“I always dream of a pen that would be a syringe.” — Jacques Derrida

BEHAVIORAL NETWORK SCIENCE

Thomas Hills
University of Warwick

Hohenheim 2023

WARWICK
THE UNIVERSITY OF WARWICK

The Alan Turing Institute

THE ROYAL SOCIETY

What we've covered so far

- How to represent networks
- Network metrics
- Comparing networks
- Null models of networks (ER random graph, etc)
- Fitting models to networks

Today

- Agent-based modelling (opinion dynamics)
- Proof-of-principle simulations (aging mind)

Social contagion

- We know one of the primary influences on behaviour is the behaviour of others (Social proof)

- People also select for information that is consistent with what they already know (confirmation bias, motivated reasoning, biased assimilation)

Perspectives on Psychological Science
1–8

DOI: 10.1177/1745691618803647
www.psychologicalscience.org/PPS
SAGE

The Dark Side of Information Proliferation

Thomas T. Hills
Department of Psychology, University of Warwick

Abstract
There are well-understood psychological limits on our capacity to process information. As information proliferation—the consumption and sharing of information—increases through social media and other communications technology, these limits create an attentional bottleneck, favoring information that is more likely to be searched for, attended to, comprehended, encoded, and later reproduced. In information-rich environments, this bottleneck influences the evolution of information via four forces of cognitive selection, selecting for information that is belief-consistent, negative, social, and predictive. Selection for belief-consistent information leads balanced information to support increasingly polarized views. Selection for negative information amplifies information about downside risks and crowds out potential benefits. Selection for social information drives herding, impairs objective assessments, and reduces exploration for solutions to hard problems. Selection for predictive patterns drives overfitting, the replication crisis, and risk seeking. This article summarizes the negative implications of these forces of cognitive selection and presents eight warnings that represent severe pitfalls for the naive “informavore,” accelerating extremism, hysteria, herding, and the proliferation of misinformation.

Bayesian Truth Serum

- “What do other people think?”
- Second-order inference

False consensus effect

Most people tend to overestimate the number of people who share their views.

Ross, Greene, and House (1977) found is that people’s beliefs about others were strongly biased in favor of their own views.

People who themselves favored public spending on an unmanned space program believed that view was more popular than it actually was, and vice versa.

Can we model this as a function of opinion dynamics?

Schelling Segregation Model

(People only need a preference for equal representation to get segregation)

Two individuals chosen at random, swap positions if they are in the minority.

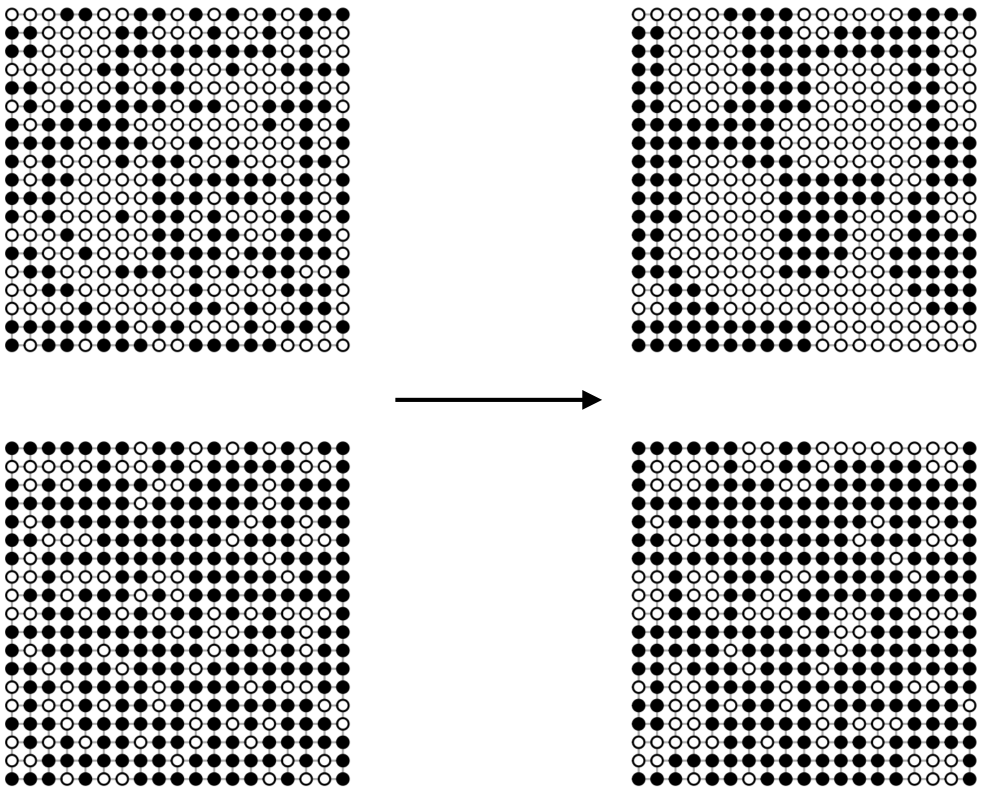
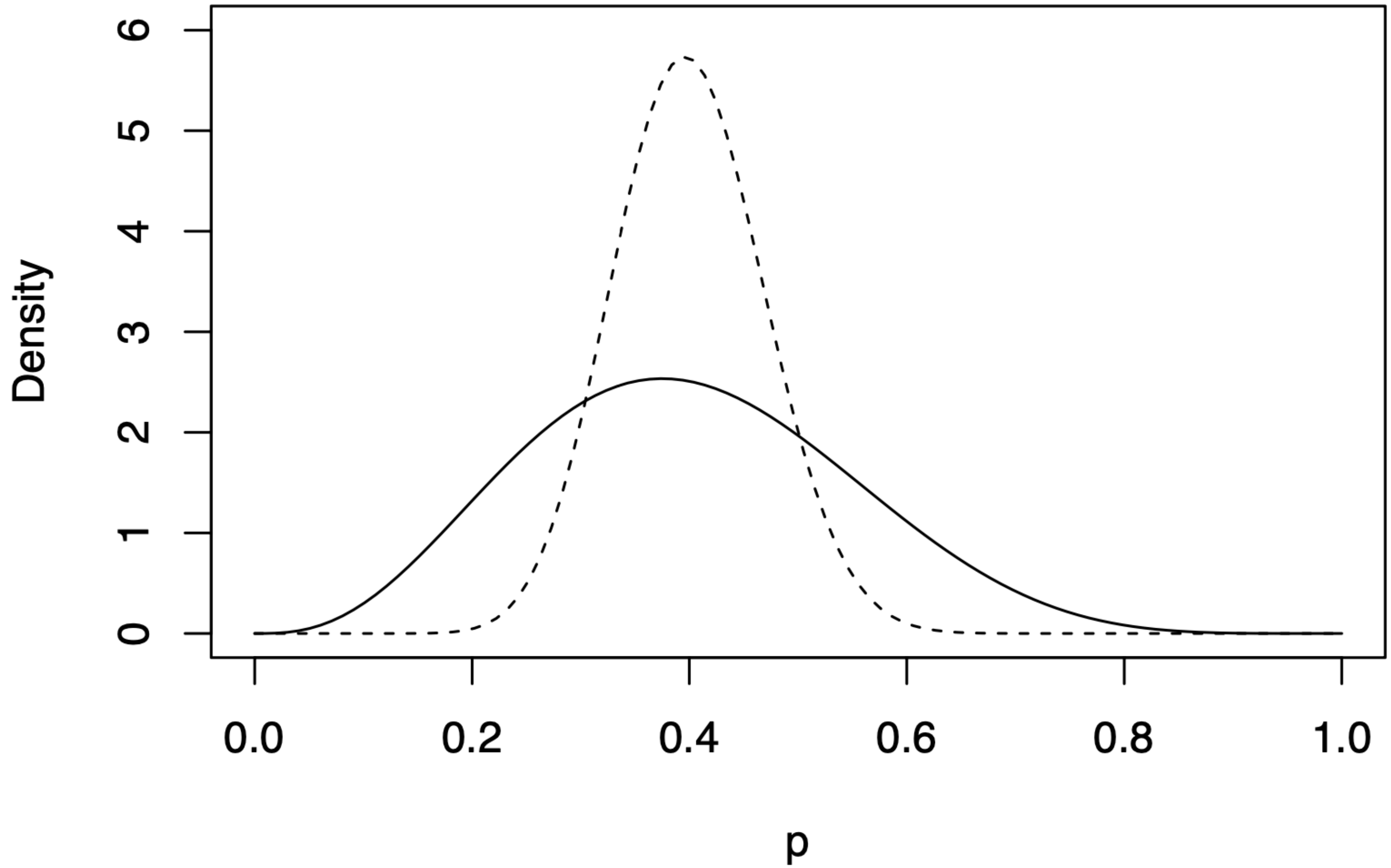


Figure 1: Thomas Schelling's segregation model. All agents prefer to move if they are in the minority color among their neighbors. At each timestep, one white and one black agent who are willing to move are chosen at random and swapped. Panels on the left show the random starting arrangements. Panels on the right show the arrangments after all available swaps are made to the corresponding panel on the left. The top panels show a 50:50 distribution of white to black nodes. The bottom panels are 30:70, white to black.

