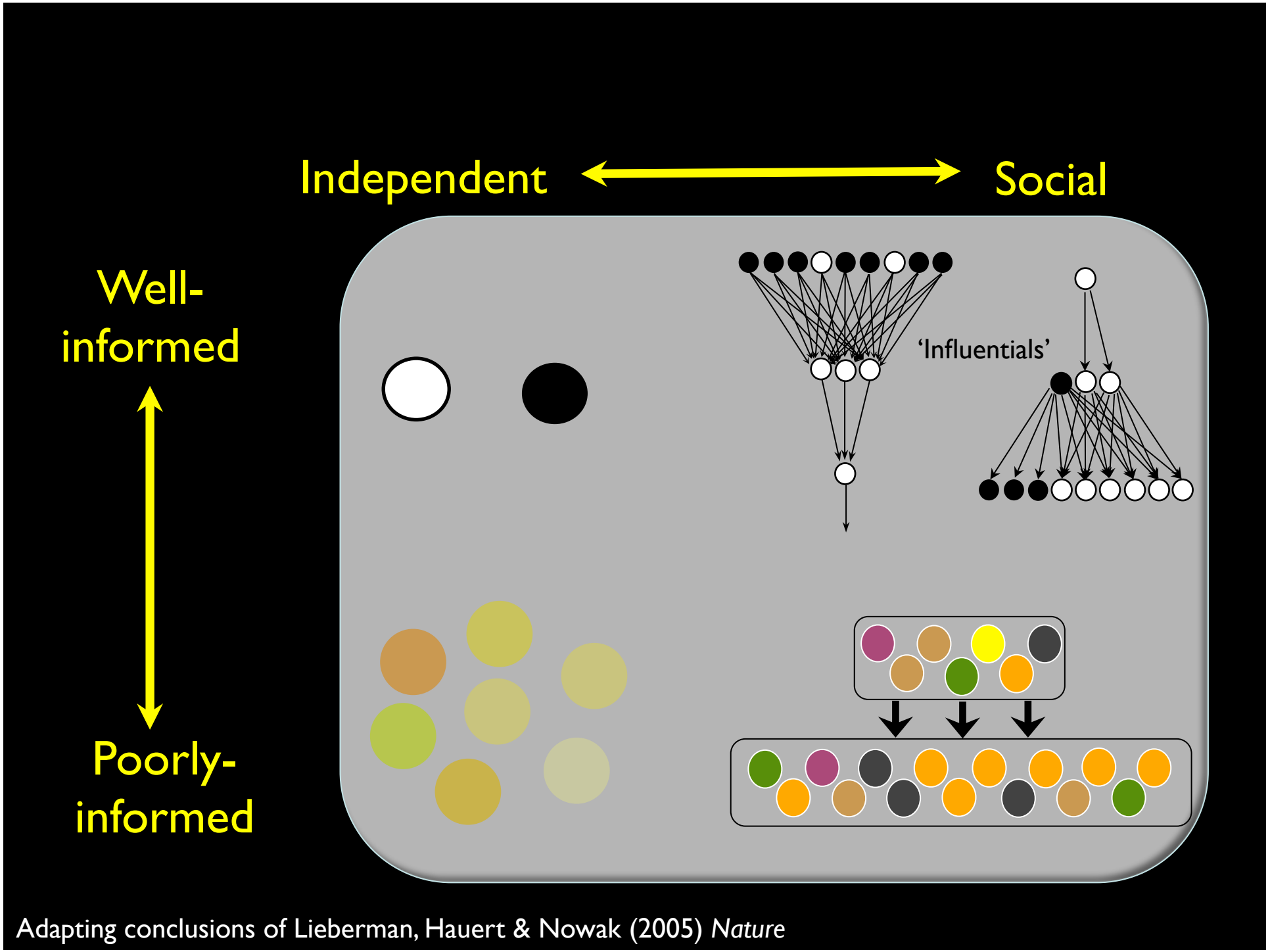
Leskovic, Backstrom, Kleinberg (2009)

$$n_t = At^{cq}e^{-t}$$

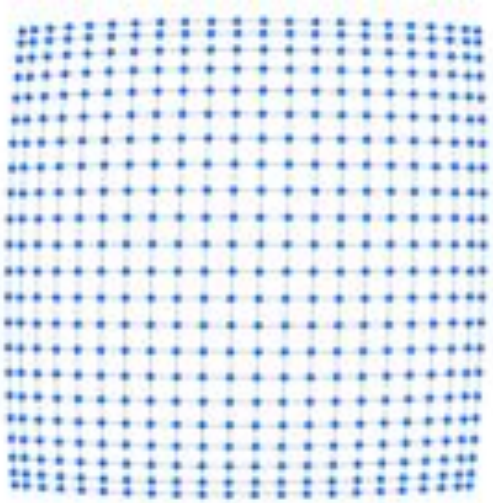




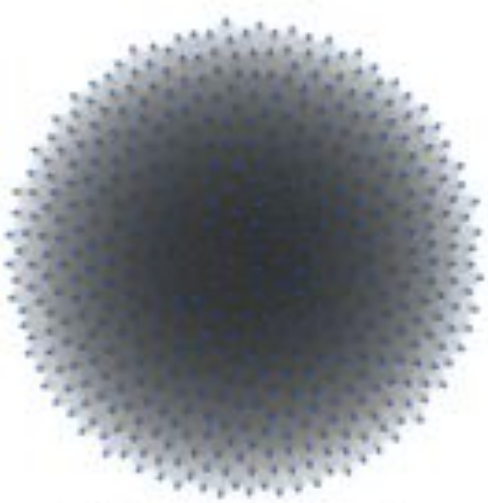
Also Gureckis &Goldstone (2009) *Topics Cog. Sci*



