SANTA FE INSTITUTE

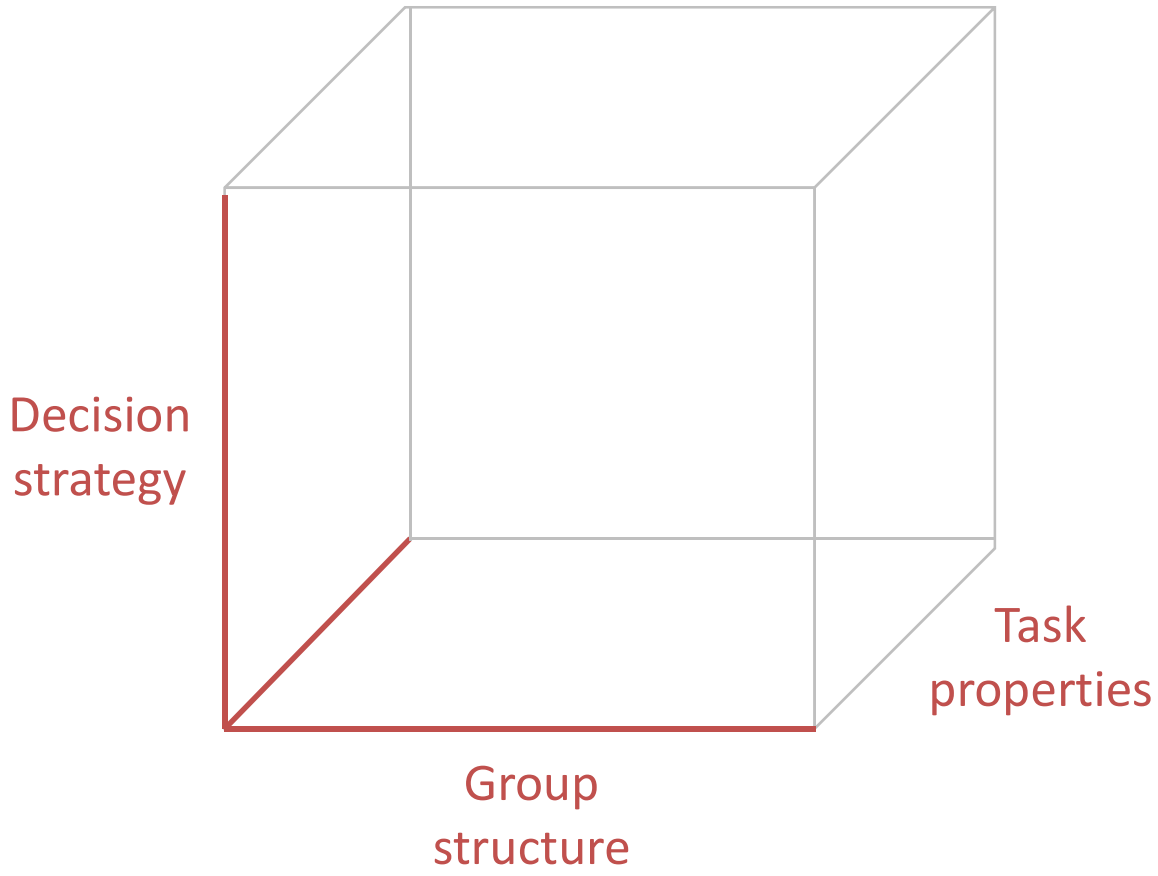
Social learning strategies
Collective problem solving

Mirta Galesic

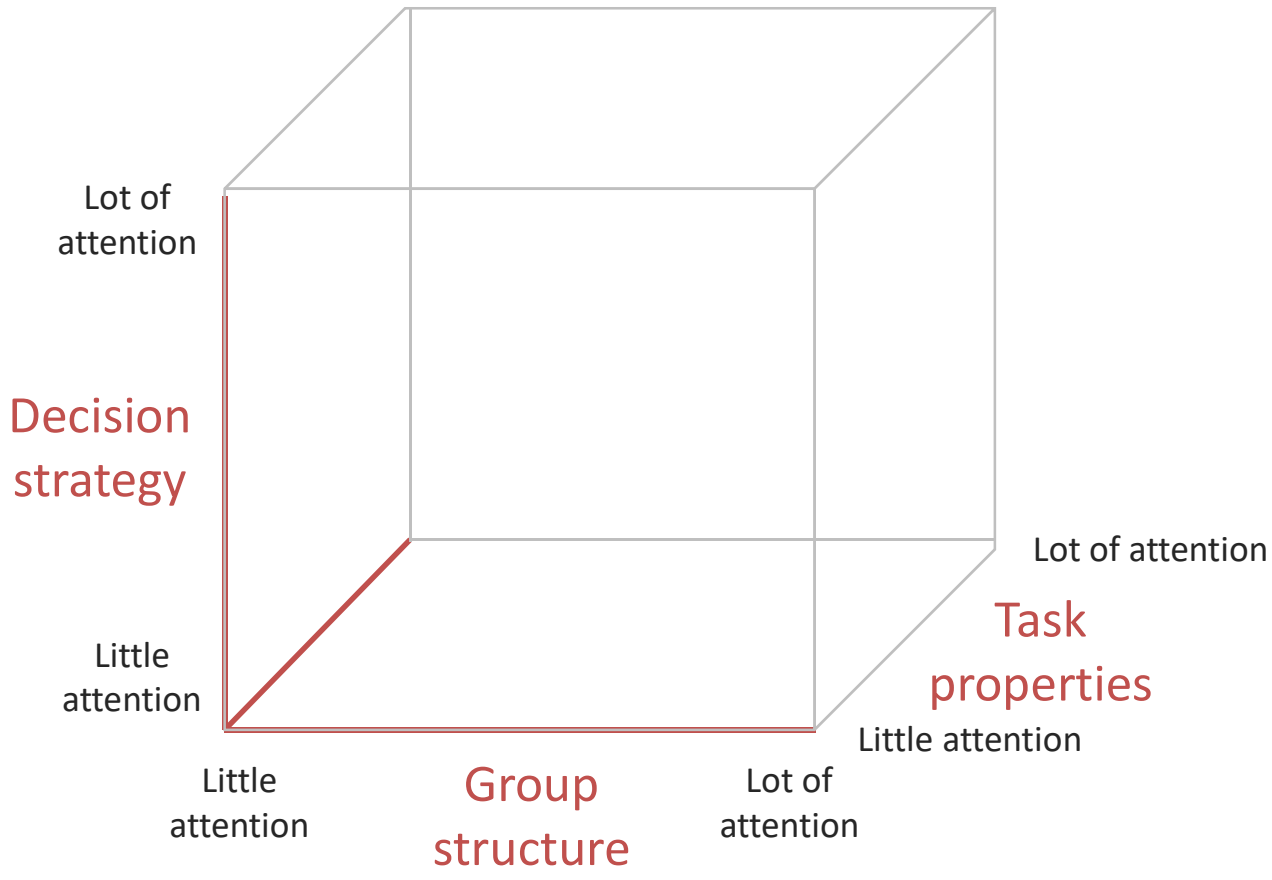
Cowan Chair in Human Social Dynamics,
Santa Fe Institute

Collective problem solving

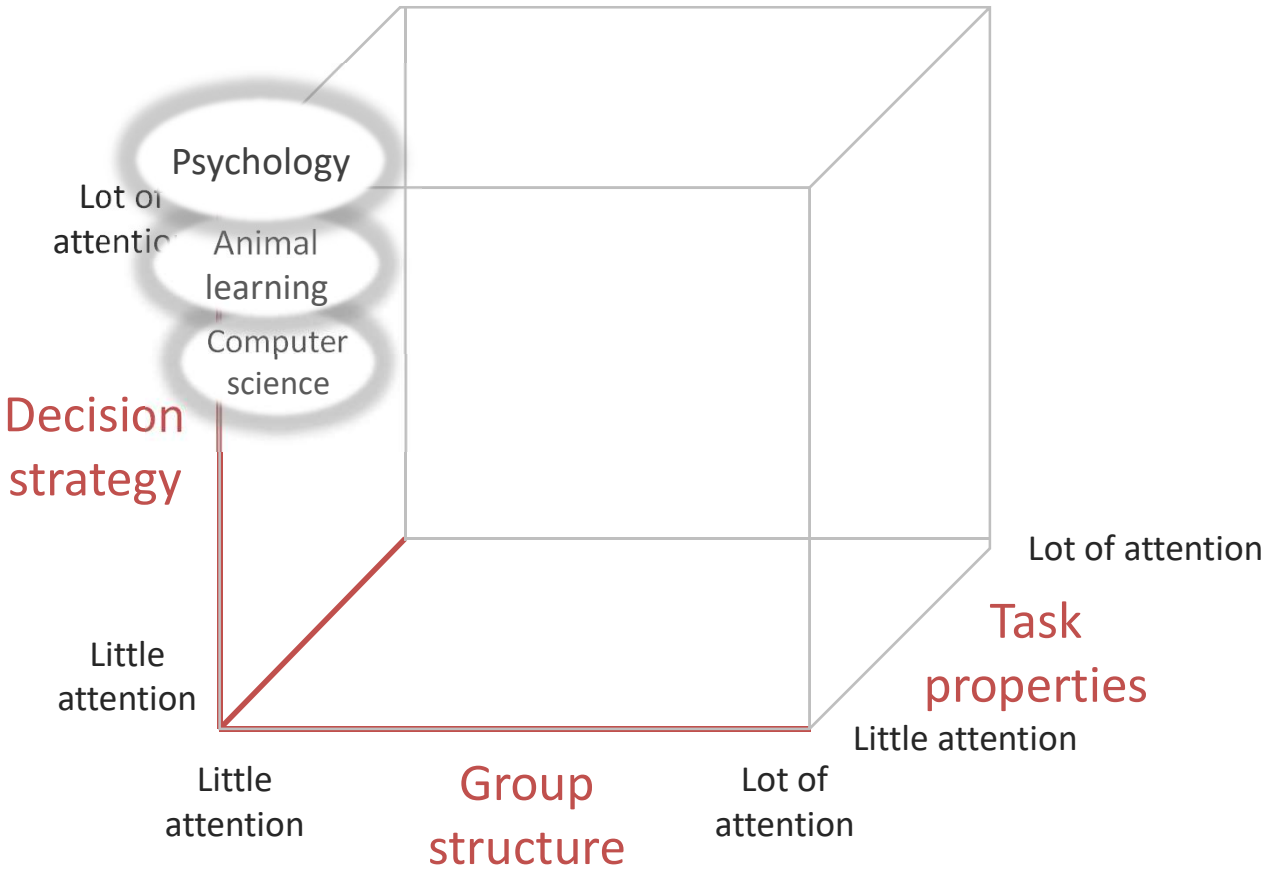
Dimensions of collective problem solving