Social Sampling Theory

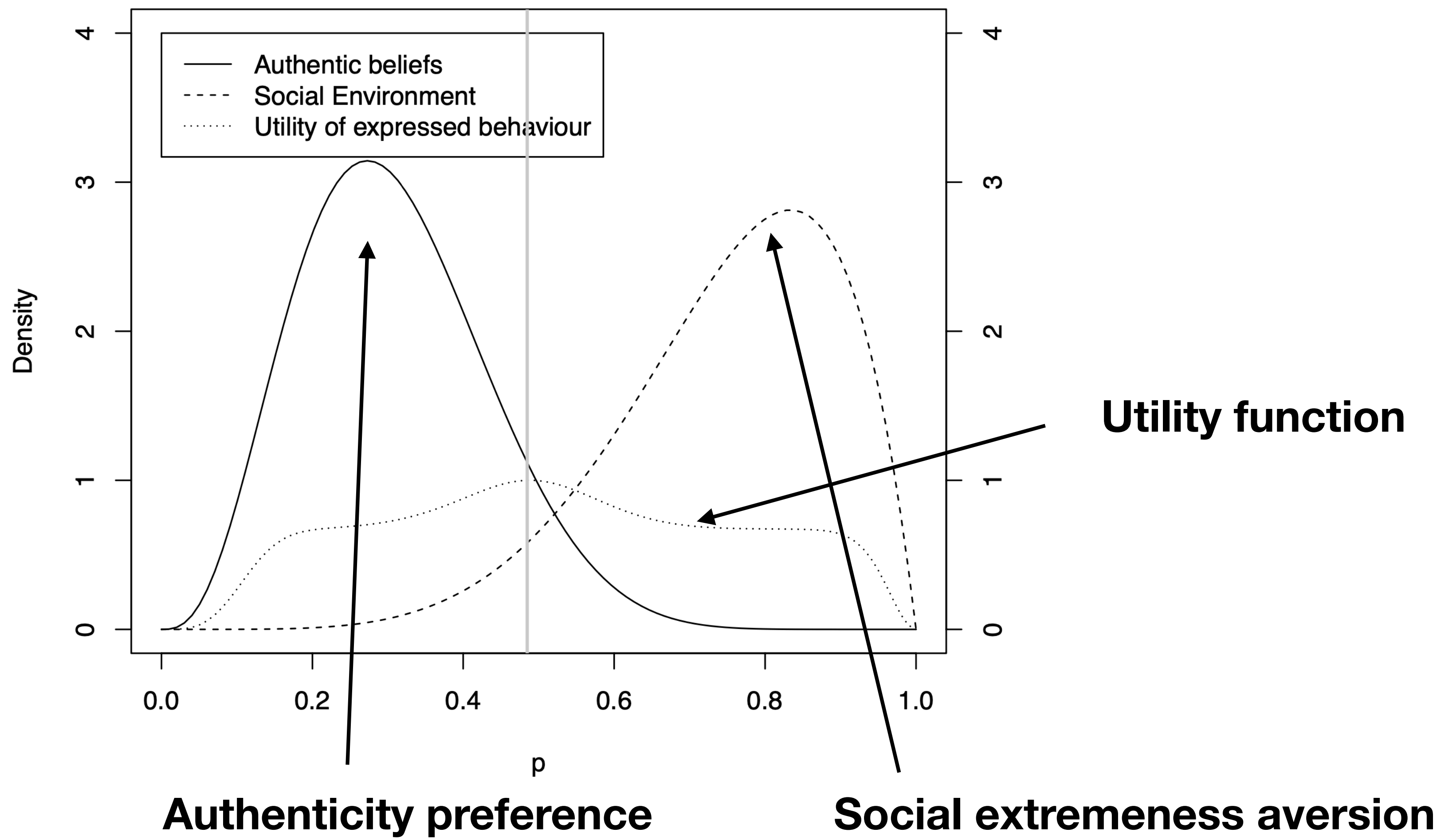
Brown, Lewandowsky, and Huang (2022)



Beliefs are beta distributions

Behaviours are a function of authentic beliefs and social environment

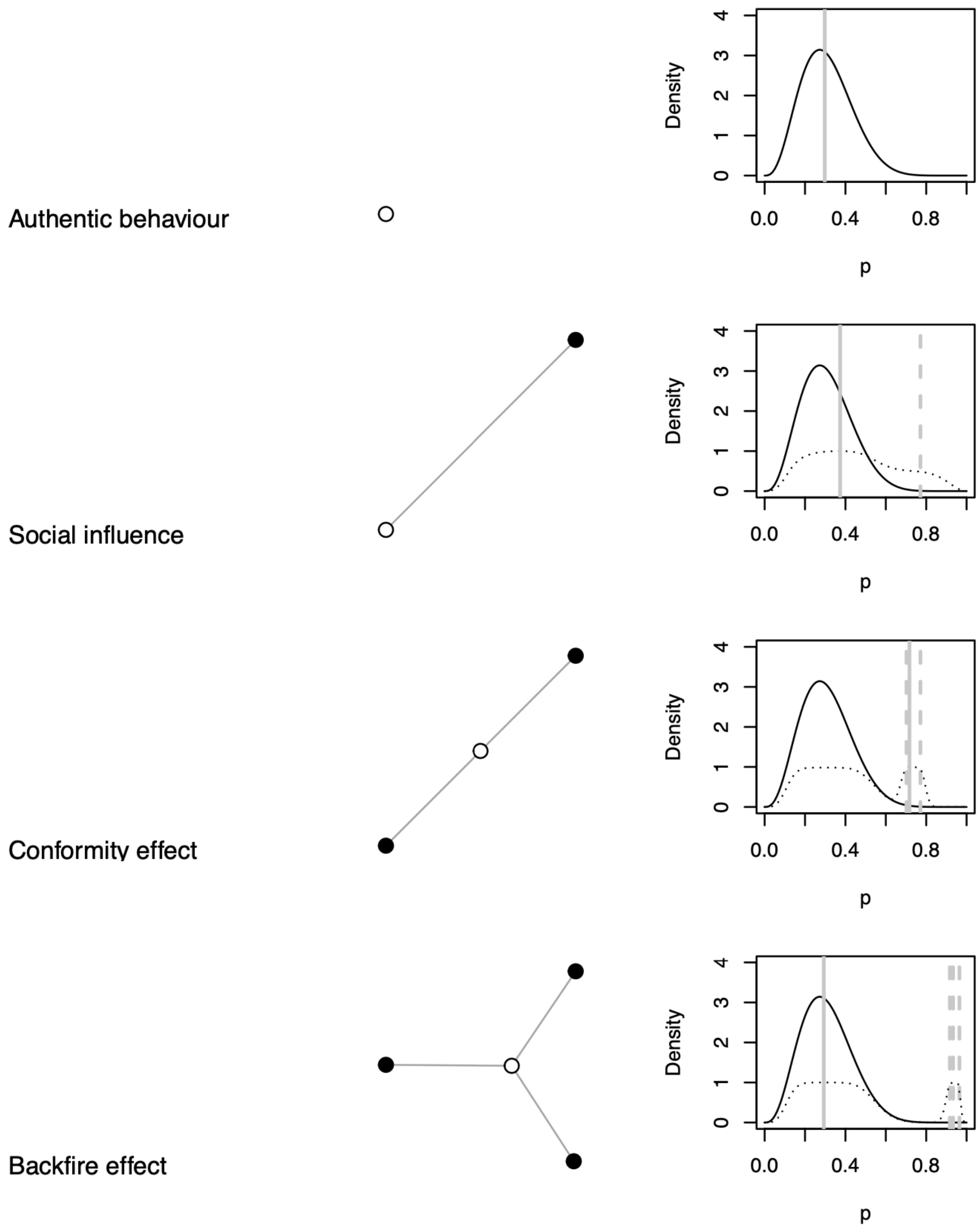
Authenticity preferences plus social extremeness aversion



Social sampling theory effects

- *Social influence or social norm effects* is presented here as the influence of one or a few individuals who influence behaviour in the direction of the social norm.
- *Conformity effects* arise when an individual's neighbors share a consensus. This was famously studied by Asch (1955) in his series of conformity experiments in which an individual was placed in a setting where they had to make a judgment with a clear an apparent objective truth.
- *Backfire effects or boomerang effects* are observed when individuals become more committed to their beliefs after receiving information that is inconsistent with those beliefs.

Social Sampling Theory



Social beliefs on a lattice

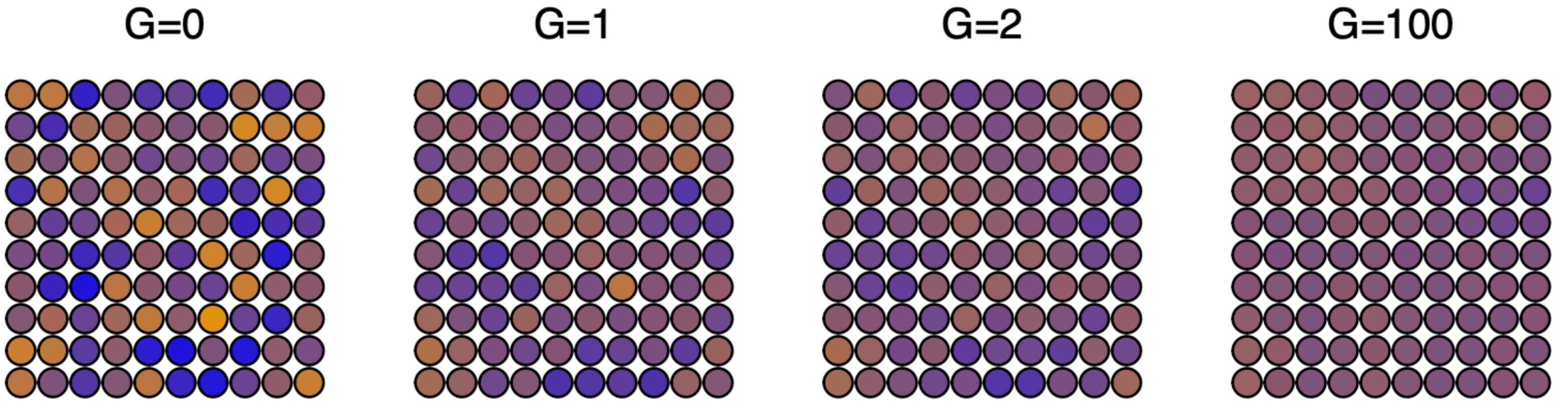


Figure 5: Social Sampling Theory over three generations. Individuals express their median authentic behaviour in generation G=0. All agents then update their expressed behaviour based on their environment in generations 1, 2, and 3. Careful inspection reveals how the iterative nature of the adaptation to one's neighbors, with them adapating at the same time, creates a subtle oscillation. An individual perceives its neighbors to be more extreme, so it expresses a more extreme behaviour. But they perceive their neighbors to be less extreme, so they express a less extreme behaviour. Now the first individual experiences less extreme neighbors, and so on. This oscillation rapidly settles down after a few generations, with individuals eventually expressing a behaviour that is both a reflection of their authentic beliefs and their social environment. (w=.4 and gamma=5).