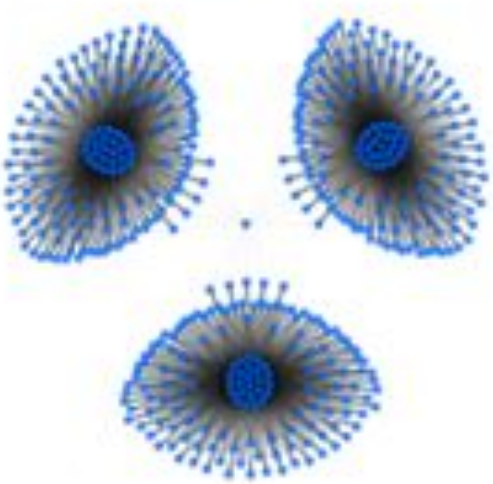
Adapting conclusions of Lieberman, Hauert & Nowak (2005) *Nature*



(a) Squarelattice n = 22



(b) Connected network n = 475

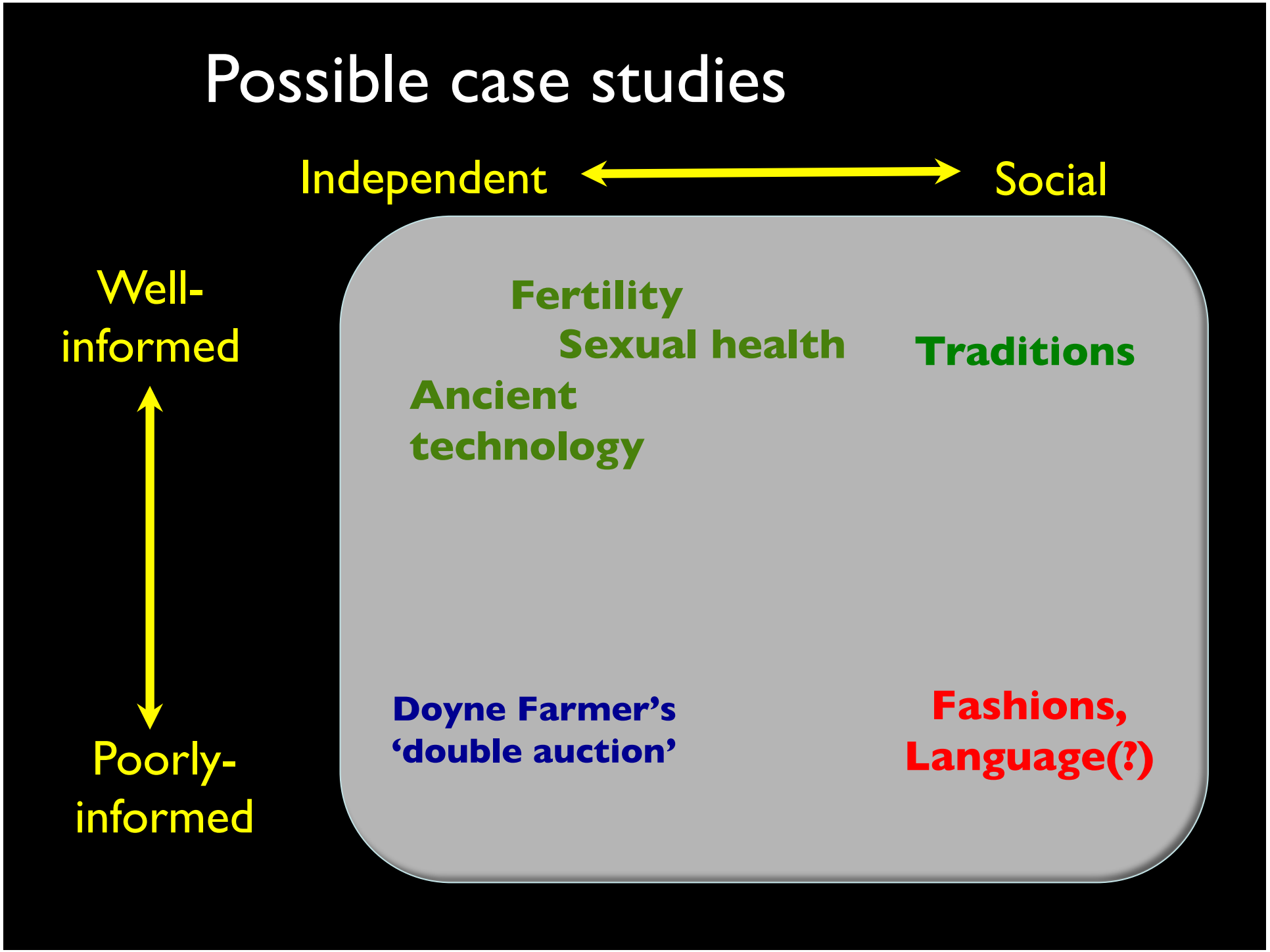


Neutral model on 'drift-promoting' and 'selection-promoting' networks:

(a) Popularity distributions
(b) Lifespan distributions
(c) Turnover in top y
(d) 'Bass' diffusions?

Bassel Tarbush

Possible case studies



Philip Garnett
Durham U.

Michael O'Brien
U. Missouri

Alberto Acerbi
Stockholm U.

Paul Ormerod
Volterra Consulting

William Brock
U. Wisconsin

Mark Earls
Anomaly Communications

Stephen Shennan
UCL

Dirk Brockmann
Northwestern U.

Dan Hruschka
Arizona State U.

David Robertson
U. Manchester, U.K.

Bassel Tarbush
Oxford U.

Matthew Hahn
Indiana U.