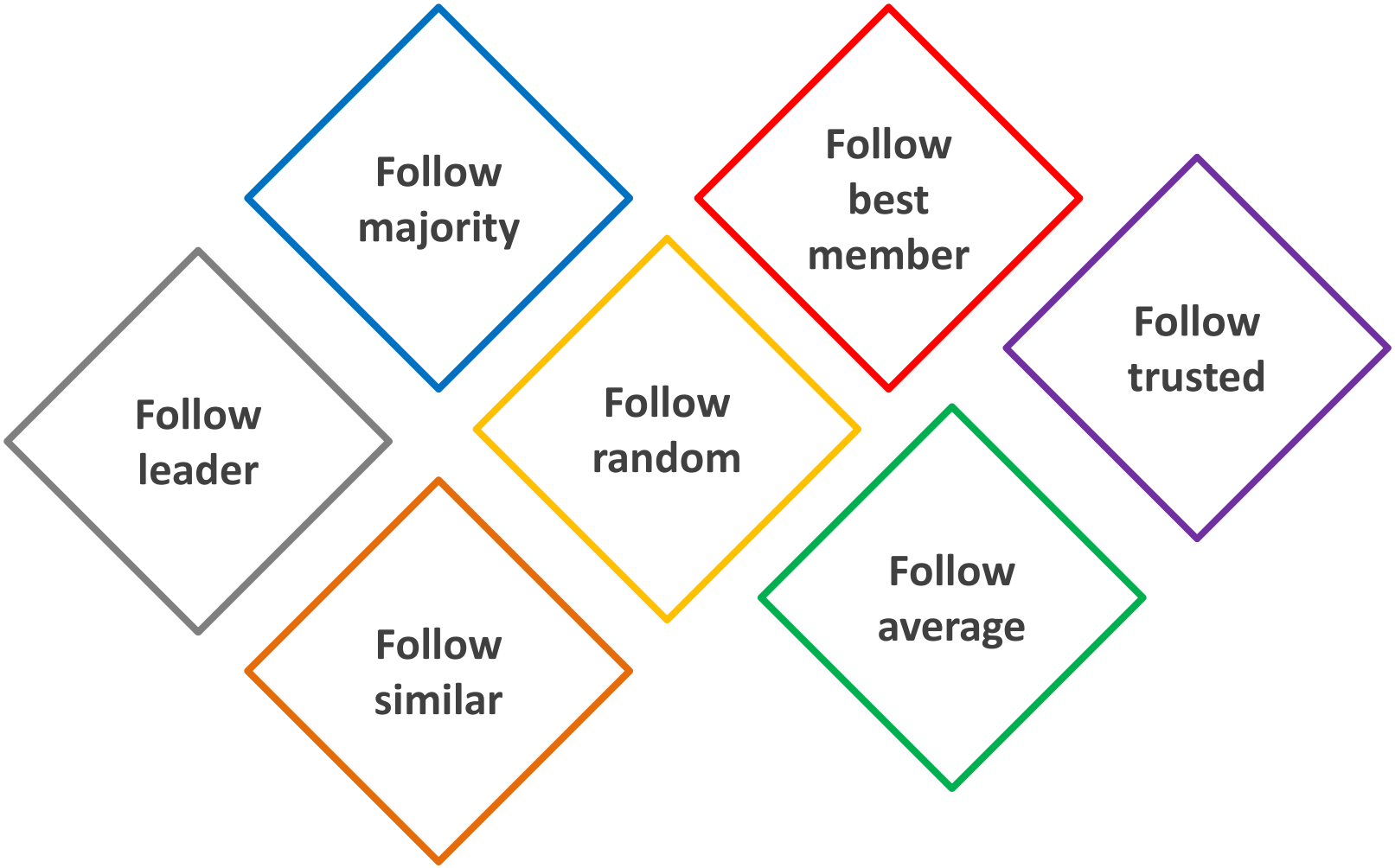
Views of different disciplines



Views of different disciplines

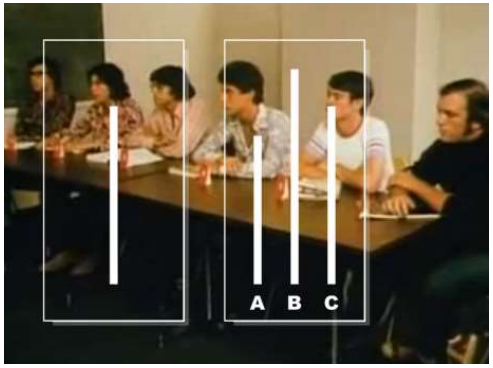


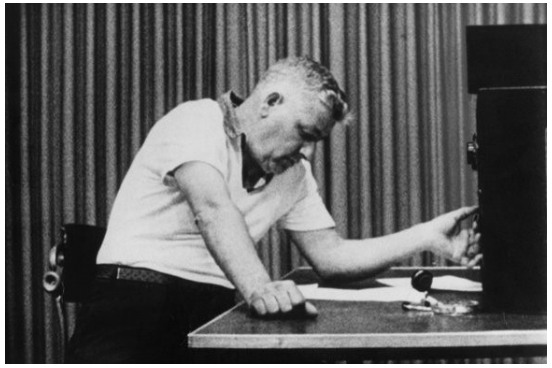
Group decision strategies



… in social psychology

- The dark side of conformity
  - Asch (1955) and Milgram (1963) experiments

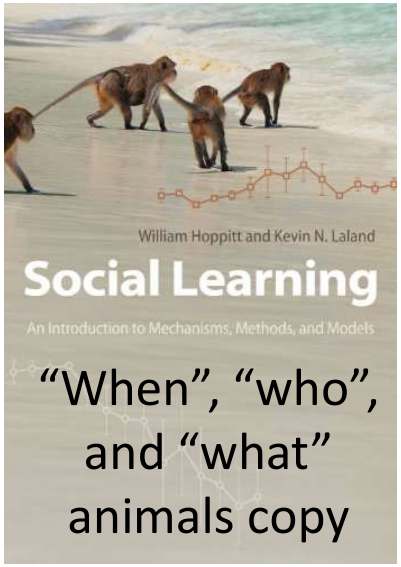




- And the good sides…

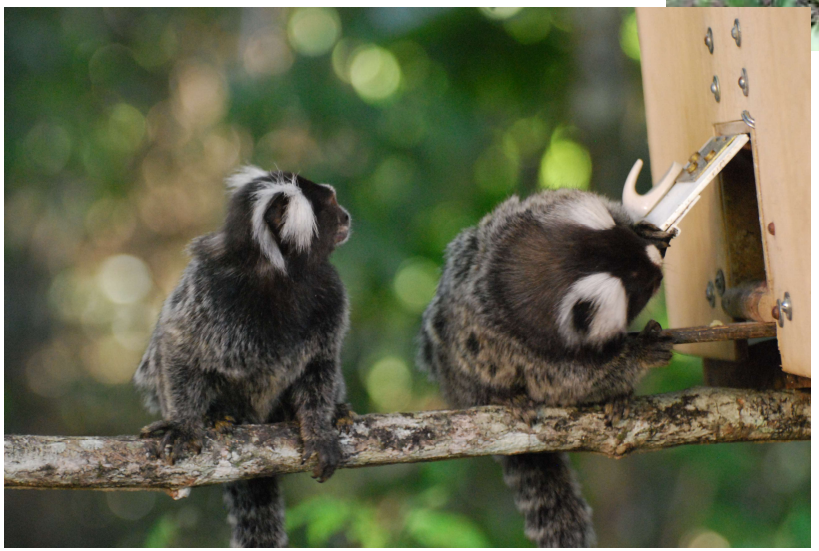
  - Social norm interventions

… in animal learning



"When", "who", and "what" animals copy

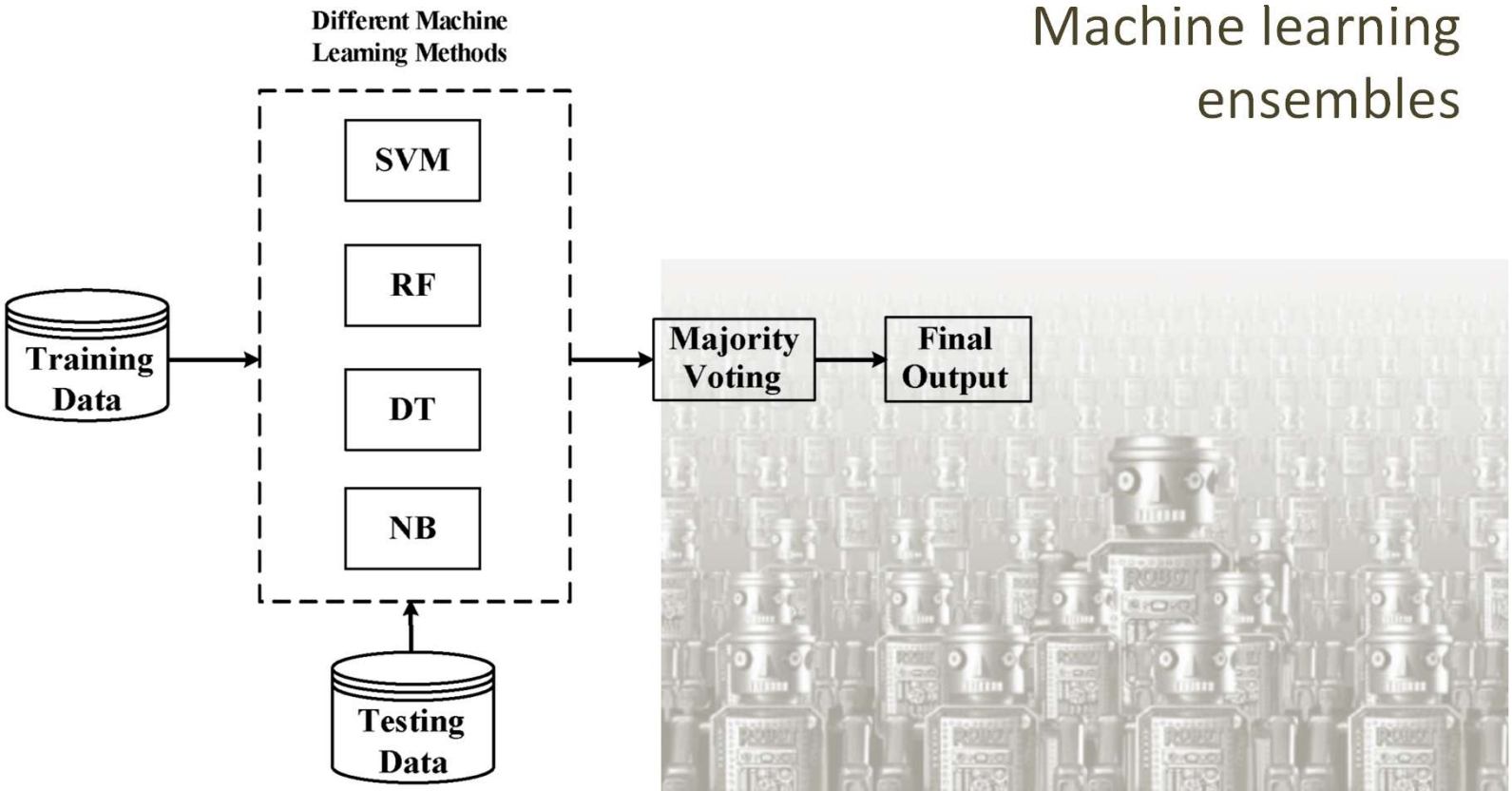
Van de Waal, 2013, Science



Gunhold, Animal Behavior, 2014

Aplin, 2015, Nature

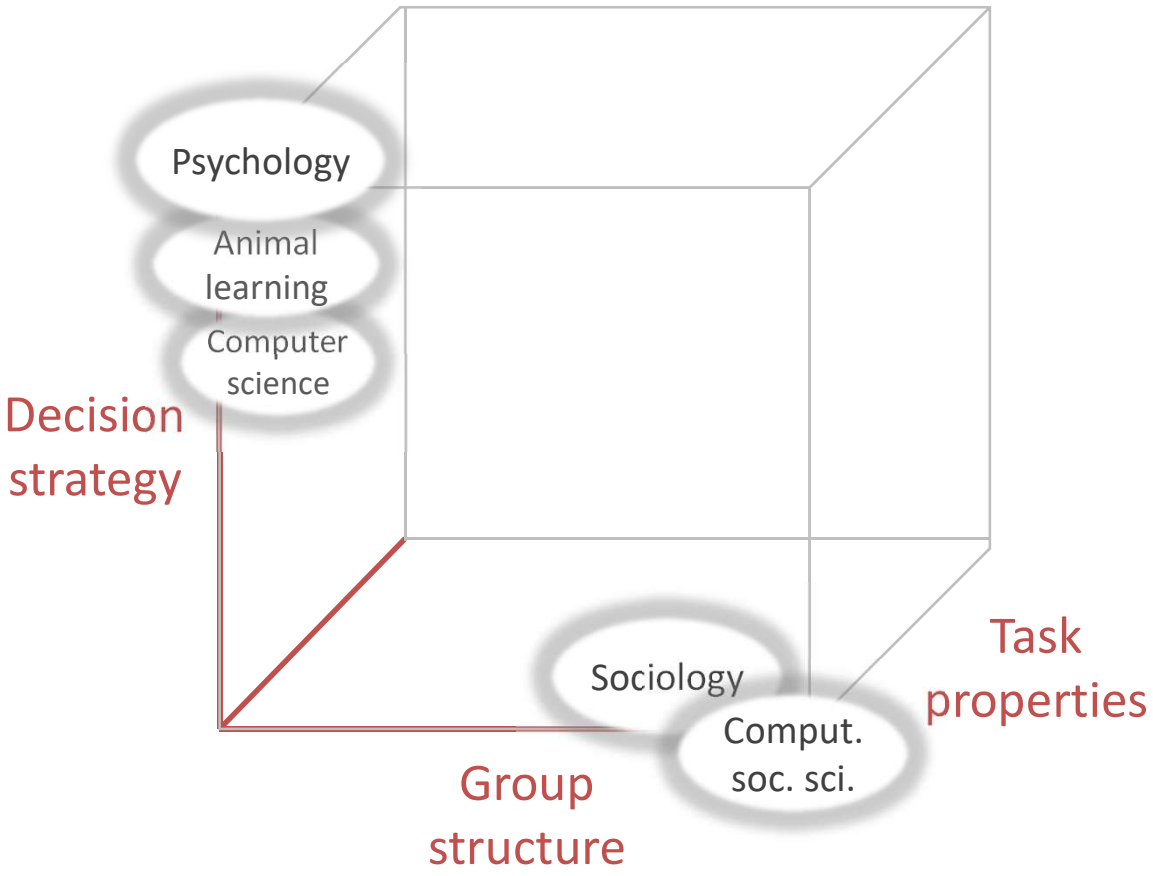
… in computer science



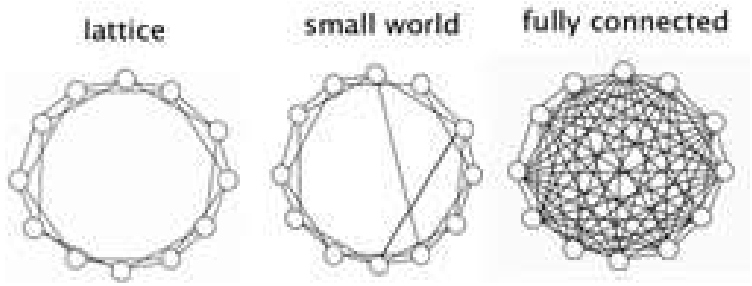
Machine learning ensembles

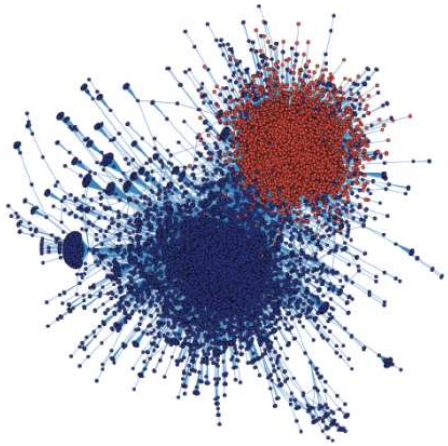
Saha et al, 2014; Rokach, 2012

Views of different disciplines

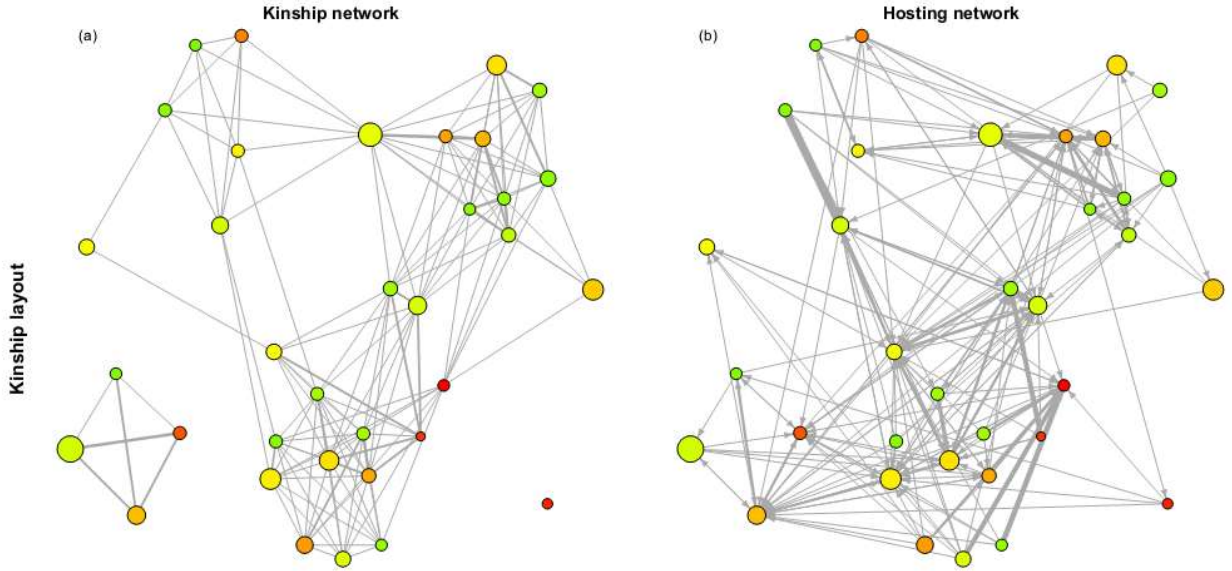


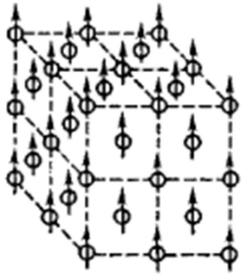
Group structure





Conover et al (2011)

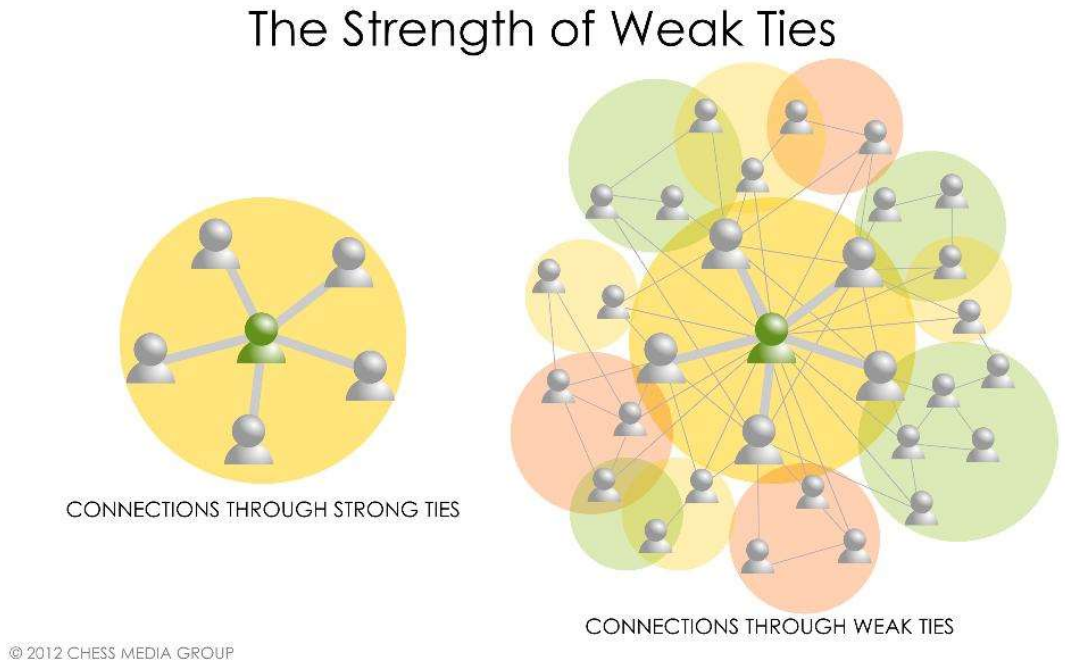




Hooper, DeDeo, Caldwell-Hooper, Gurven, & Kaplan (2013, Entropy)