Social beliefs with Schelling Migration

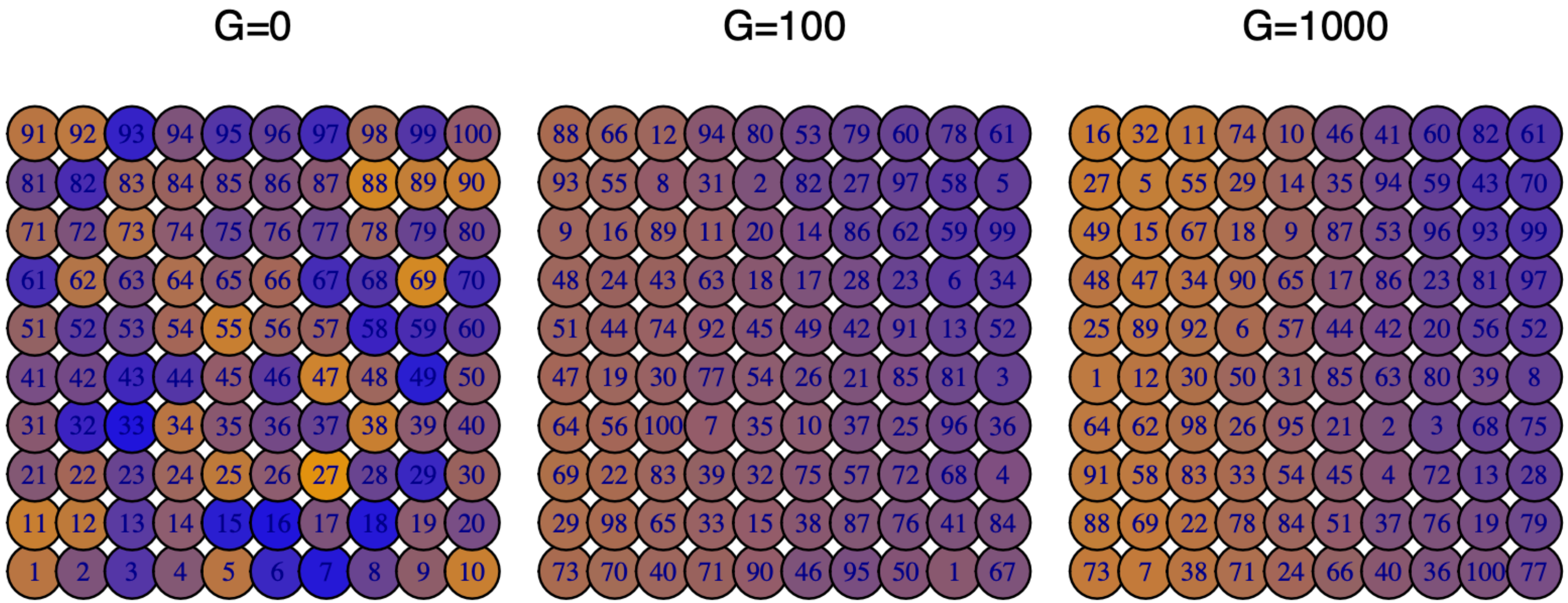


Figure 6: Social Sampling Theory with Schelling migration. Individuals are sampled in pairs, approximately 5% each generation, and allowed to evaluate the expressed behaviour among their partner's neighbors. If both parties would have higher utility if they swap position, then they replace one another on the network. In the figure, G=0 shows the behaviour individual's would express if they expressed their authentic beliefs. Each generation therafter includes two phases. The first phase involves mediating the trade-off between authenticity preferences and social extremeness aversion. The parameters are w=.4 and gamma=5, slightly favoring social extremeness aversion. The second phase involves migration. Individuals that swap position move to their new location with the expressed behaviour they would have in that location and with their corresponding utility. Numbers are preserved across panels to show how individuals can move to locations at G=1000 that better allow them to reveal their authentic behaviour shown at G=0. However, even at G=1000 there are still many individuals who would benefit from migration (e.g., number 56).

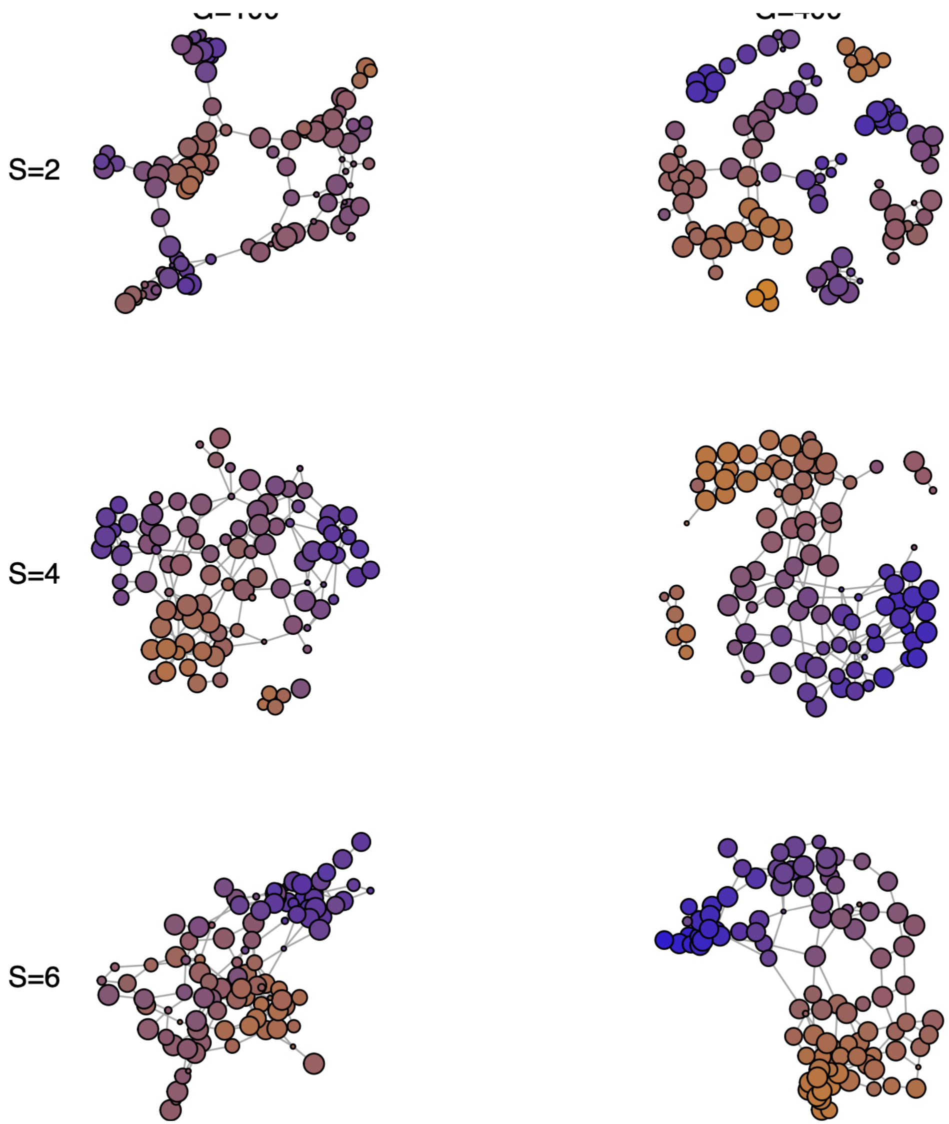


Figure 7: Search and Social Sampling Theory. Starting with a 10 x 10 lattice, agents adjust expressed behaviour to their local neighbors and authentic attitude. Then agents search through local neighbors to identify the individual with the expressed behaviour that is the smallest distance from their authentic attitude. If this behaviour is closer to their authentic attitude than the most distance behaviour among their existing neighbors, they rewire the edge to ther existing neighbor, swapping it to the more authentic distance neighbor. Agents search to a path length, S, of 2,4,6, or 10 edges. Colors indicate the expressed behaviour and size indicates the relative utility.

Perceived Normality

$$\mathcal{N}_i = 1 - |1 - \frac{2(r_i - 1)}{k_i}|$$

Table 1: Impact of search on social dynamics. Search depth and belief after 200 iterations, 50 network simulations each. Schelling migration picks two individuals at random as in Figure 6.

Depth	Utility	SD	Variance	SD.1	Components	SD.2	Normality	SD.3
No movement	0.778	0.000	0.034	0.000	1.00	0.000	0.448	0.000
Schelling migration	0.783	0.014	0.065	0.022	1.00	0.000	0.607	0.095
2	0.829	0.013	0.153	0.021	4.90	1.403	0.619	0.030
4	0.843	0.010	0.158	0.016	1.68	0.713	0.650	0.033
6	0.846	0.010	0.156	0.018	1.36	0.525	0.650	0.032

Group problem Solving

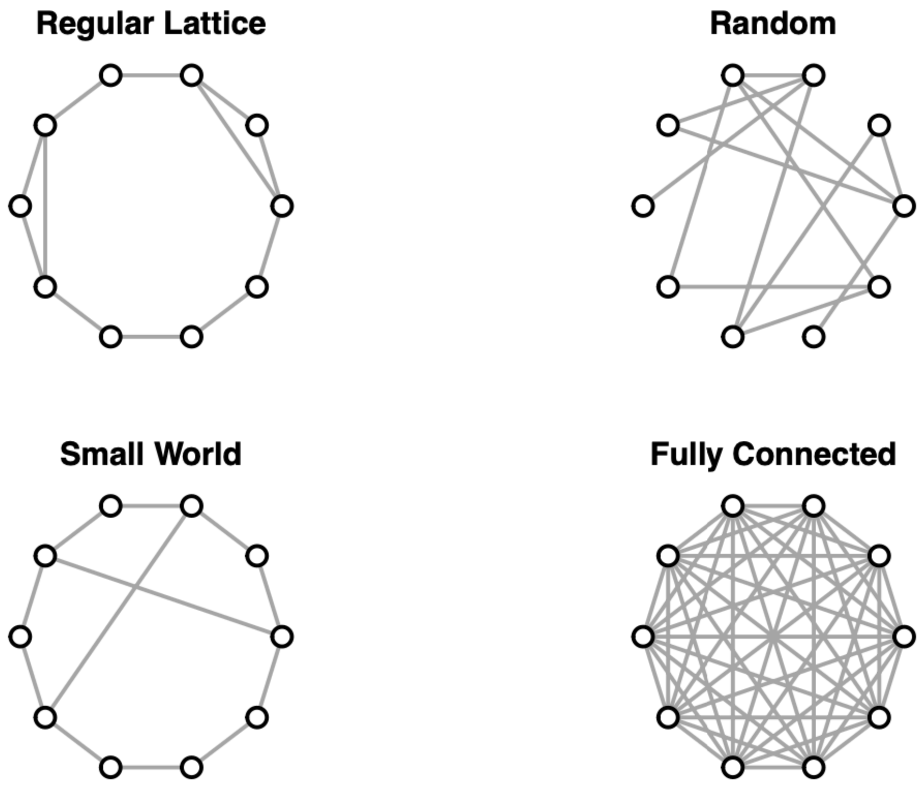


Figure 2: The four different treatment groups used in Mason et al 2008. Each network represents a different configuration for information sharing. Nodes represent individuals in the experiment and edges indicate the ability to see another individual's guesses and performance. Images recreated based on the description provided in Mason et al. 2008.