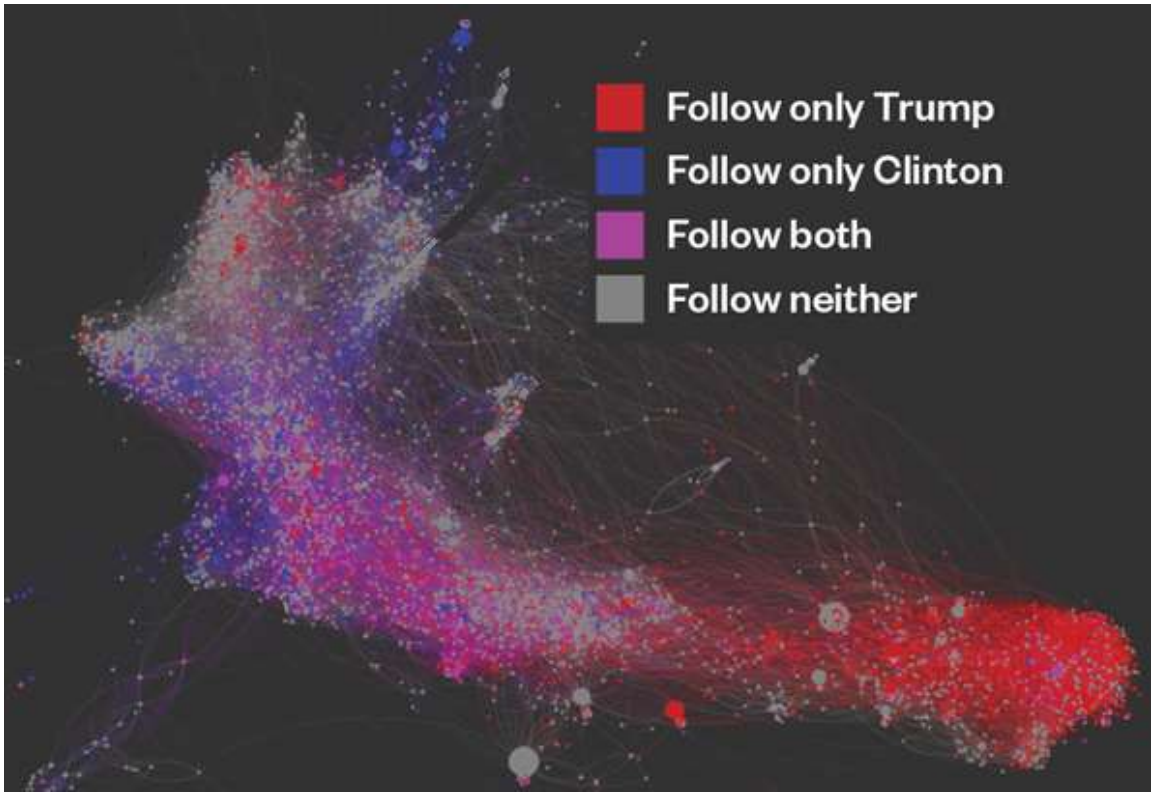
… in sociology

- Mechanisms of social contagion
  - Granovetter (1973, 1978), Centola (2007, 2010)



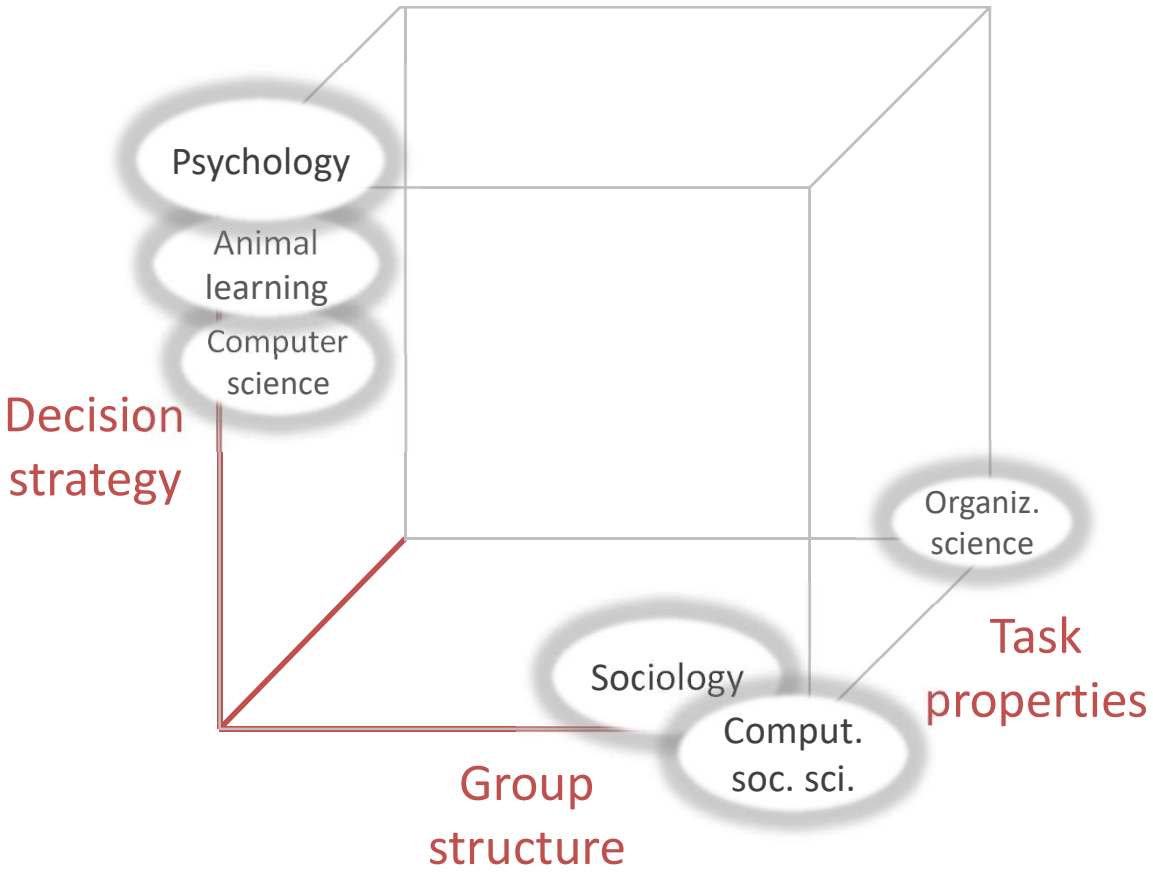
… in computational social science

- Reconstructing social networks from big data

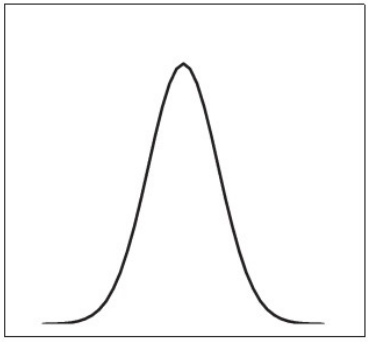


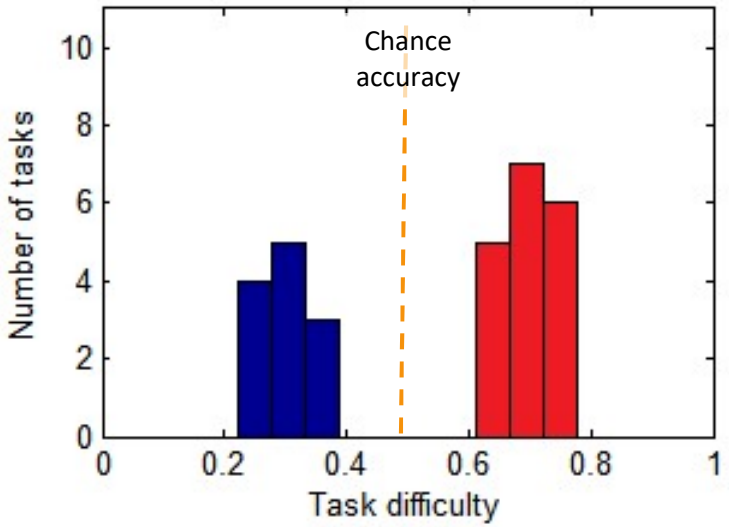
http://www.electome.org/, Twitter+MIT

Views of different disciplines



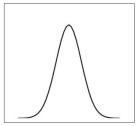
Task properties



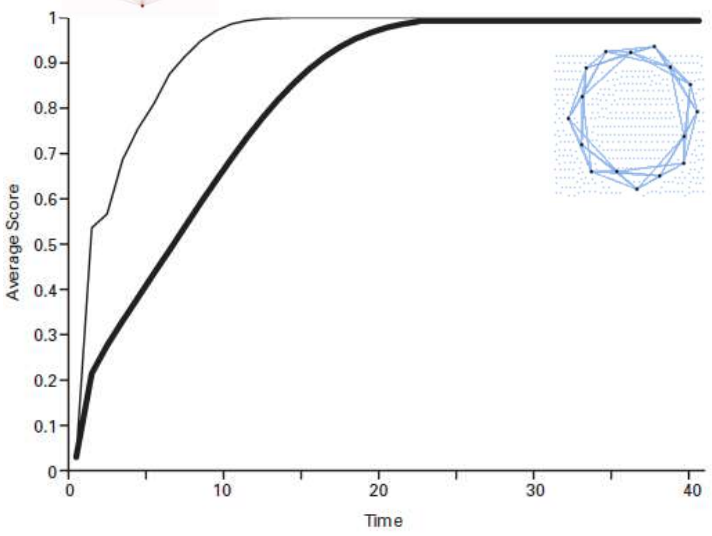


… in organization science

Performance of different group structures

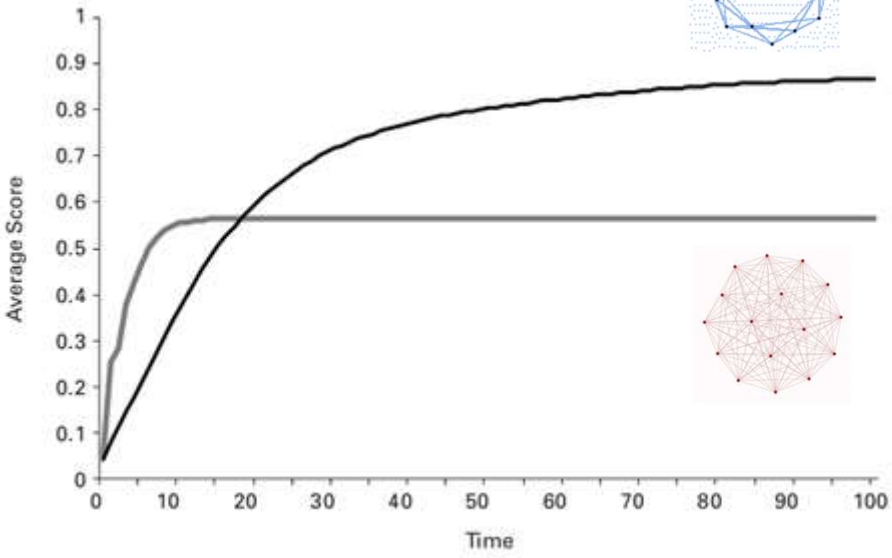


in simple tasks

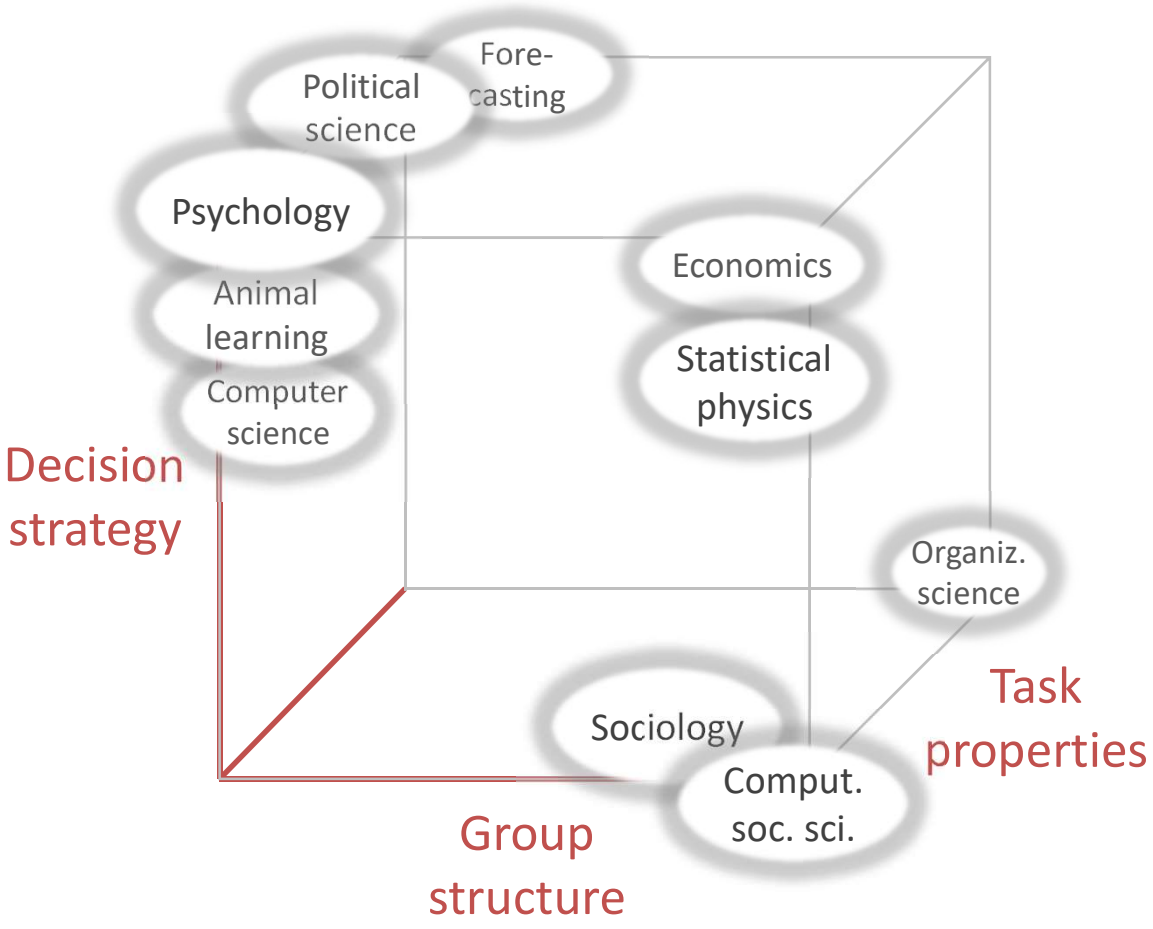


Lazer & Friedman, 2007

in complex tasks



Views of different disciplines

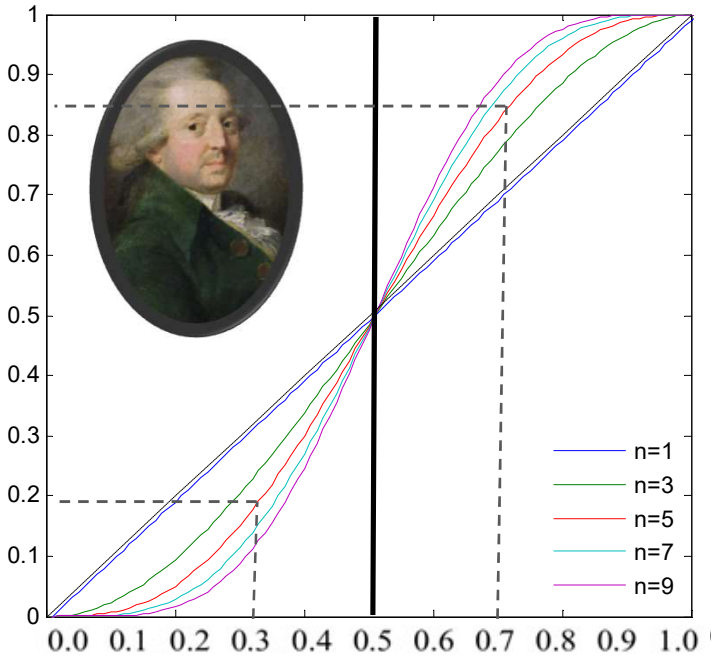


Political science

- Voting models for aggregation of preferences and information
  (Dewan & Shepsle, Annu Rev Polit Sci, 2011)