Smooth and Rugged Landscapes

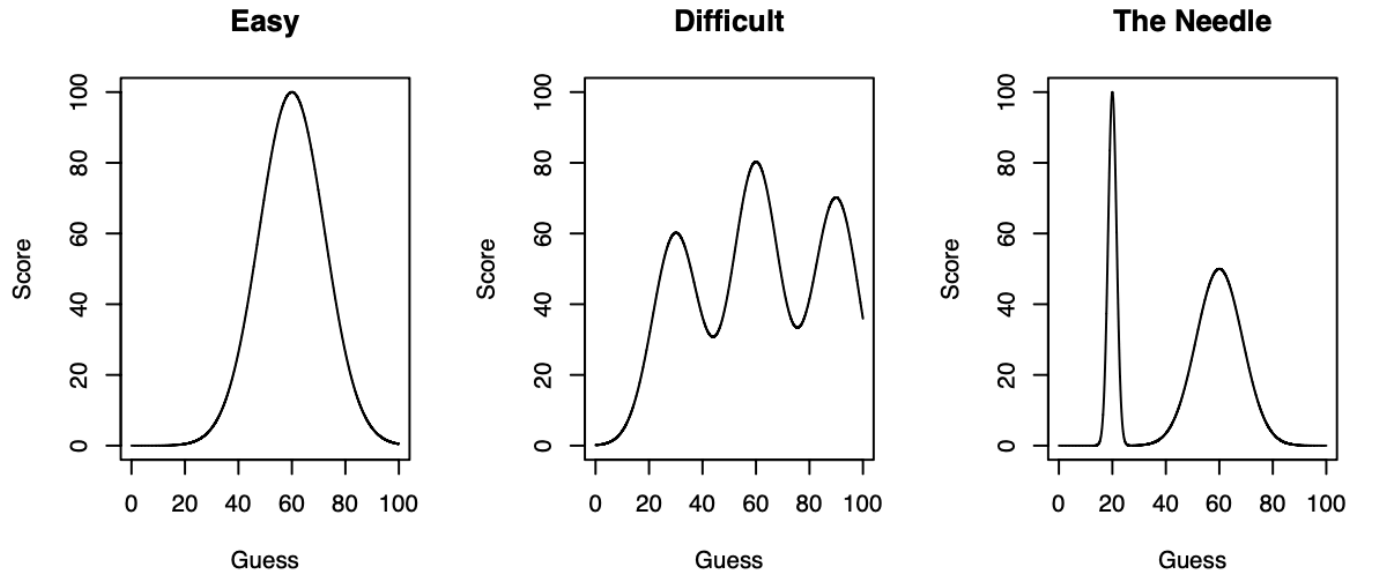
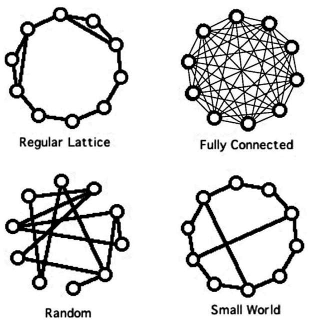
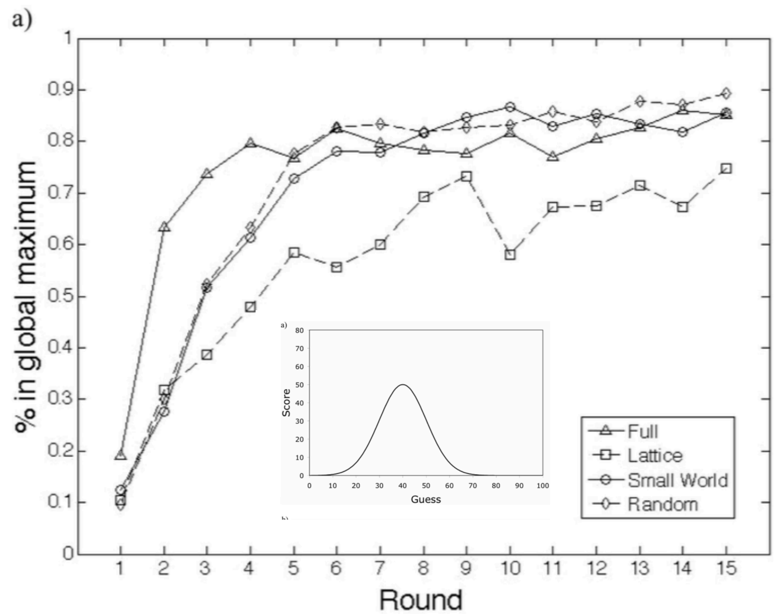
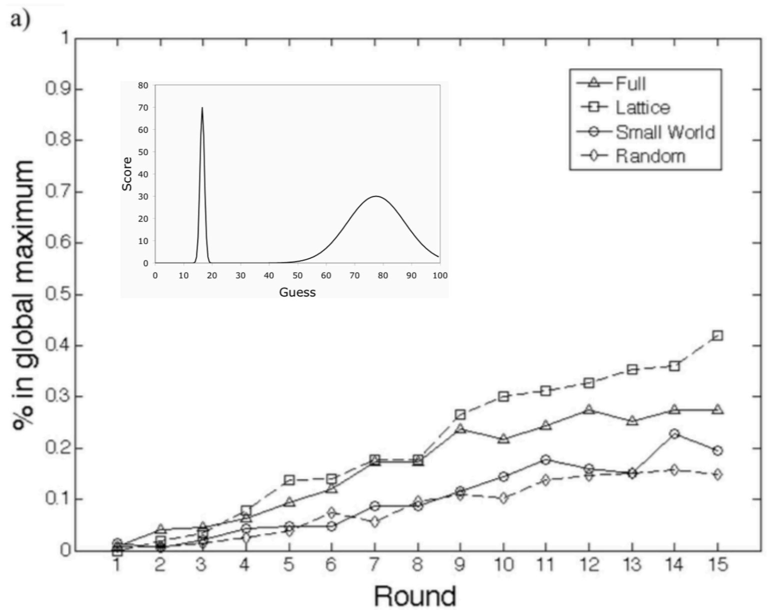


Figure 3: Payoff functions for three collective search problems in Mason et al. 2008. Participants would each guess a number and receive the payoff associated with that number.

NETWORK CONNECTIVITY MODULATES EXPLORATION/EXPLOITATION







Mason, Jones, & Goldstone (2008)

People are connected together in different configurations(so they see other people's performance. They are trying to choose the best guess (a number) to get the highest score (see inset).

When the problem is easy, more connections are good

When the problem is hard, fewer connections is better

Smooth and Rugged Landscapes

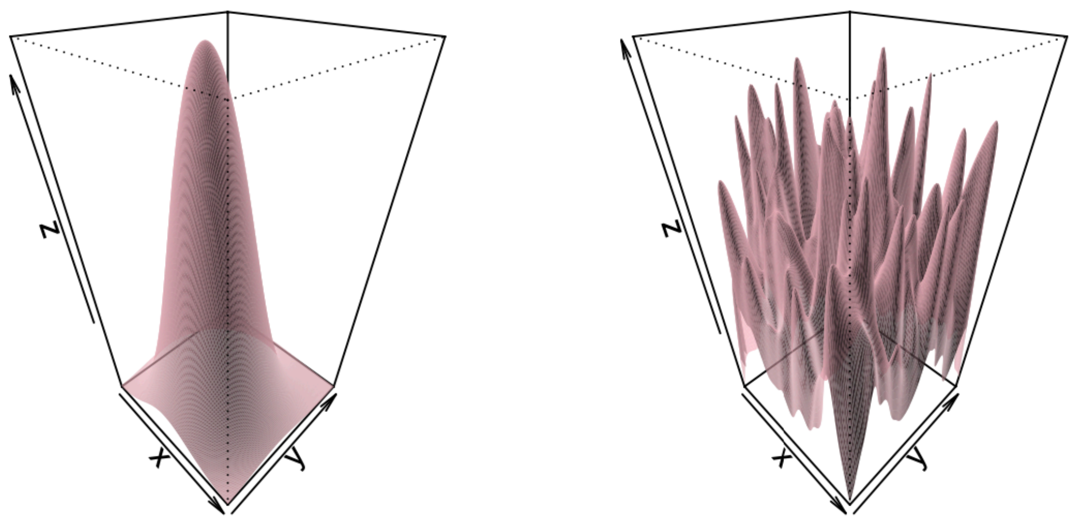


Figure 1: Visual metaphors of smooth and rugged landscapes. The smooth landsdcape on the left always indicates the path to the global maximum by its gradient. Moving up the gradient leads to the global maximum. The rugged landscape on the right requires exploration. Most peaks are local maxima: following the gradient leads to a less than optimal solution. But there is no straightforward algorithm for finding the global maximum.