- Example: Condorcet Jury Theorem

Accuracy of group majority



Average individual accuracy

$$M = \sum_{i=m}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

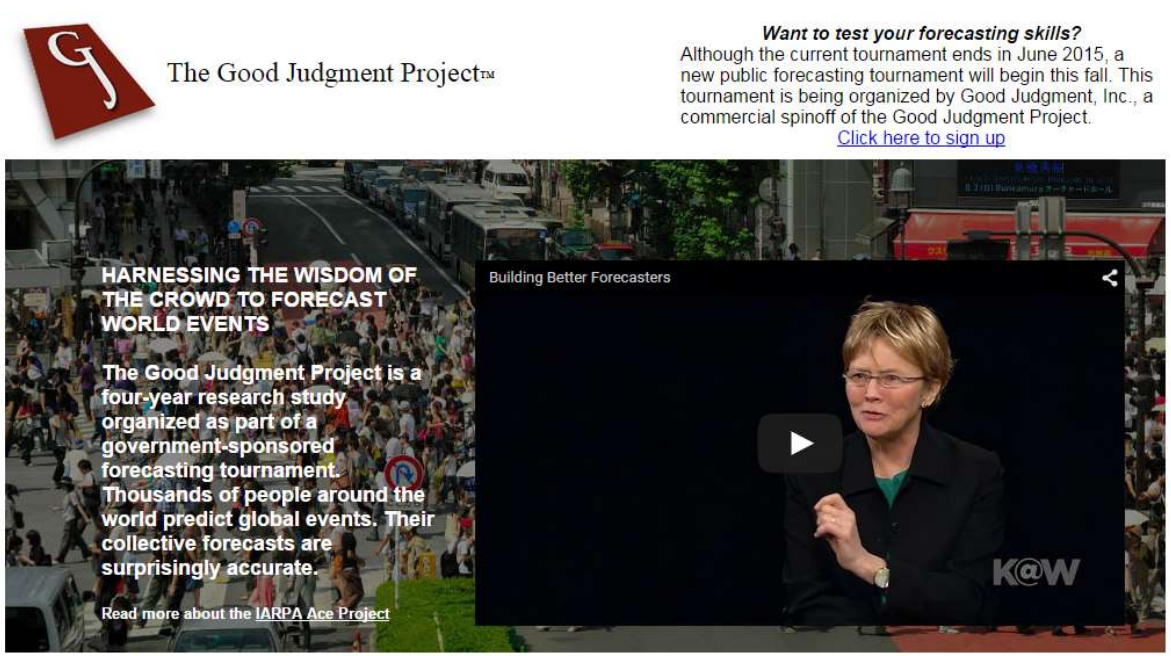
Grofman, Owen, & Feld (1983, Th Dec). 13 Theorems in Search of the Truth.

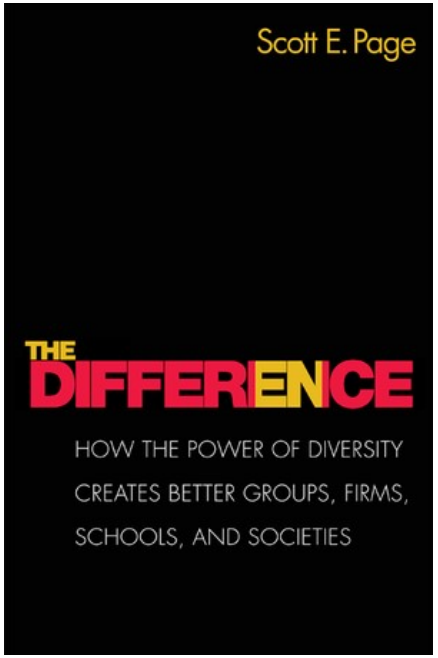
List & Goodin (2001, J of Pol Phil)

Forecasting

When and why groups perform better than individuals?

How to elicit and aggregate expert forecasts?





Economics

Opinion formation on networks

- Bayesian updating models

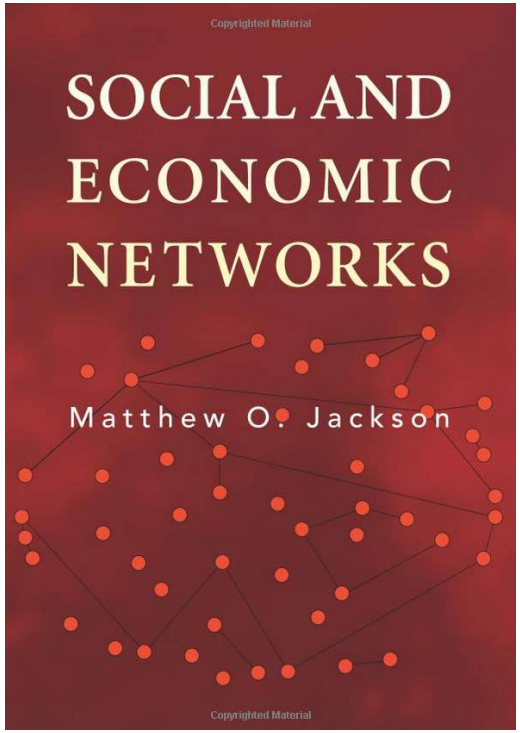
$$P(\theta|s) = \frac{P(s|\theta)P(\theta)}{P(s)}$$

$\boldsymbol{P(\theta)}$ – prior belief
$\boldsymbol{s}$ – social signal

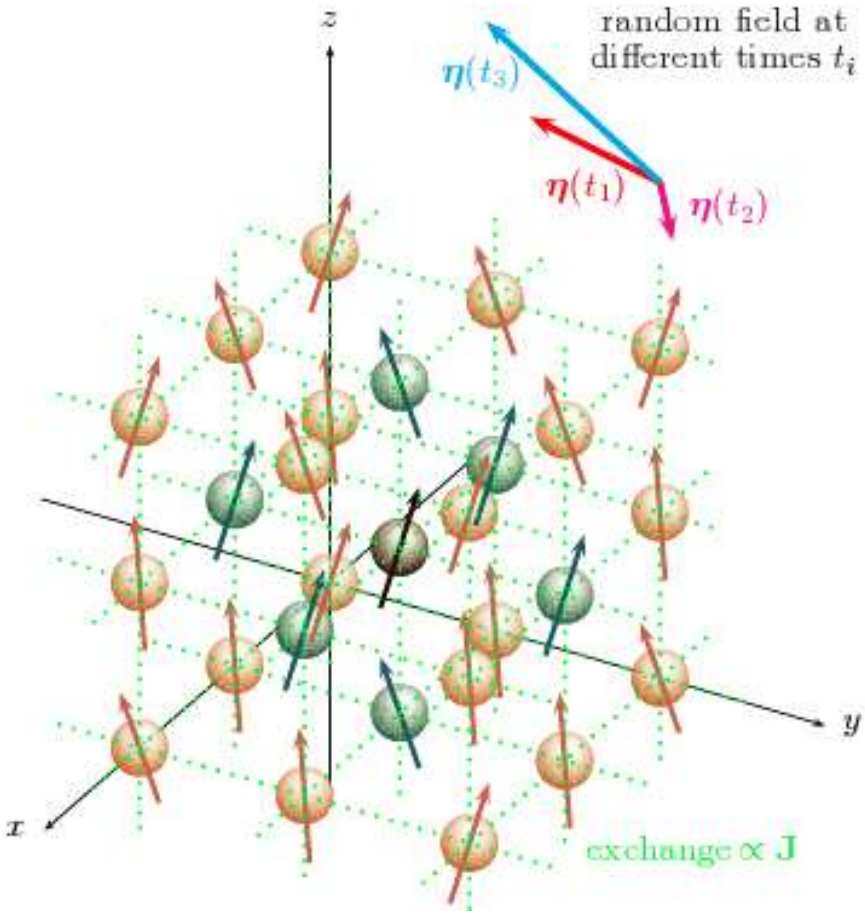
- Non-Bayesian models
  - DeGroot model

$$x_{i,t+1} = \sum_{j=1}^{n} T_{ij} x_{j,t}$$

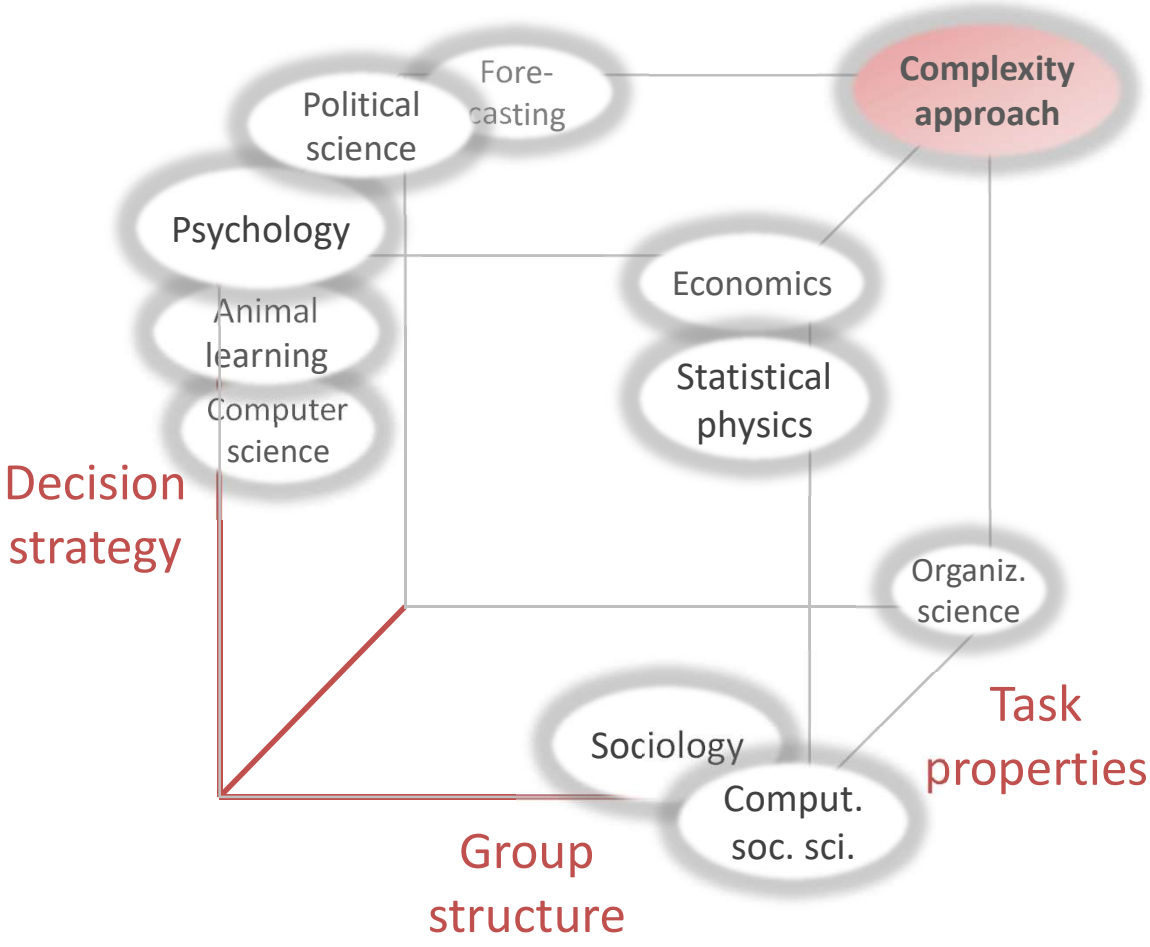
x_i– belief of agent i
T_{ij} – trust of i in j's signal



Statistical physics

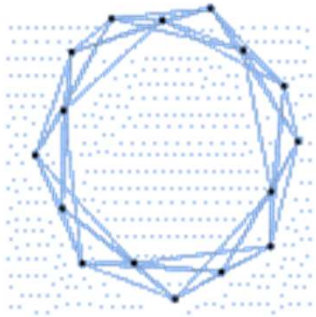


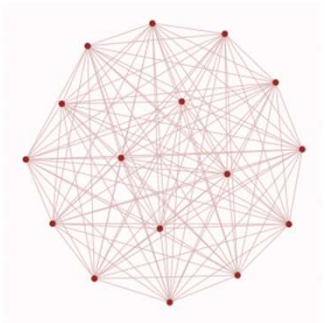
Need for complexity approach



Need for complexity approach

Example 1: Which networks are best for collective problem solving?



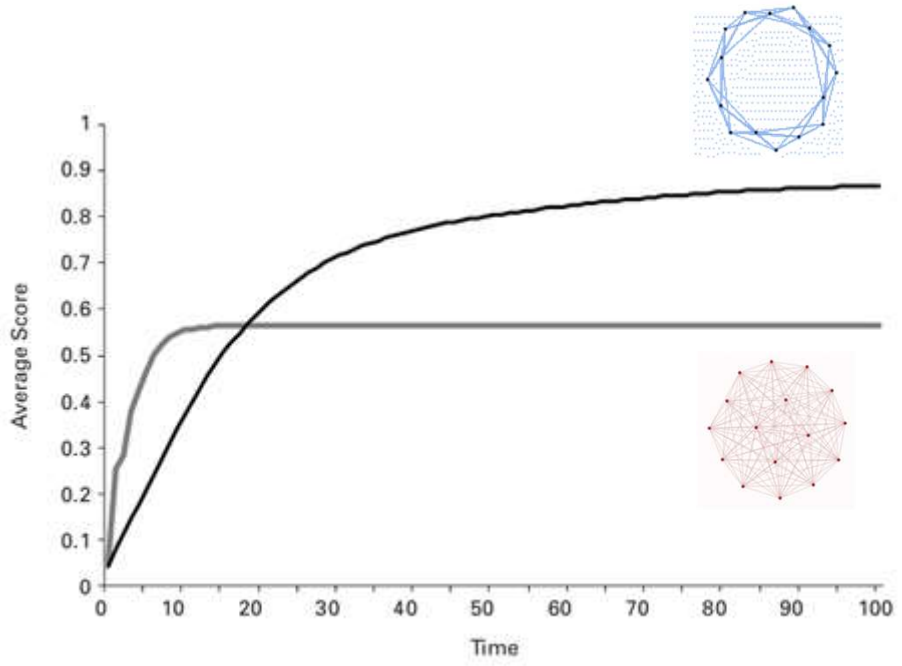


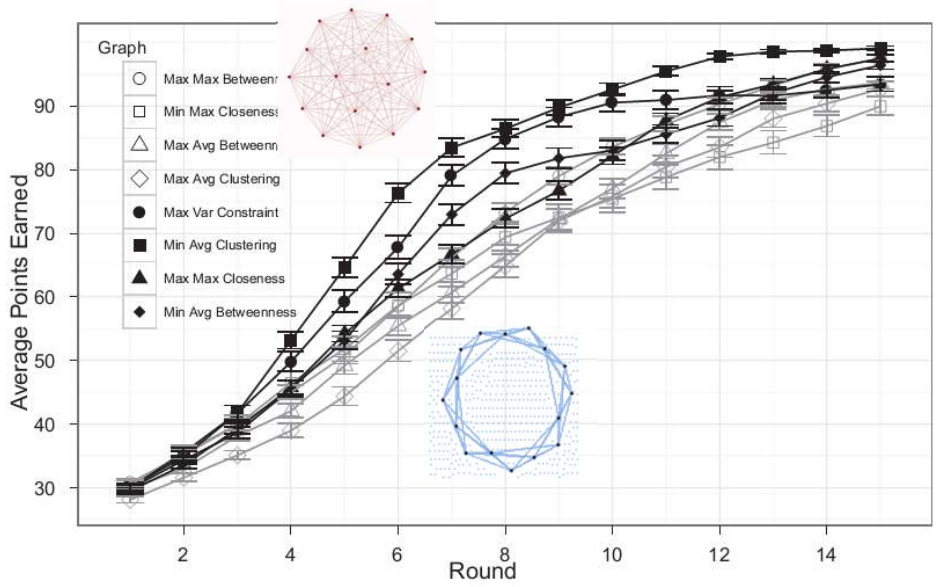
Studies disagree:

- Poorly-connected, slow networks better (Lazer & Friedman, 2007, ASQ; Derex & Boyd, 2016, PNAS)