Fully connected groups get trapped in local minima

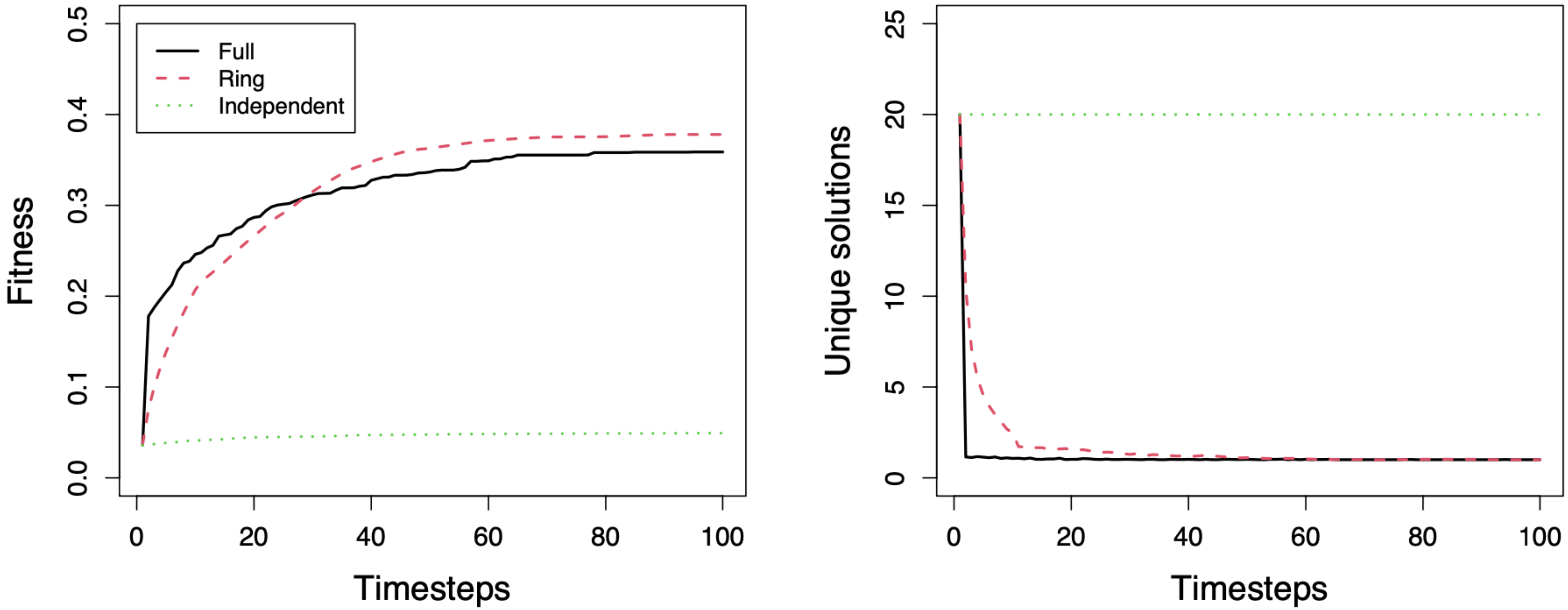


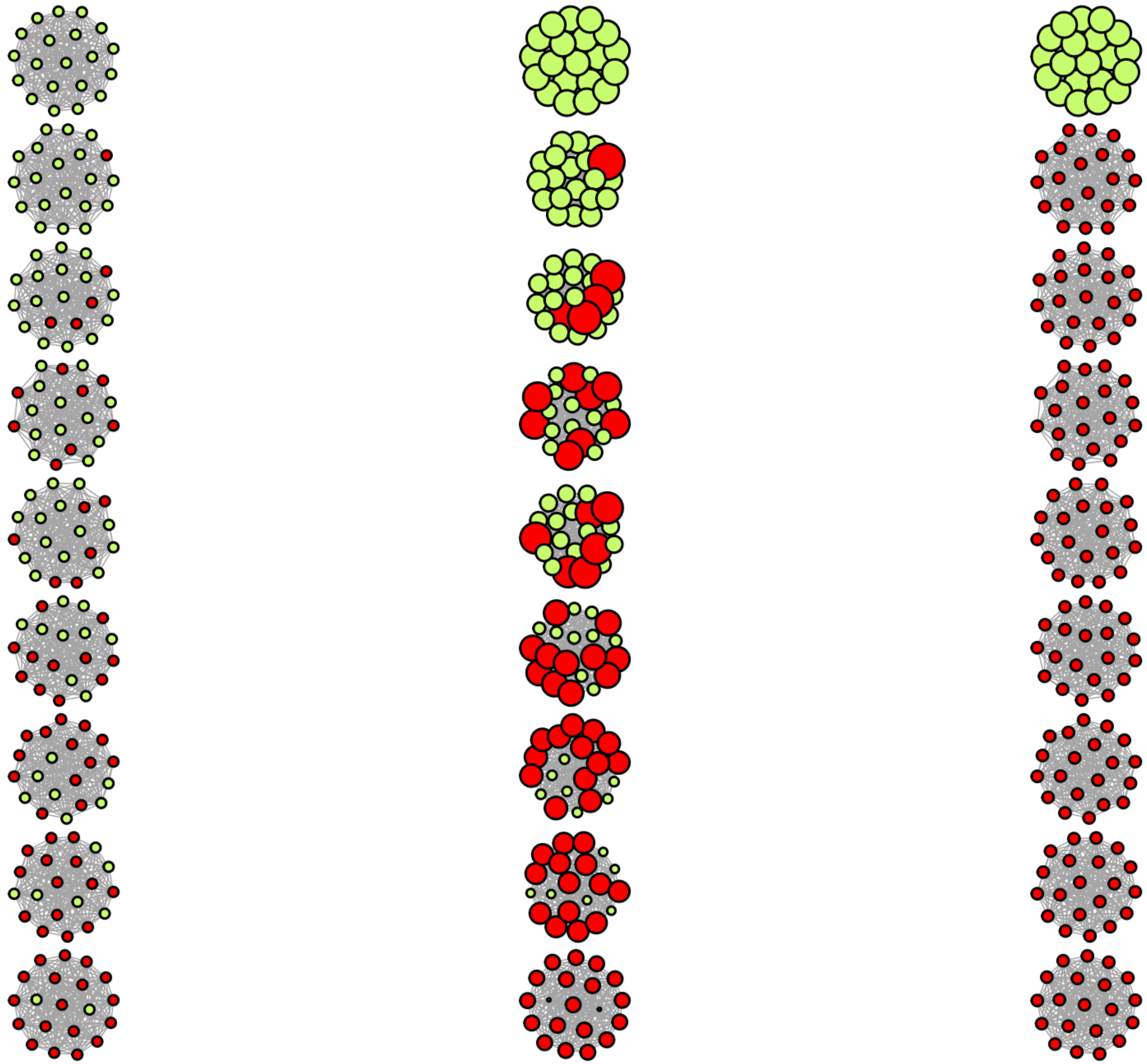
Figure 6: Performance trajectories for each of the network configurations. Fitness represents the average performance of all the individuals in the group. Unique solutions gives an indication of the exploratory potential of the group.

Prisoner’s dilemma

	Cooperate	Defect
Cooperate	R=3	S=0
Defect	T=5	P=1

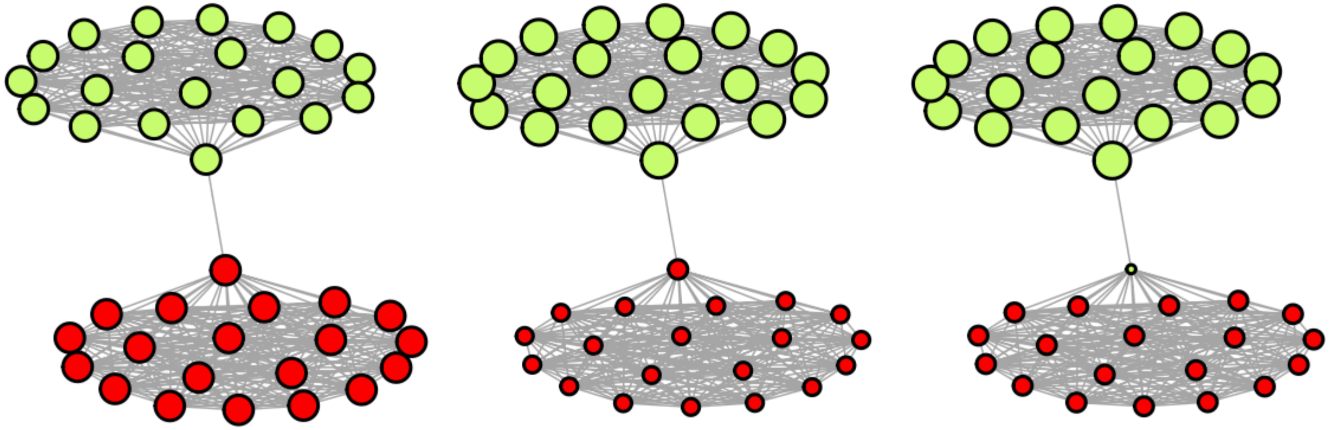
$$T > R > P > S$$

Mean field approximation (Fully connected network)



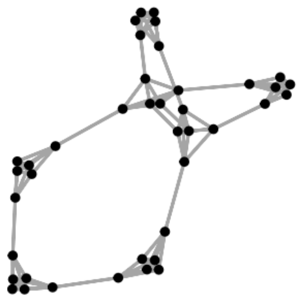
Social Viscosity

Assortativity

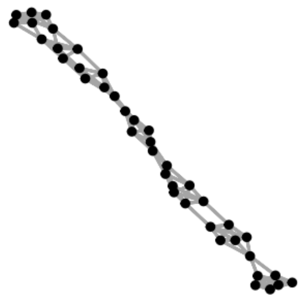


Mason Watts Networks

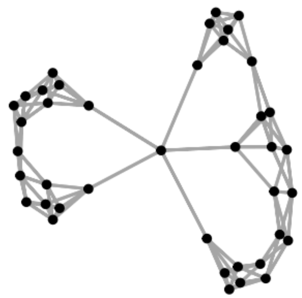
Max CC



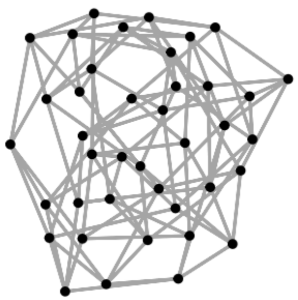
Max mean B



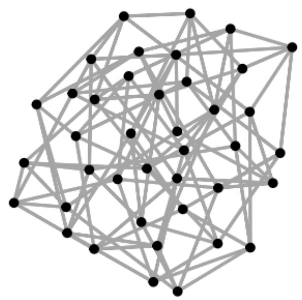
Max max B



Min CC

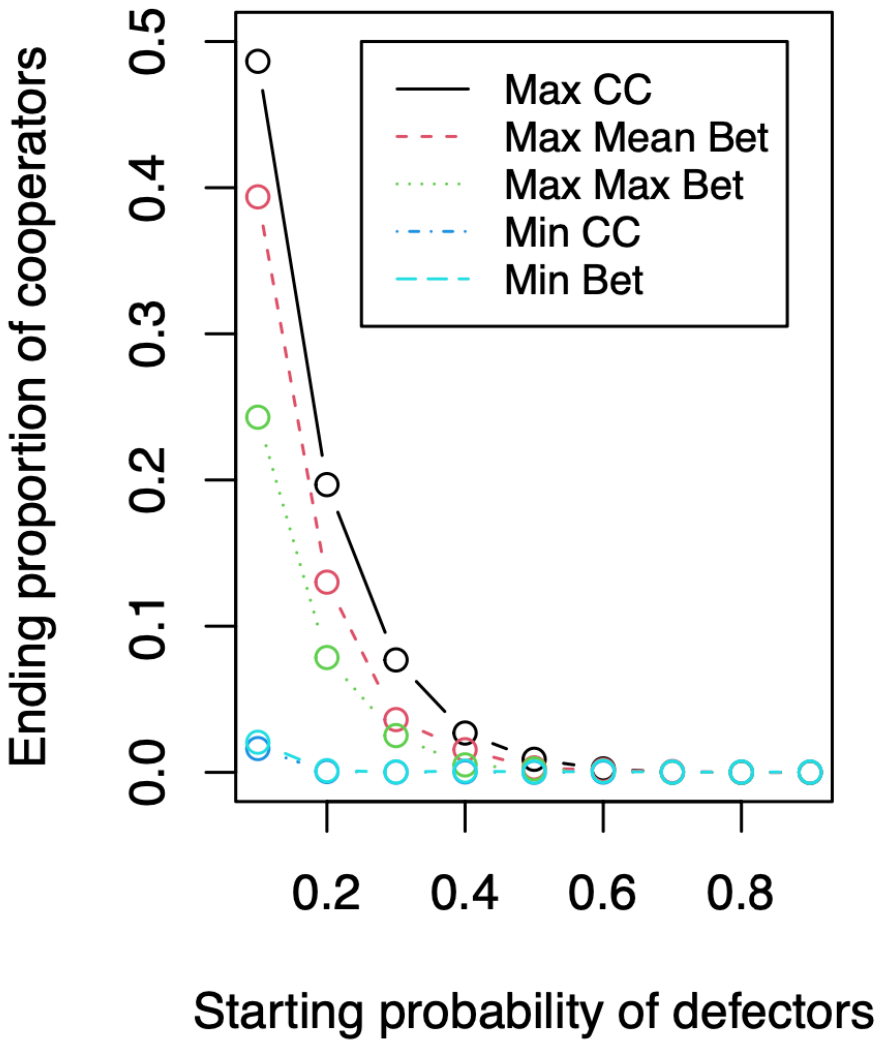


Min B

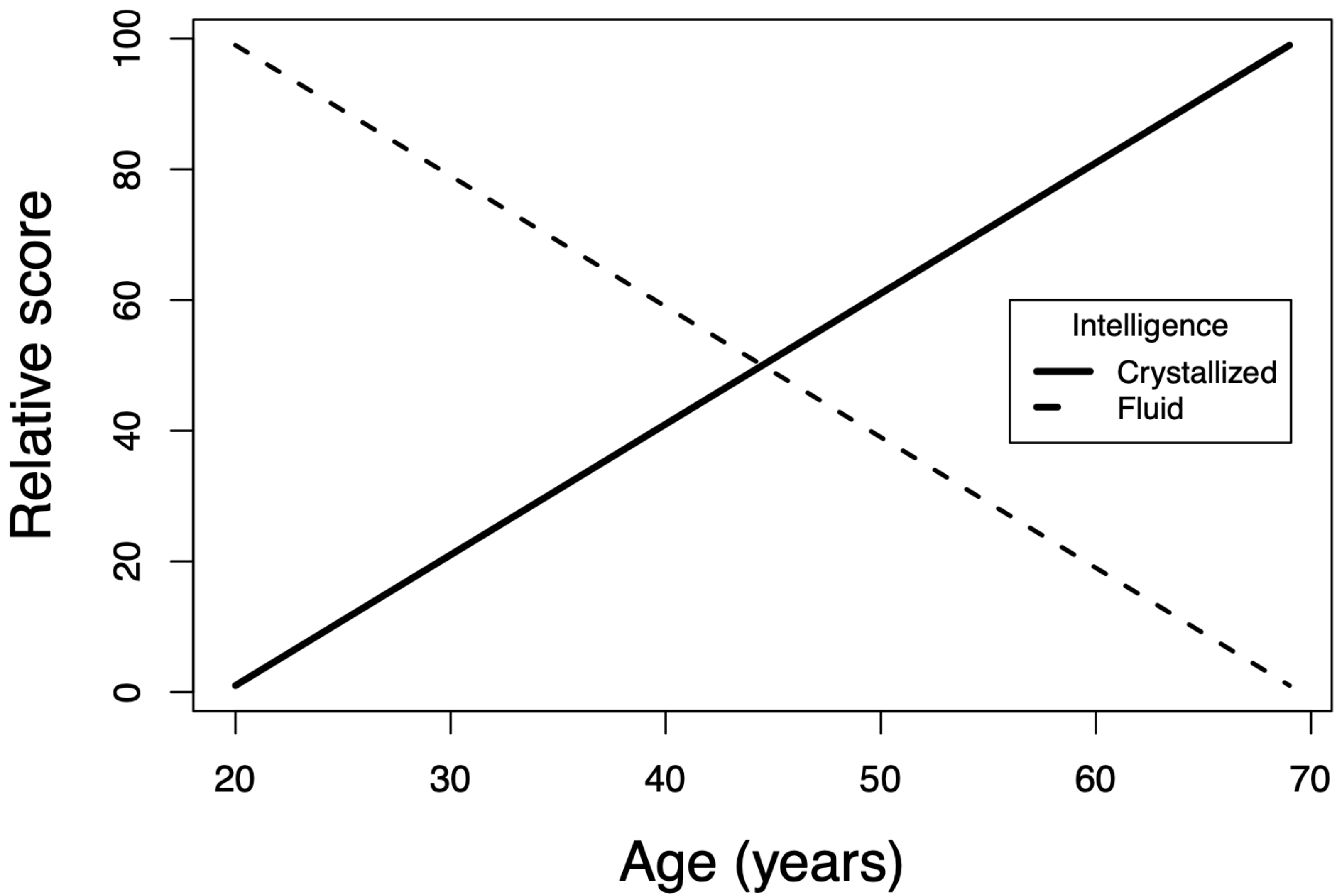


**These networks maximise a specific feature.
This uses a Markov search process to identify a metric maximizing/minimizing Feature.**

More C means more cooperation



Aging


Relative score
0
20
40
60
80
100
20
30
40
50
60
70
Age (years)
Intelligence
Crystallized
Fluid

With aging

- Words get further away in ‘conceptual space’
- People perceive objects to be less similar

Networks become more sparse with age

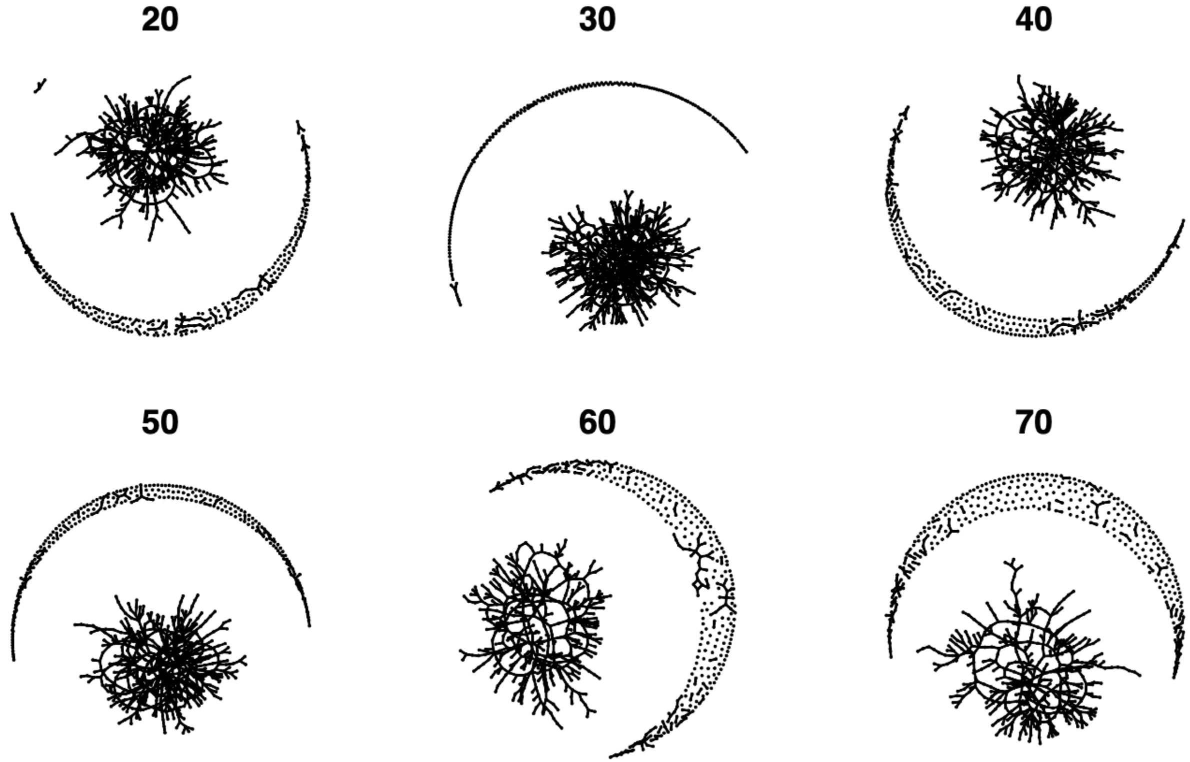
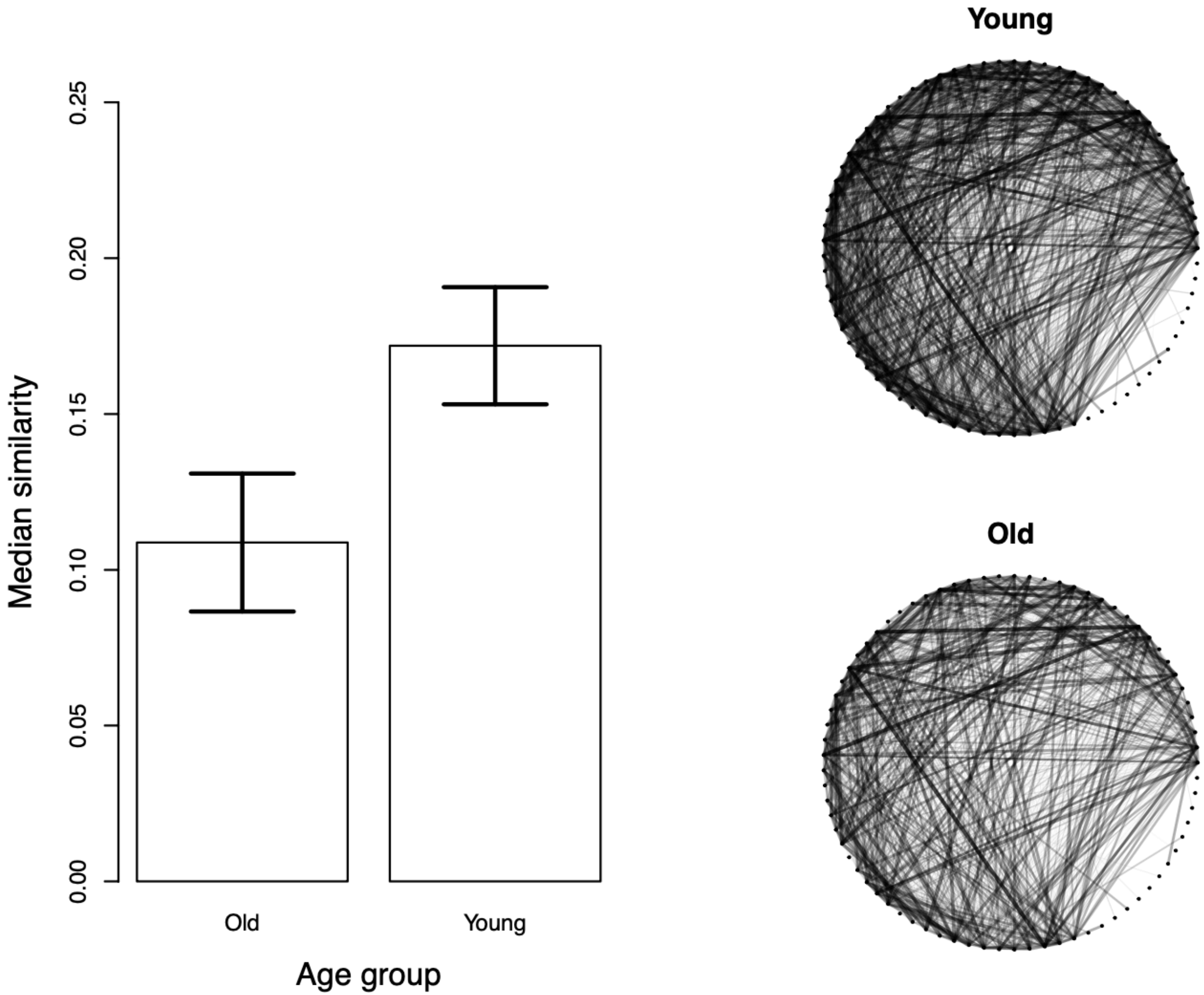