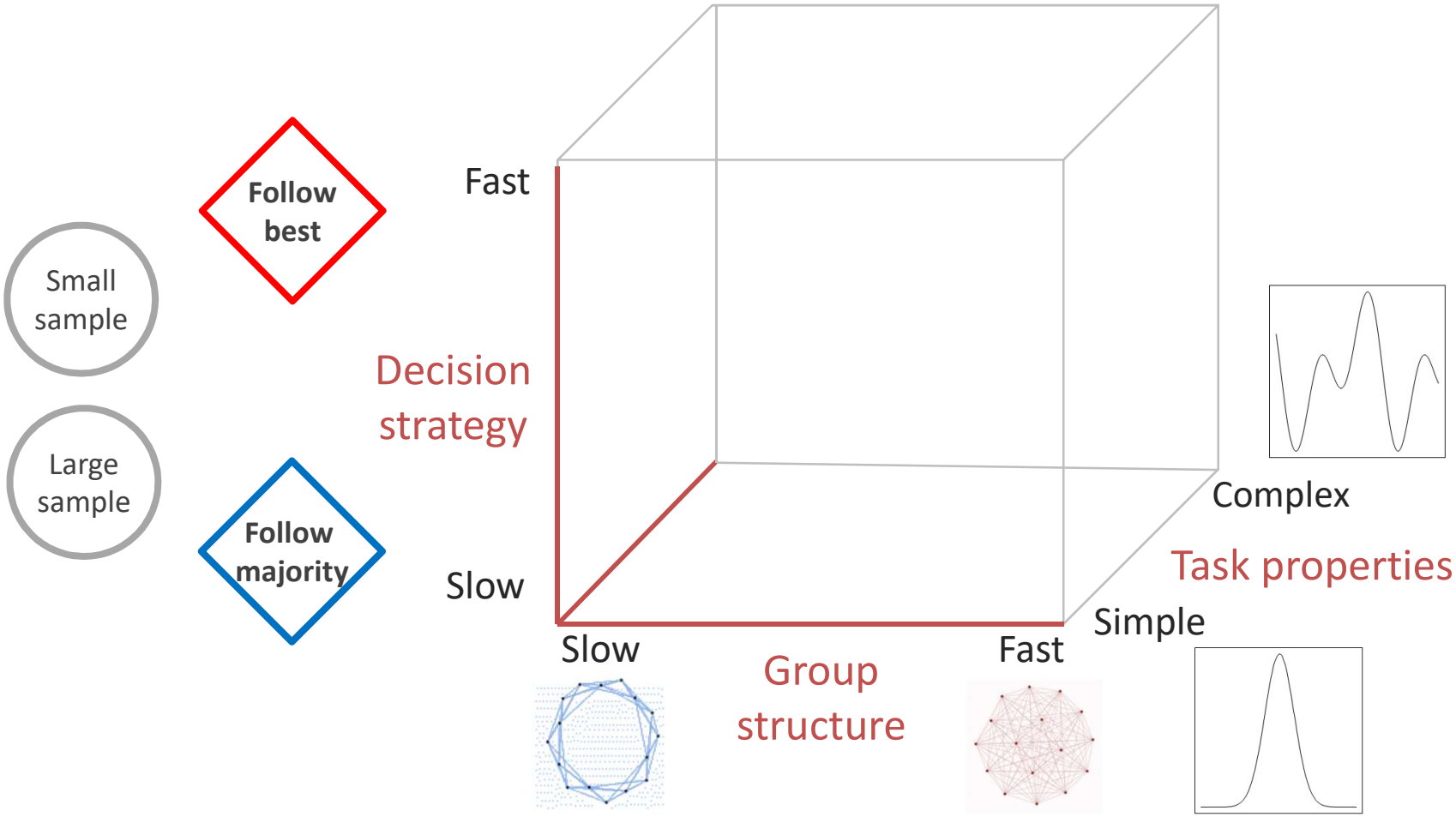
- Well-connected, fast networks better (Mason & Watts, 2012, PNAS)

→ Studies focused on group structure; disregarded decision strategies and task properties





Dimensions of collective problem solving



Computational study

100 agents, in different networks, solve simple or complex tasks over 200 time steps:

Step 1. Imitate another agent, if better

Choice of agent depends on decision strategy

Step 2. Otherwise, explore alone

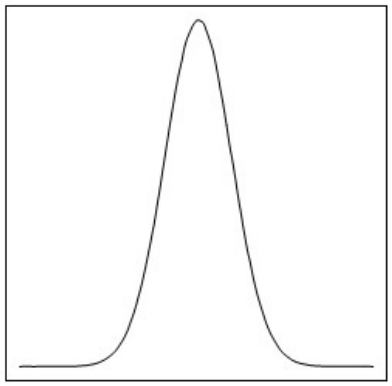
Record average payoff of agents on each time step.

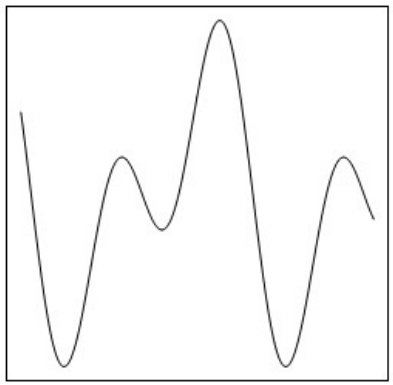
Barkoczi & Galesic (2016), *Nature Communications*

Task properties

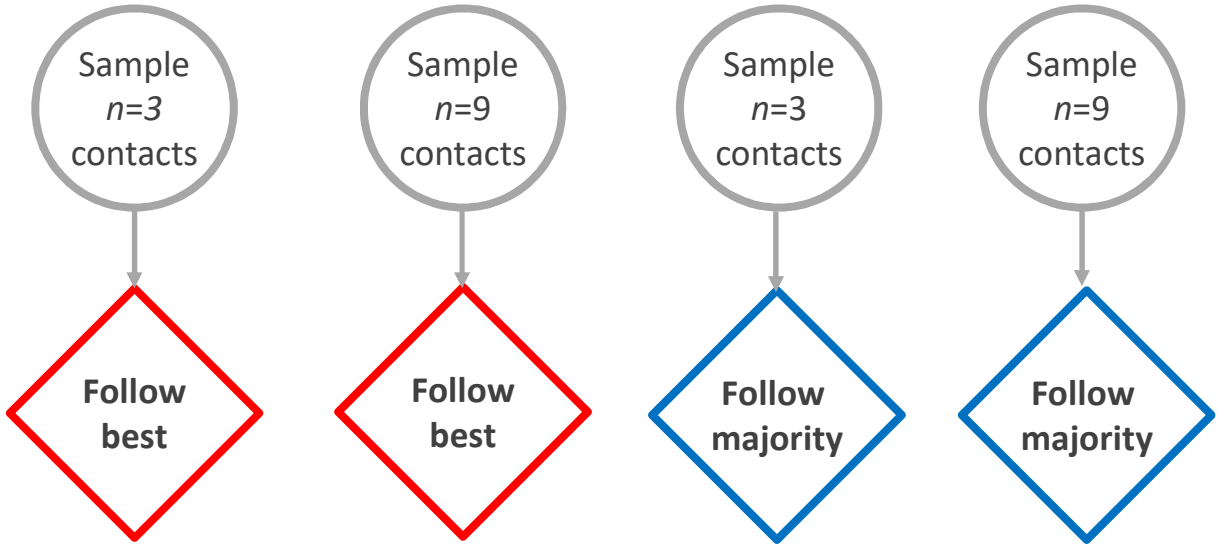
NK environments:

- Solutions are strings of N elements (000101001)
- Fitness of each element depends on K other elements
- N = 15
- K varies from 0 (“Simple environment”) to 14 (“Complex environment”)



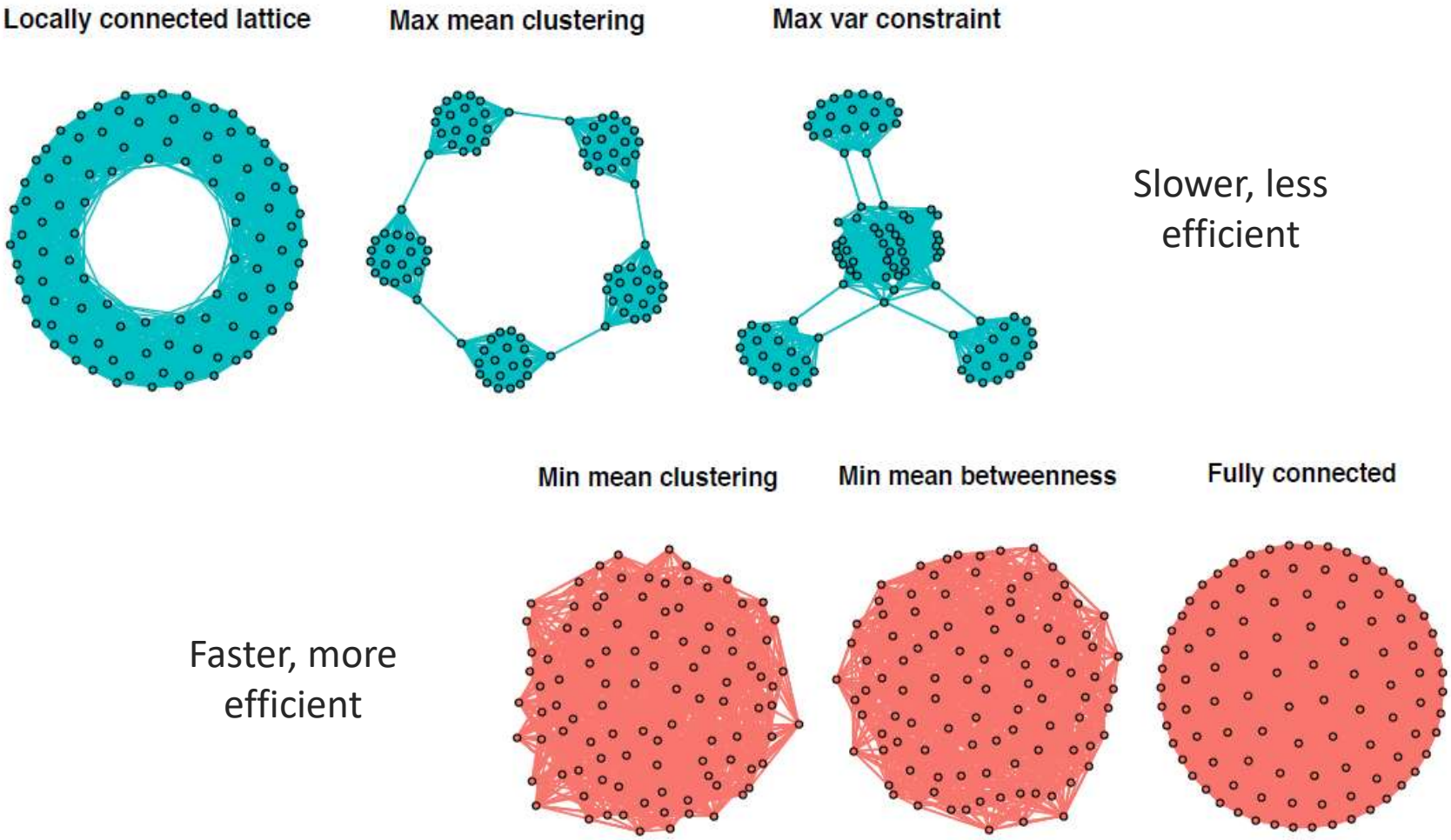


Decision strategies



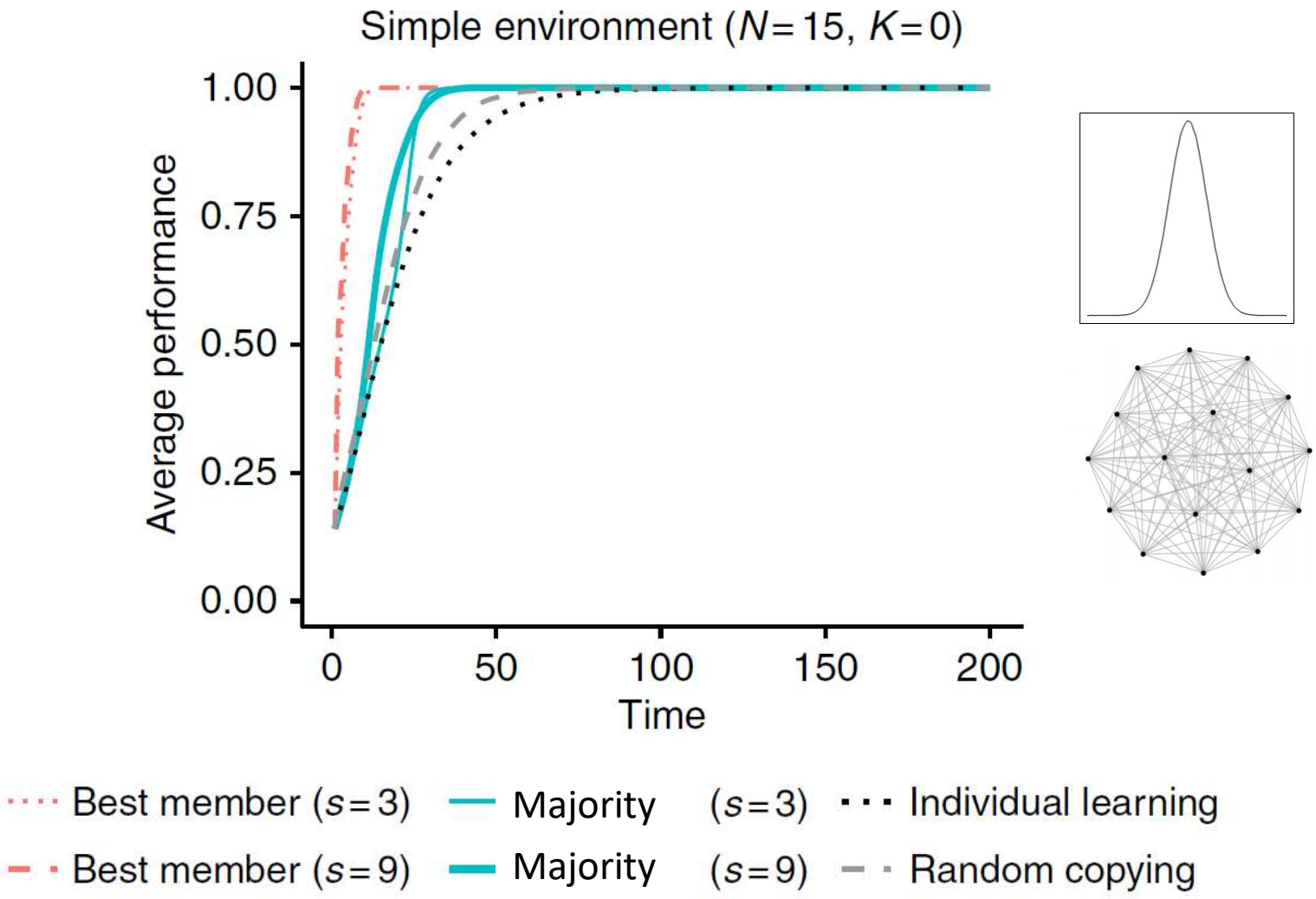
+ for comparison: random copying and individual learning

Group structure - Networks

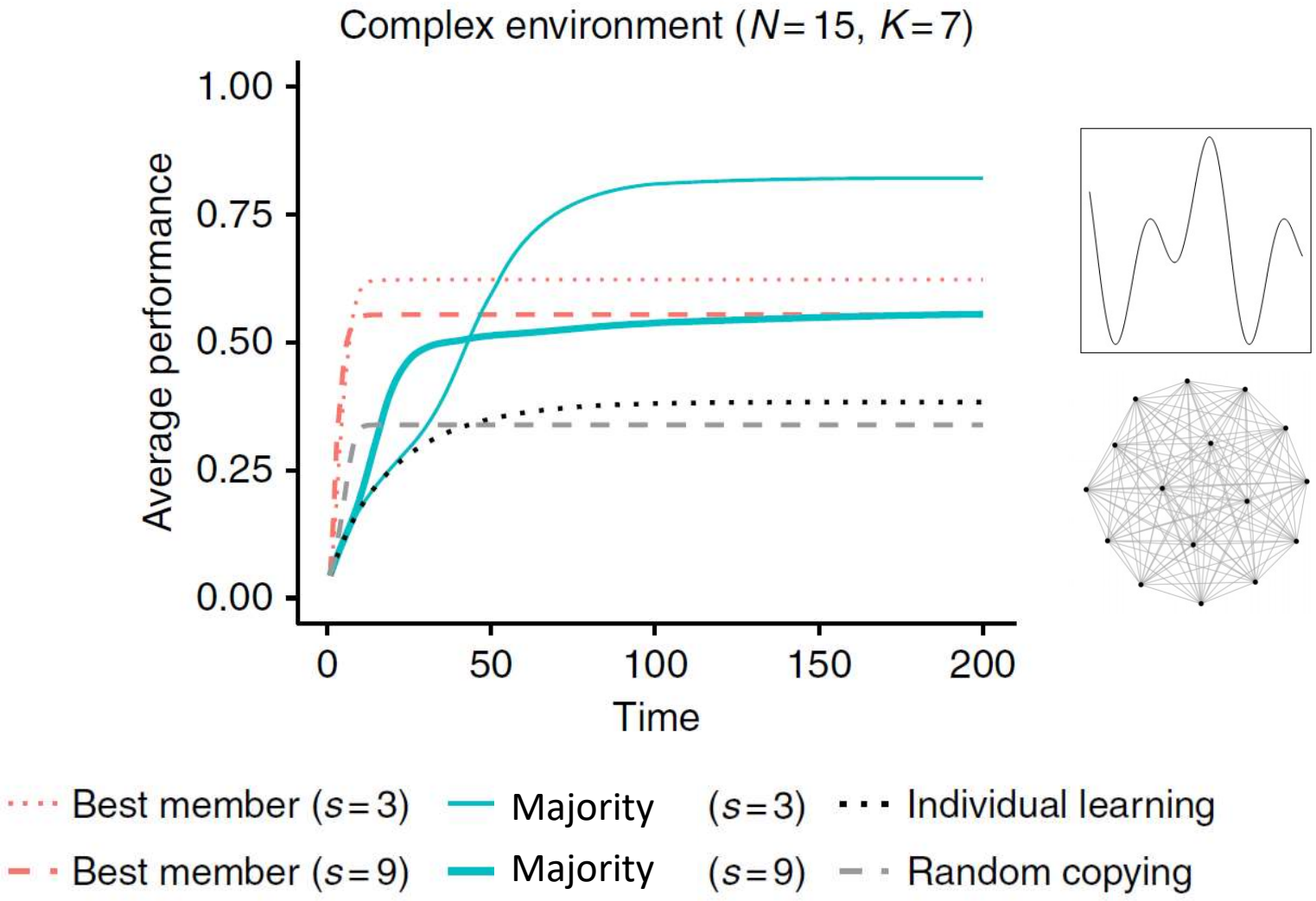


Results

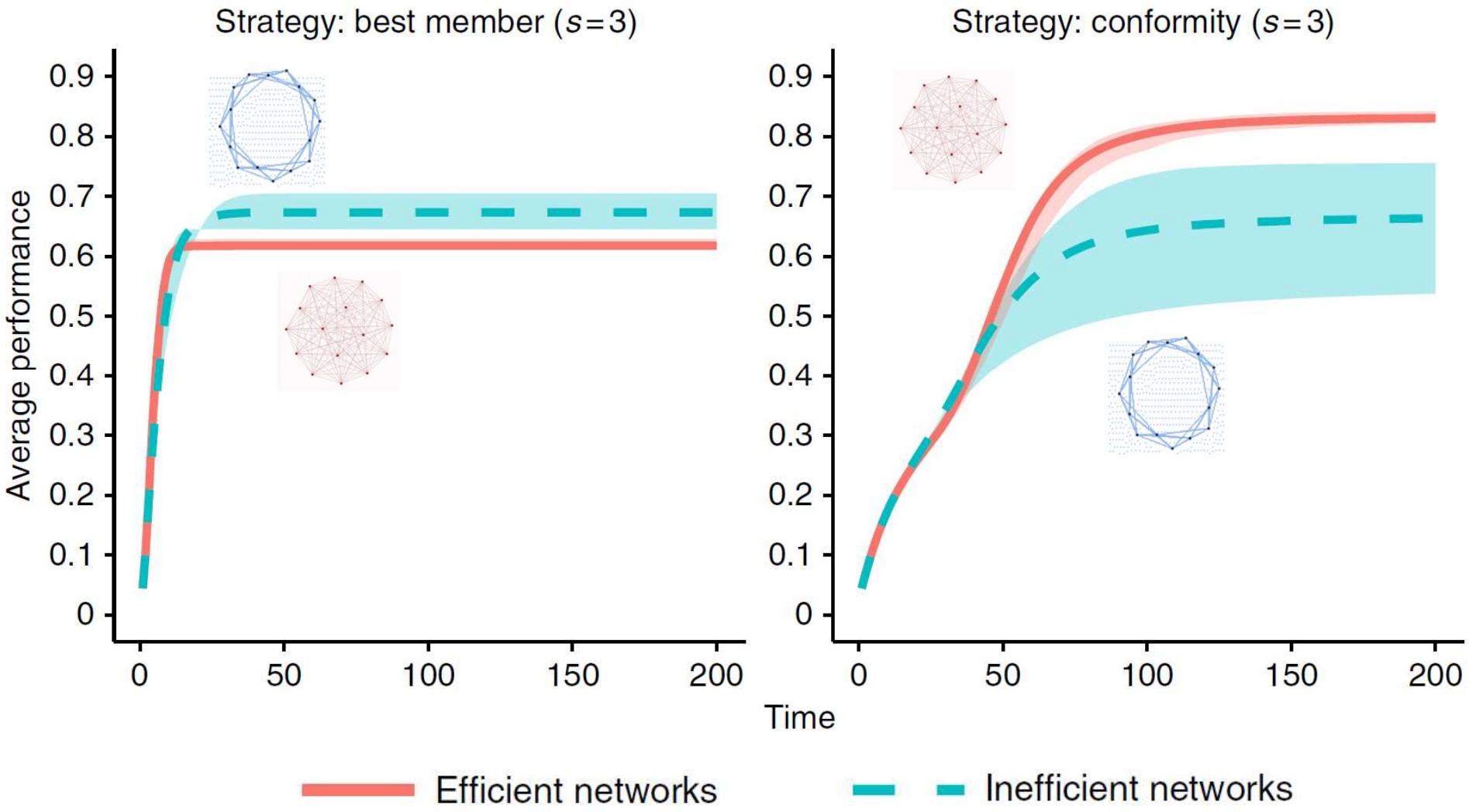
Fast strategy good when task simple



Slow strategy good when task complex

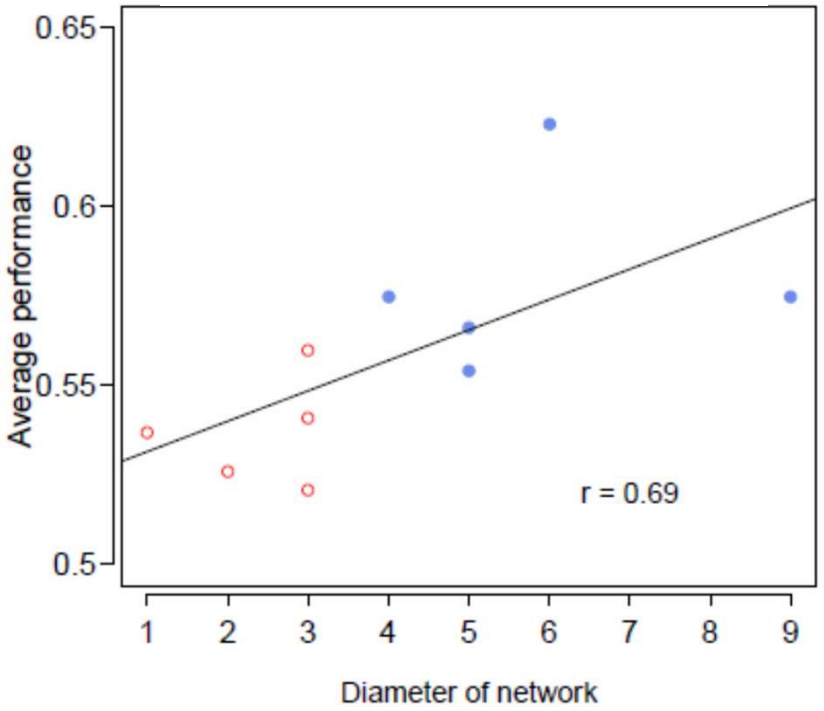


Fast networks better with slow strategies

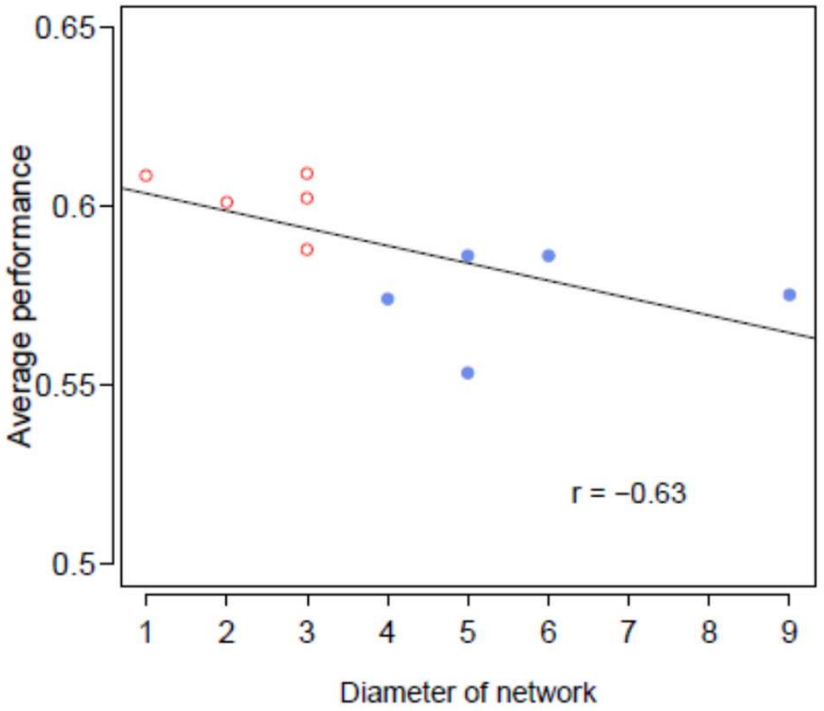


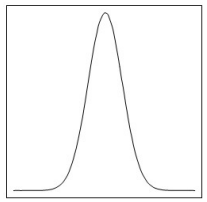
Fast networks better with slow strategies

Follow best member

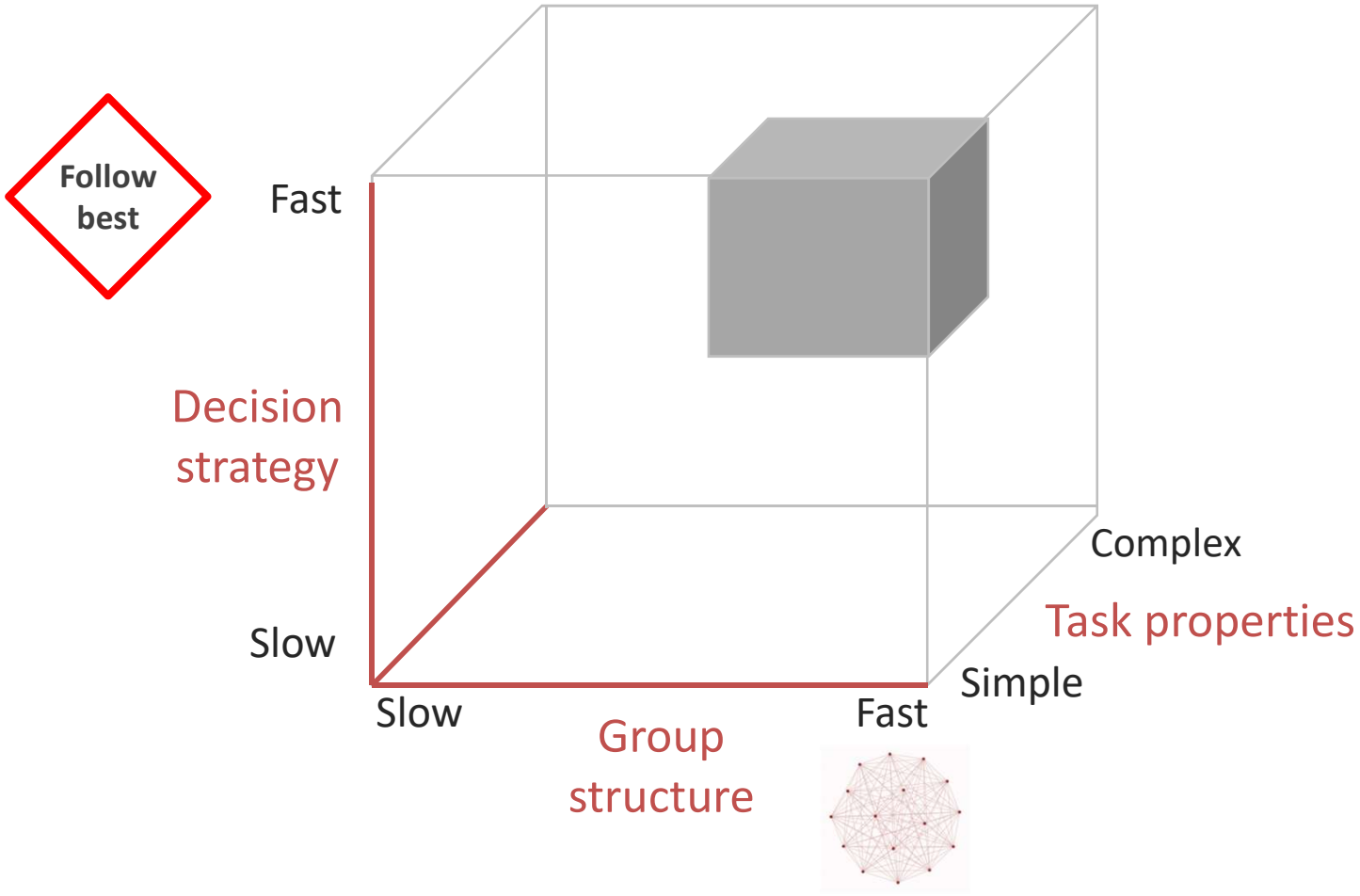


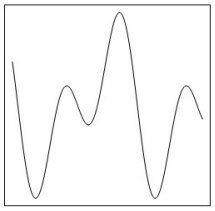
Follow majority



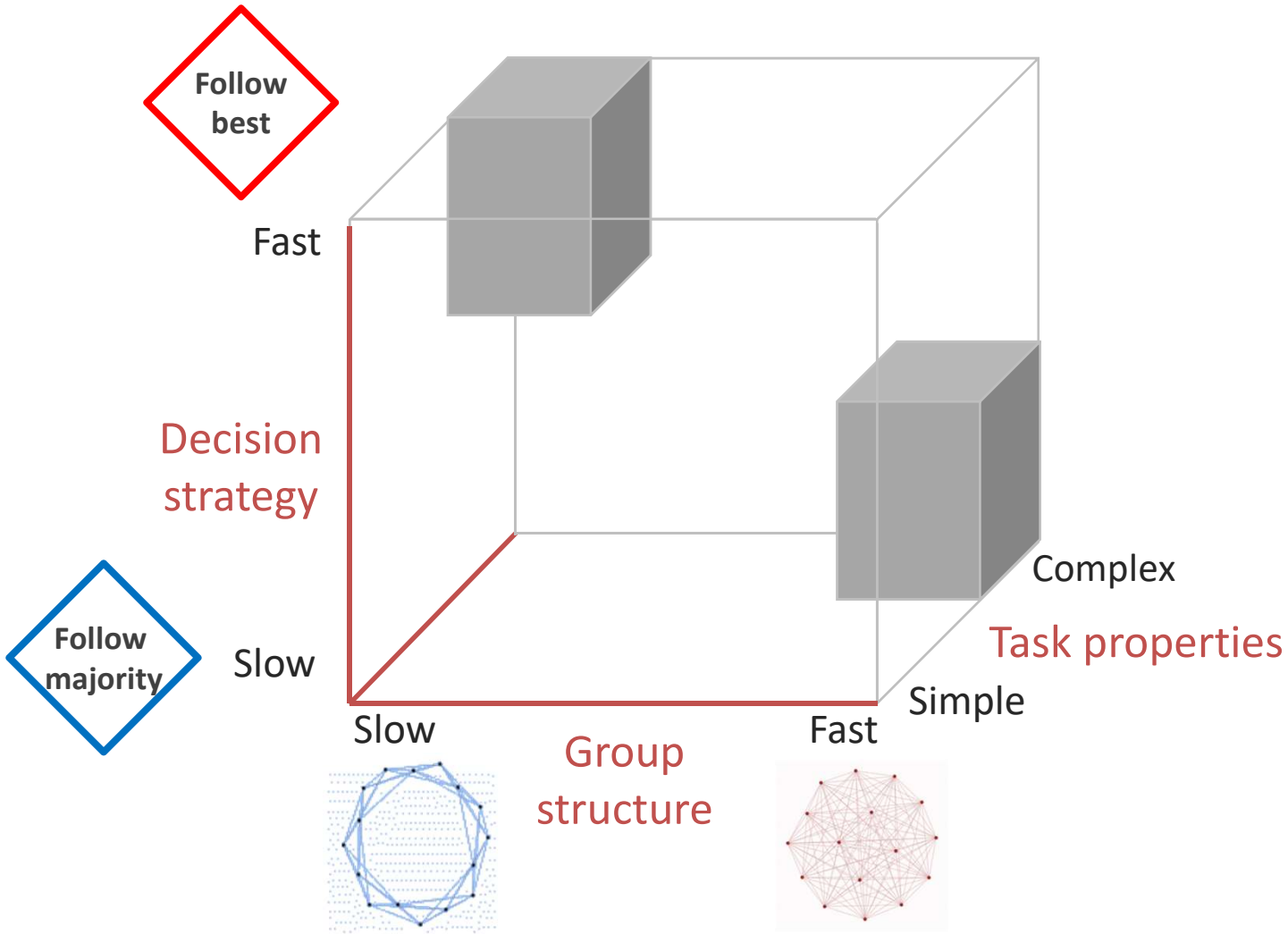


Simple tasks





Complex tasks



Implications

- There is no one best group structure
- When designing an organization, it matters what are the desired (or available):
  1. Network/communication structures
  2. Decision making procedures/strategies
  3. Task properties