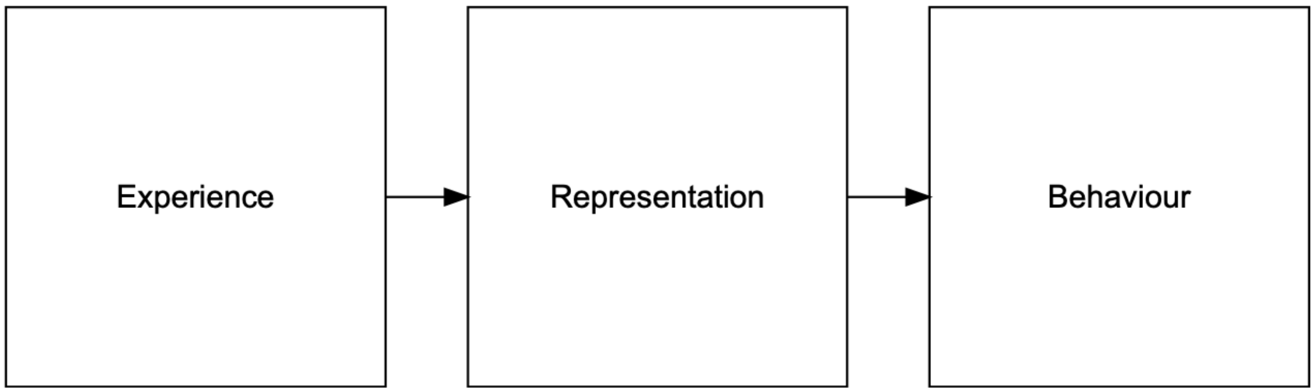
Figure 2: Adult networks of 420 cue words collected across individuals ranging in age from approximately 20 to 70. Edges are formed when cues produced similar patterns of free associations. Giant components are shown in the center of each representation, with isolates and smaller components shown in the hemispheres along the edge. From age 30 onwards Dubossarsky et al., 2017, showed that degree gradually fell as average shortest path length increased.

Older adults say that animals are less similar to one another



Can we model this effect using basic learning models?

- We will use a learning process
- That learns an environmental representation
- Then the individual will make decisions based on their representation.



Rescorla-Wagner learning

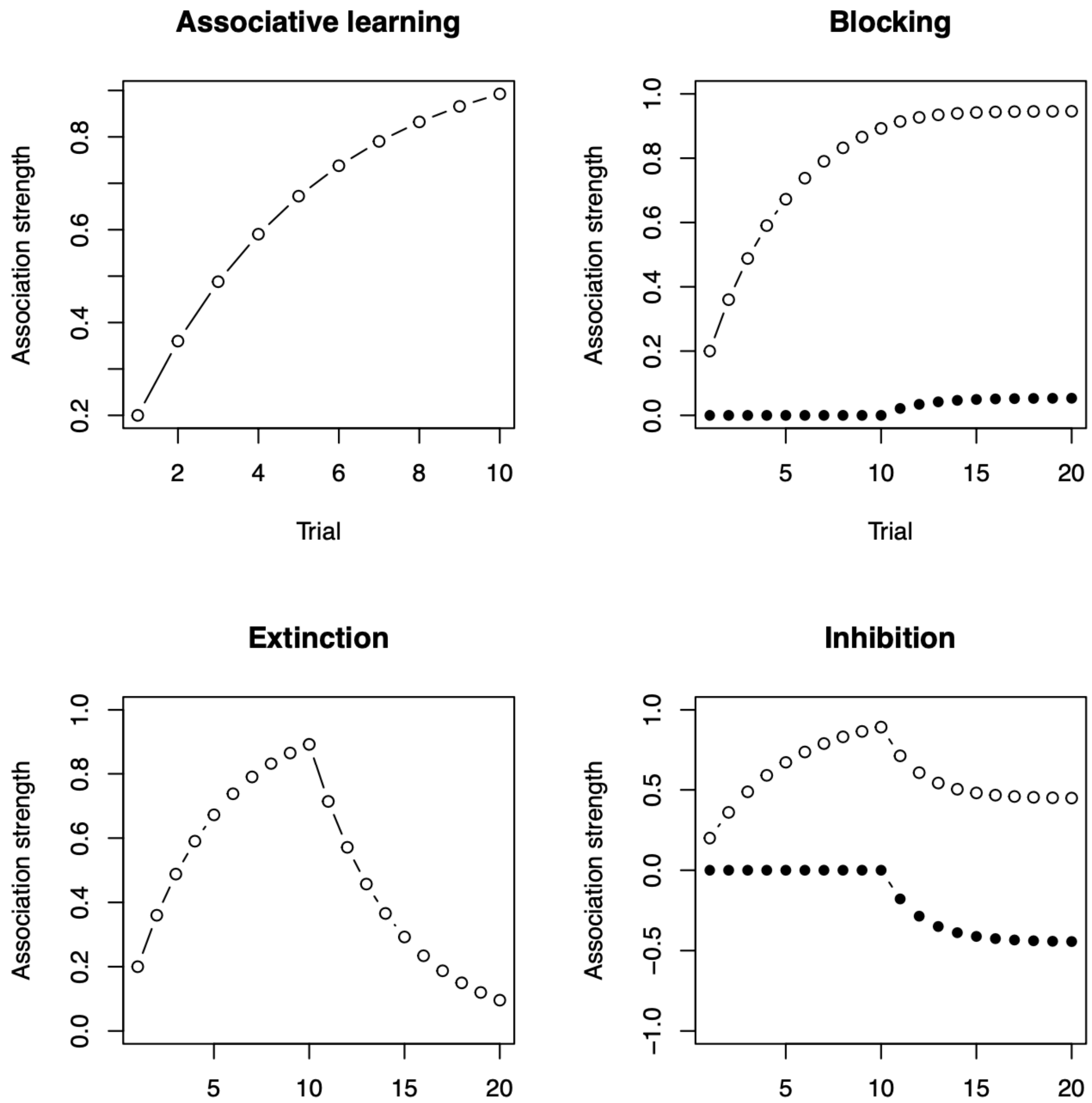
(Prediction error modelling)

$$\Delta V_{C\to U} = \alpha_C \beta_U (\lambda_U - V_{C\to U})$$

$$V_{C\to U,t+1} = V_{C\to U,t} + \Delta V_{C\to U,t}$$

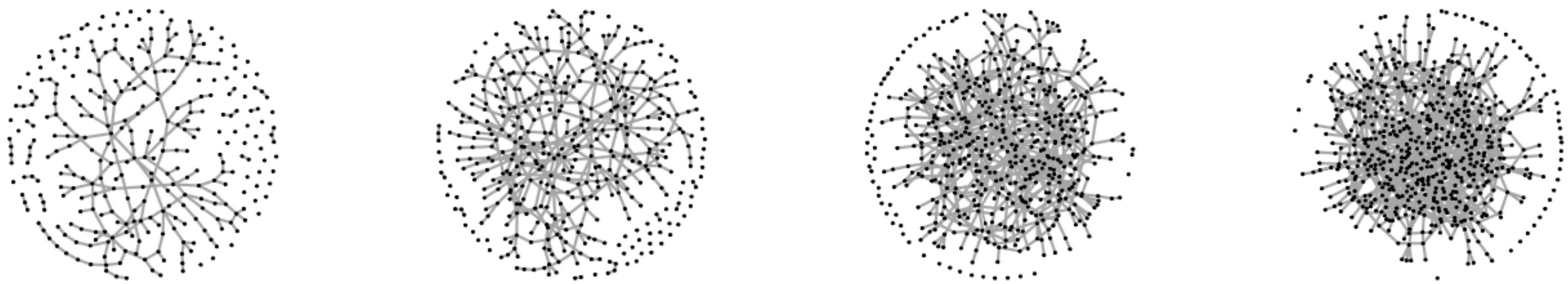
- Lambda is the observed
- Vc is the expected
- The difference is the prediction error.
- So the system gets better at predicting the actual outcome over time.