Need for complexity approach

Example 2: What is the best size of committees that need to make many decisions over time?

Typical committees

http://www.federalreserve.gov/

www.med.upenn.edu/criticalcare/

www.rhuddlantowncouncil.gov.uk/

Galesic, Barkoczi, & Katsikopoulos (2016), *Decision*

Typical committee sizes

Jury sizes in most countries: 6-15 people

Town councils in UK and Australia: 5-30

Parliamentary committees in US, EU, Australia: 20-40

US House and Senate subcommittees: 10-15

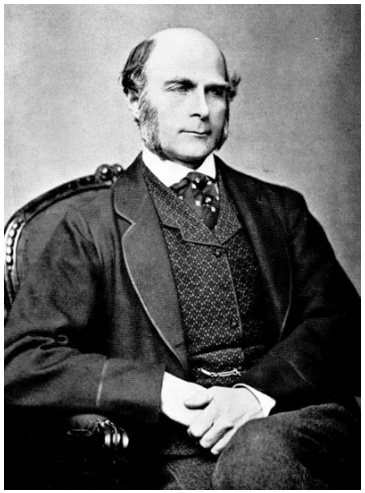
Central bank boards: 5-12

Number of close friends: 6 or less

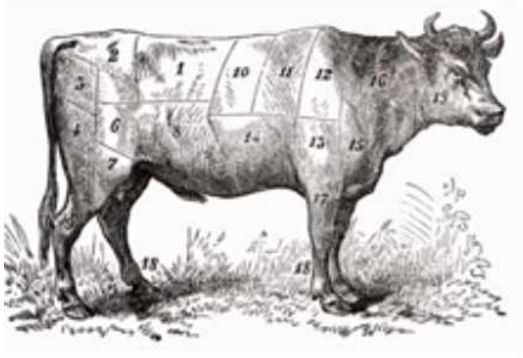
Number of online reviews read: average 5, max 30

If there is "wisdom of crowds", why are committees so small?

Wisdom of crowds



Francis Galton, 1906



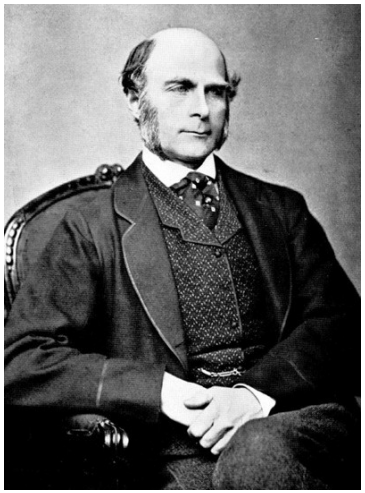
average of 800 guesses = 1,197
actual weight of the ox = 1,198

When a group estimates a quantity by averaging individual guesses
→ Larger group will always have smaller error (as long as it is diverse)

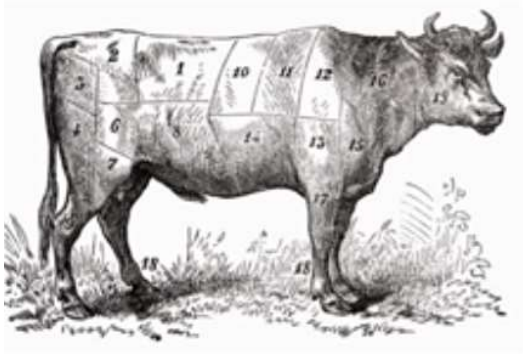
$$MSE = \overline{\text{bias}}^2 + \frac{1}{M}\overline{\text{var}} + \left(1 - \frac{1}{M}\right)\overline{\text{cov}} + \sigma_Y^2$$

→ Total error = bias + variance + covariance + irreducible error
(*M*=group size)

Wisdom of crowds



Francis Galton, 1906



average of 800 guesses = 1,197
actual weight of the ox = 1,198

When a group estimates a quantity by averaging individual guesses
→ Larger group will always have smaller error (as long as it is diverse)

But when a group needs to decide on one of several courses of action
→ Larger group is not always better

Group error components

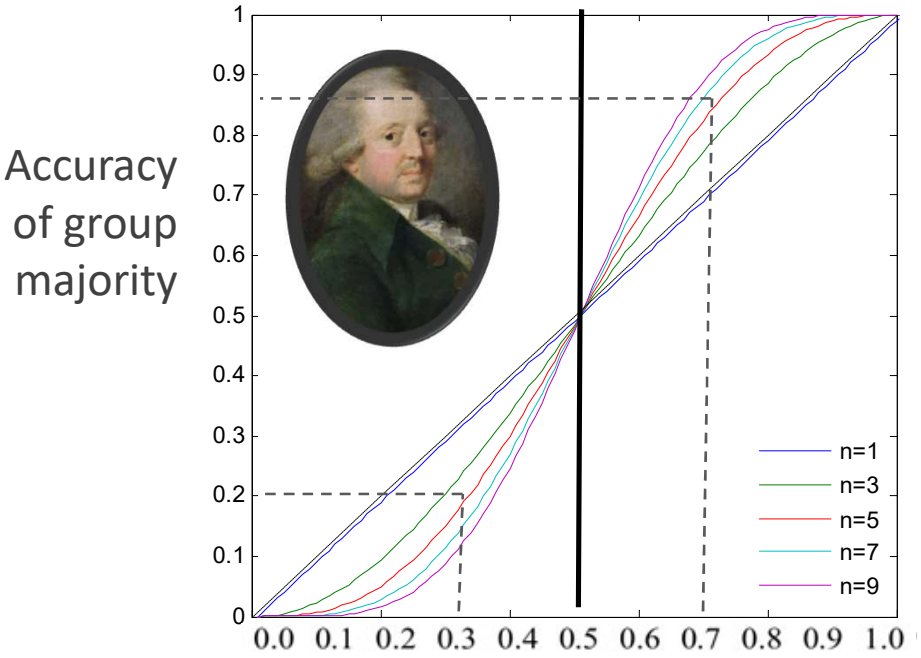
$$MSE = \overline{\text{bias}}^2 + \frac{1}{M}\overline{\text{var}} + \left(1 - \frac{1}{M}\right)\overline{\text{cov}} + \sigma_Y^2$$

Typical committee tasks

- Group of experts vote by **simple majority** to make decisions and predictions such as:
  - Will the economy grow or fall in the next period?
  - Which policy should we adopt?
  - Should we make this investment or not?
  - What is patient's diagnosis?
  - Should we attack or not?

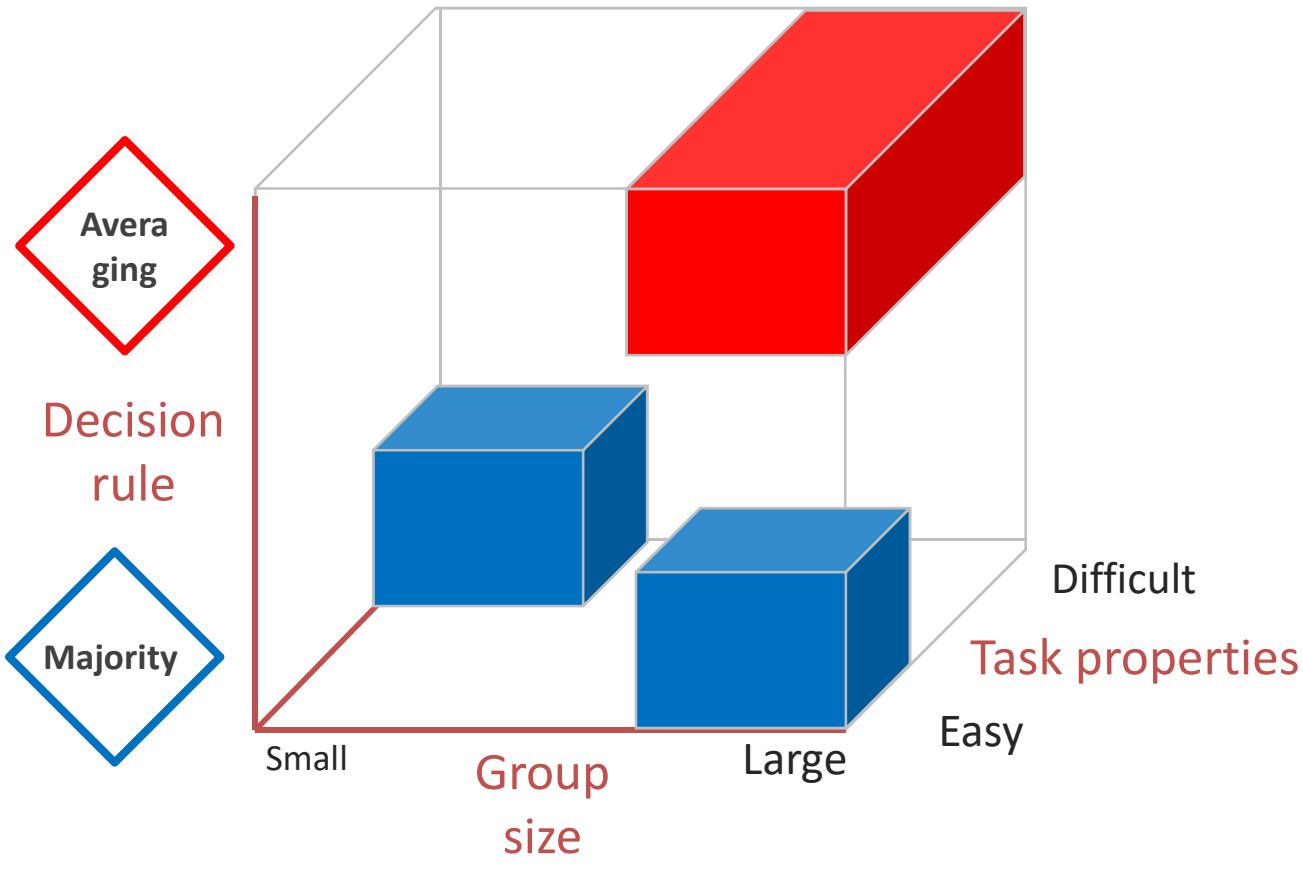
Accuracy of simple majority rule for a **single** task

Condorcet Jury Theorem



$$M = \sum_{i=m}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

Best group size depends on decision strategy and task difficulty

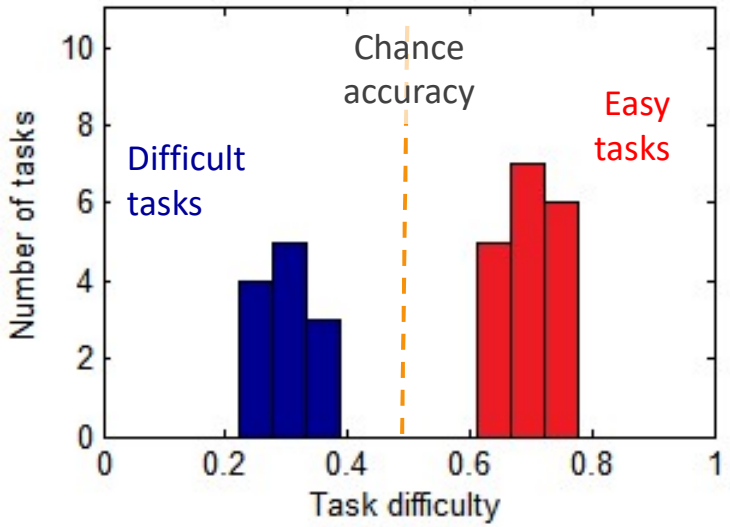


Accuracy over many tasks?