Basic learning effects with Rescorla Wagner



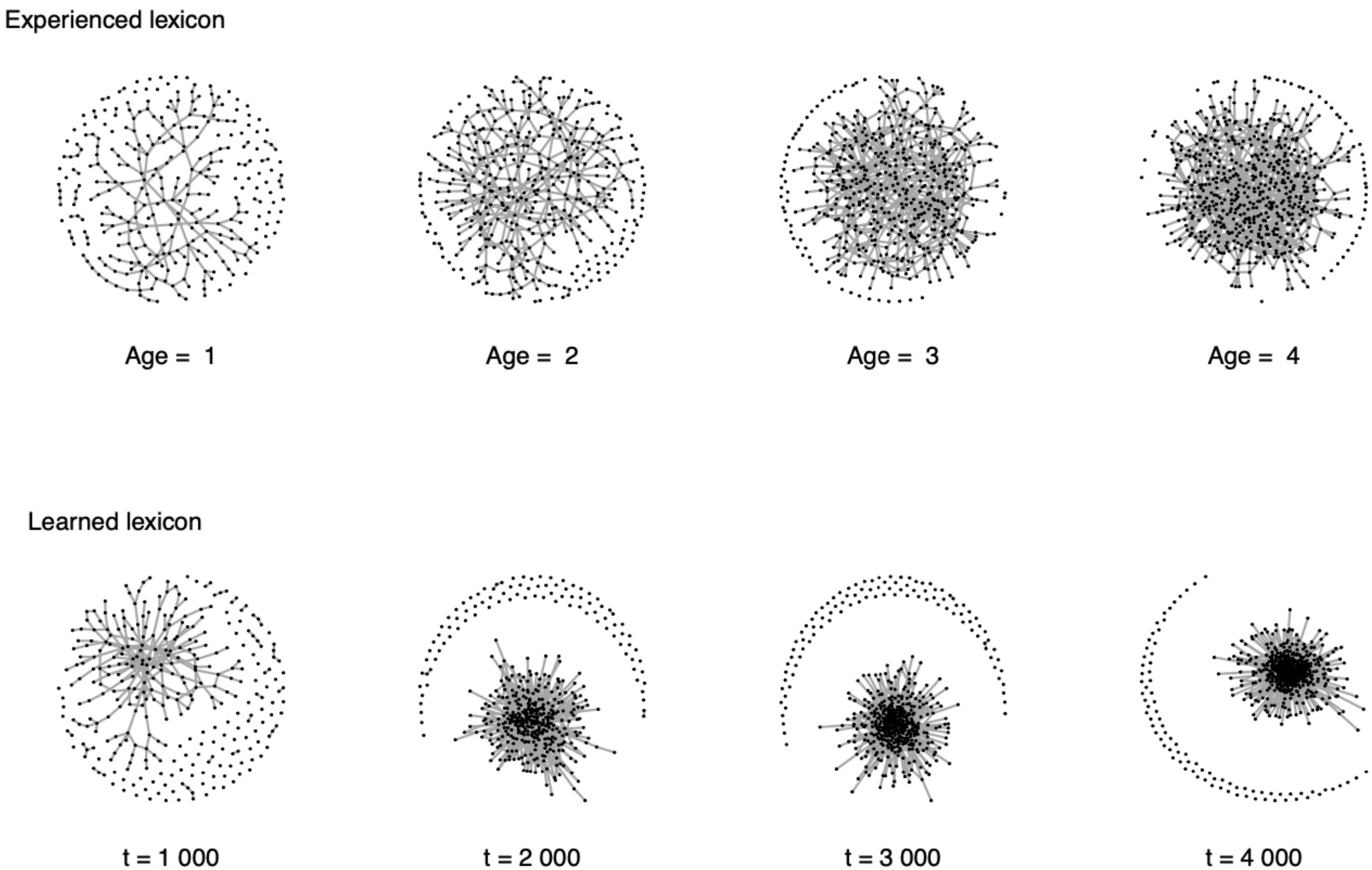
This is the environment

Experienced lexicon



Age = 1 Age = 2 Age = 3 Age = 4

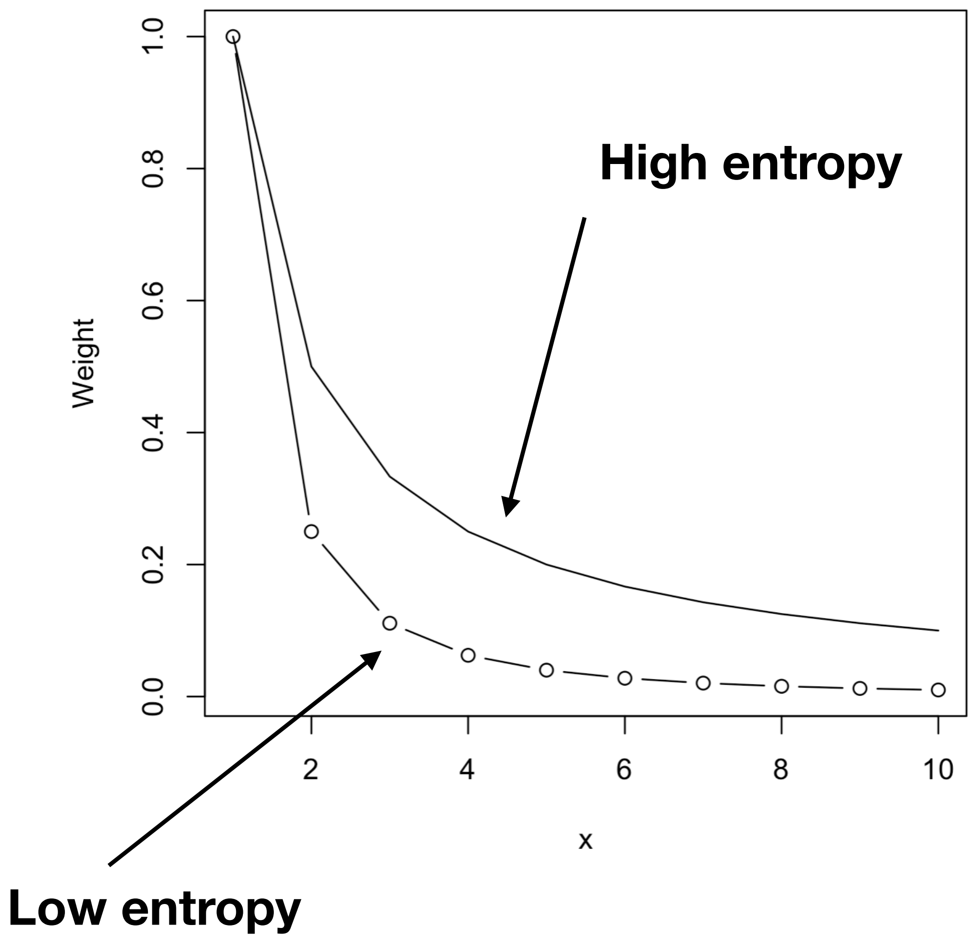
This is the learned representation



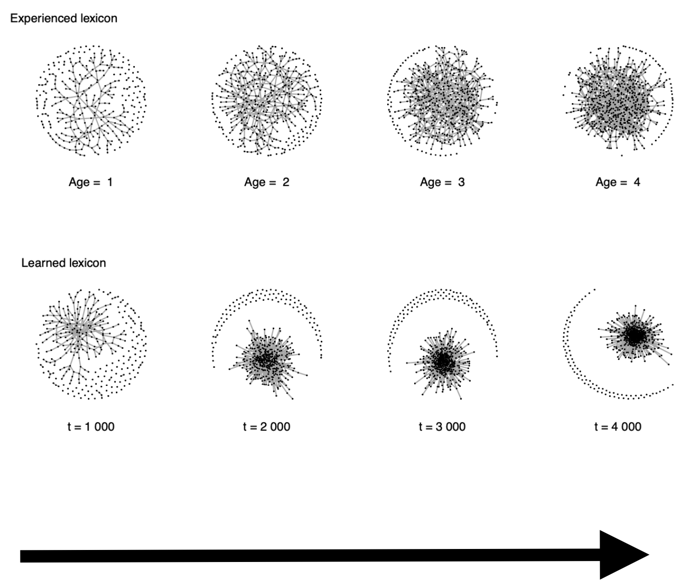
Higher entropy

$$H = -\sum_{i=1}^{k} p_i log(p_i)$$

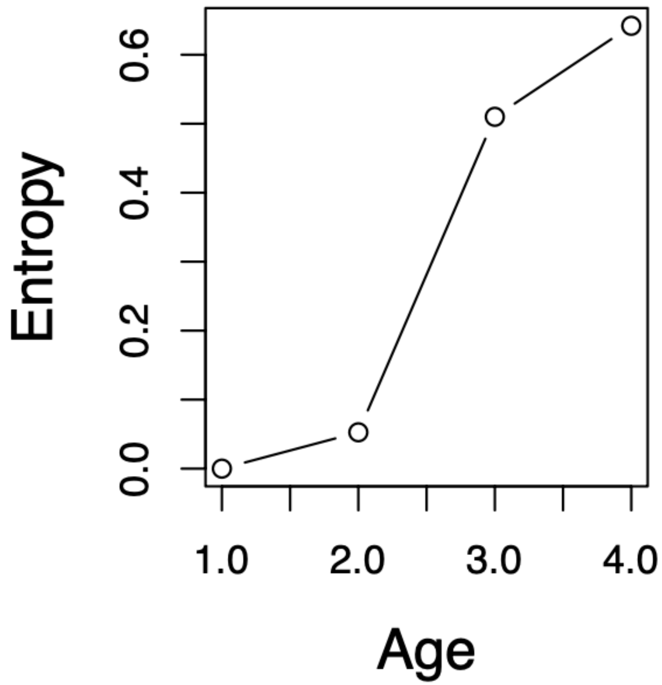
- P = vector of edge weights



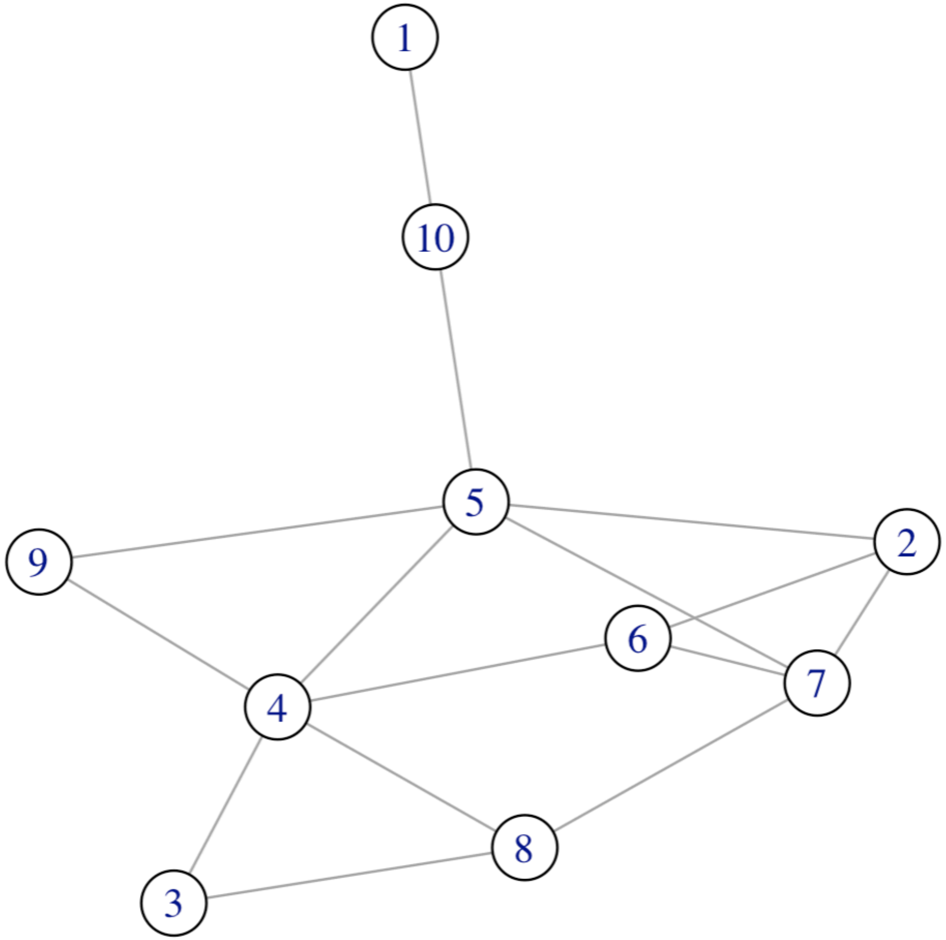
With increased learning, entropy increases



$$H = -\sum_{i=1}^{k} p_i log(p_i)$$



How do we measure similarity?



This is the result—Higher entropy (less predictability) and lower similarity

$$S = A_{j \to k} + A_{j \to k}$$