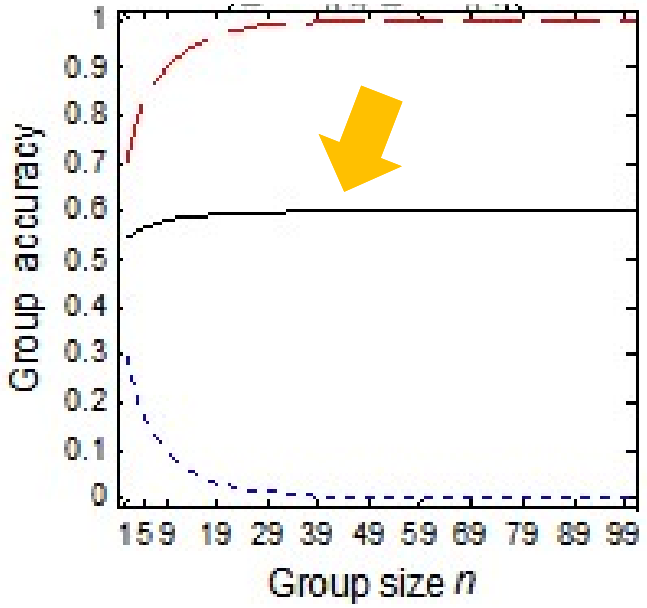
- Committees encounter many tasks over the course of their existence
  - Some tasks are easy
  - Some tasks are surprisingly difficult
  - **We can't know in advance how easy or difficult the next task is going to be**

Accuracy of simple majority rule across **many** tasks

"Neutral" task environment



Proportion of easy tasks: $e = 0.6$



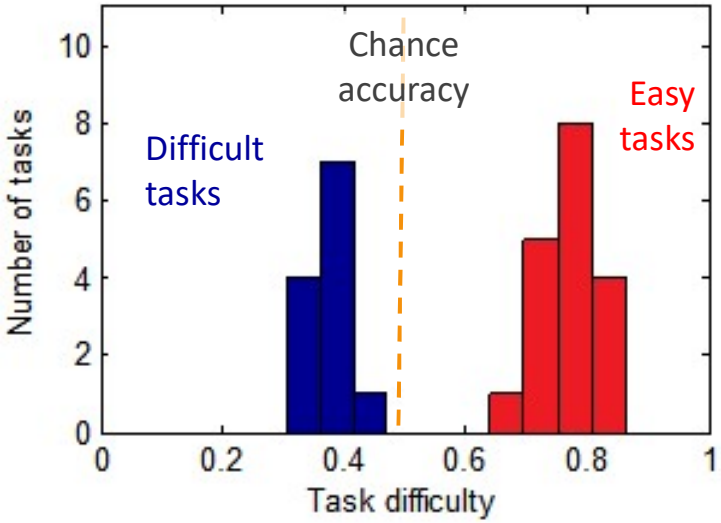
Group accuracy for easy tasks

Group accuracy for difficult tasks

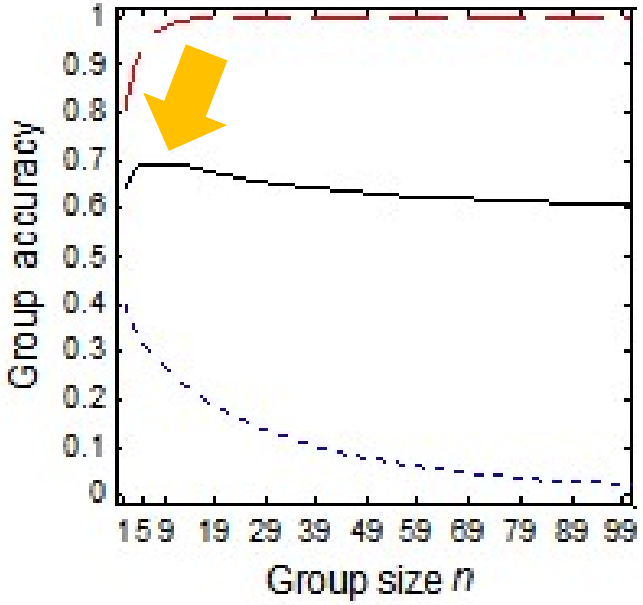
Average group accuracy

Accuracy of simple majority rule across **many** tasks

"Friendly" task environment



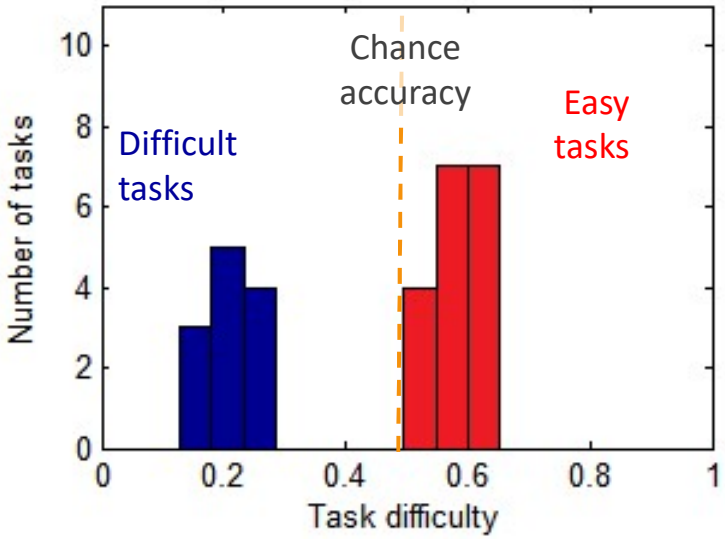
Proportion of easy tasks: e = 0.6



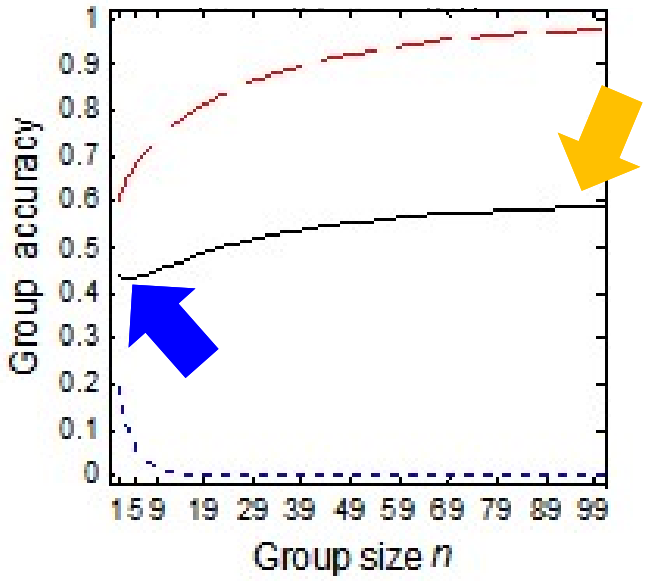
Group accuracy for easy tasks
Group accuracy for difficult tasks
Average group accuracy

Accuracy of simple majority rule across **many** tasks

"Unfriendly" task environment

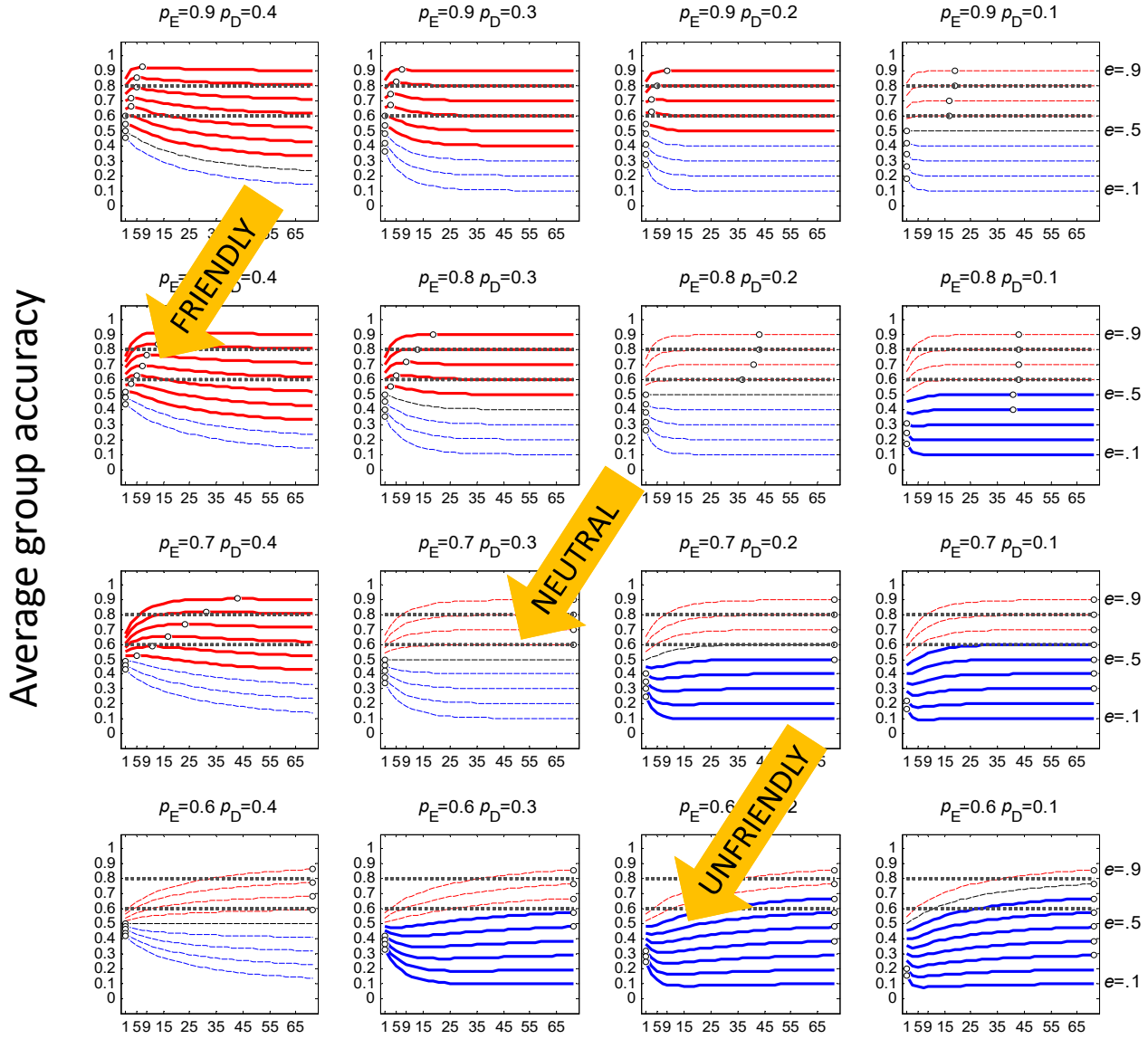


Proportion of easy tasks: $e = 0.6$

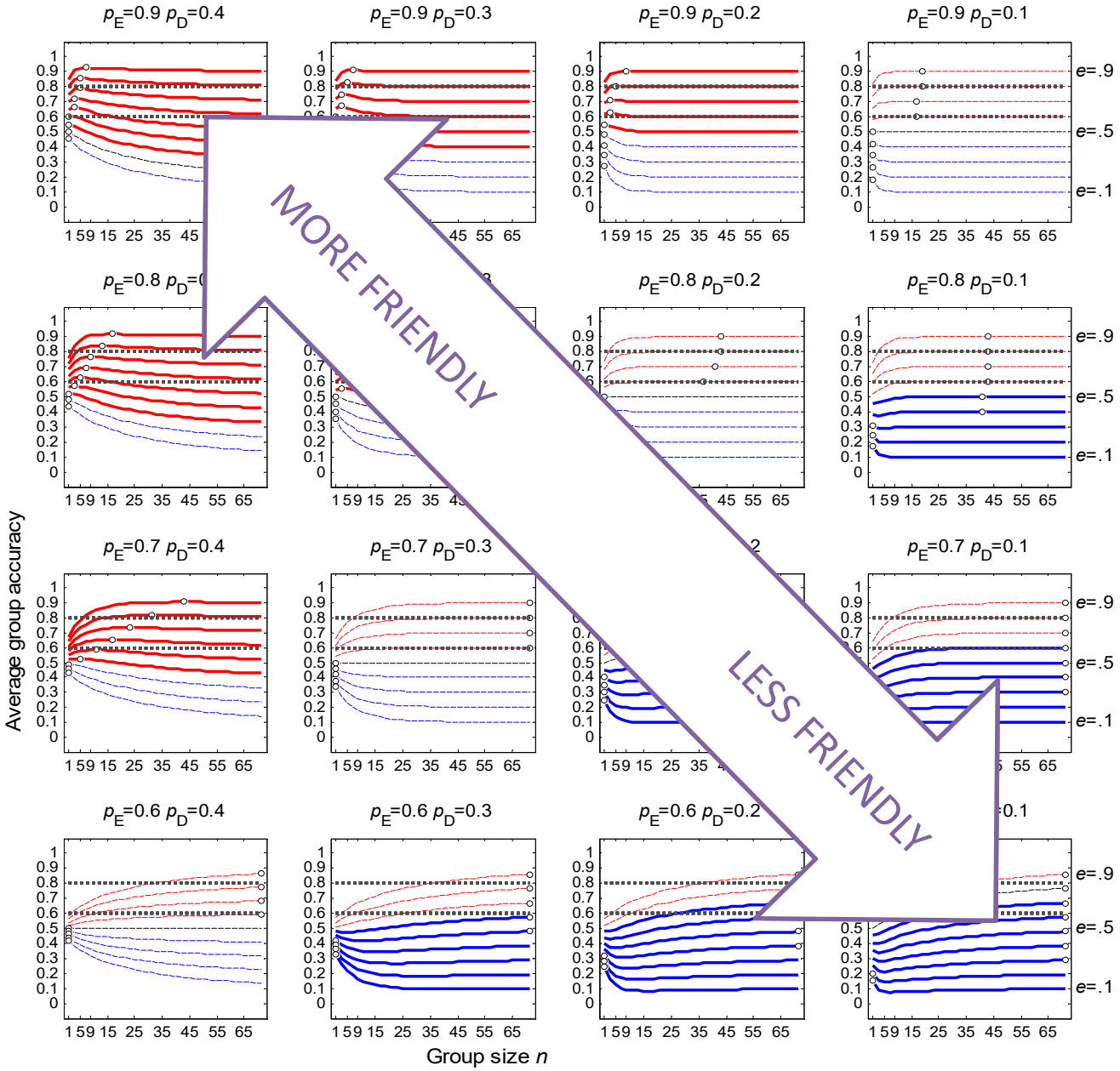


Group accuracy for easy tasks
Group accuracy for difficult tasks
Average group accuracy

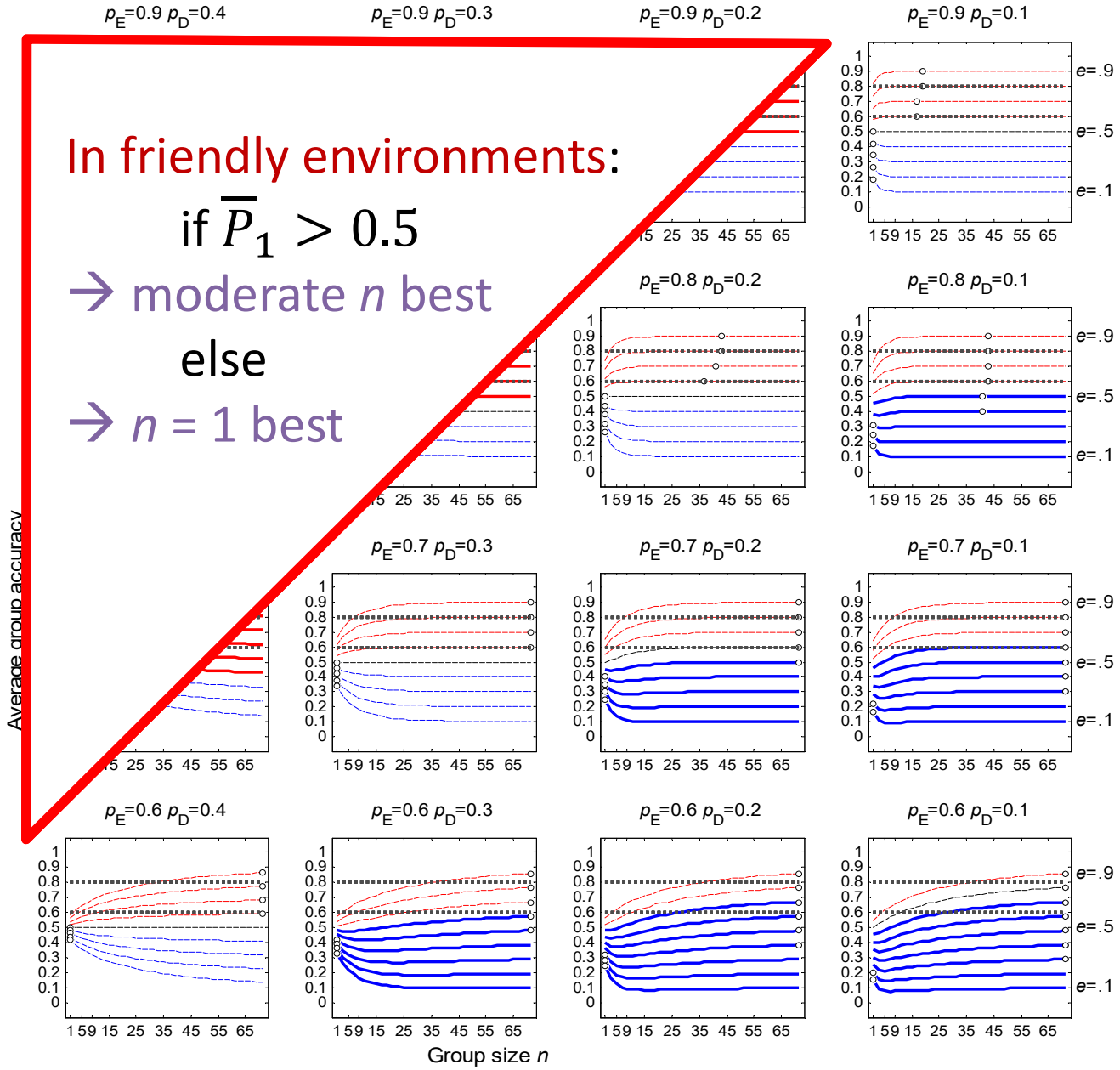
All possible task environments



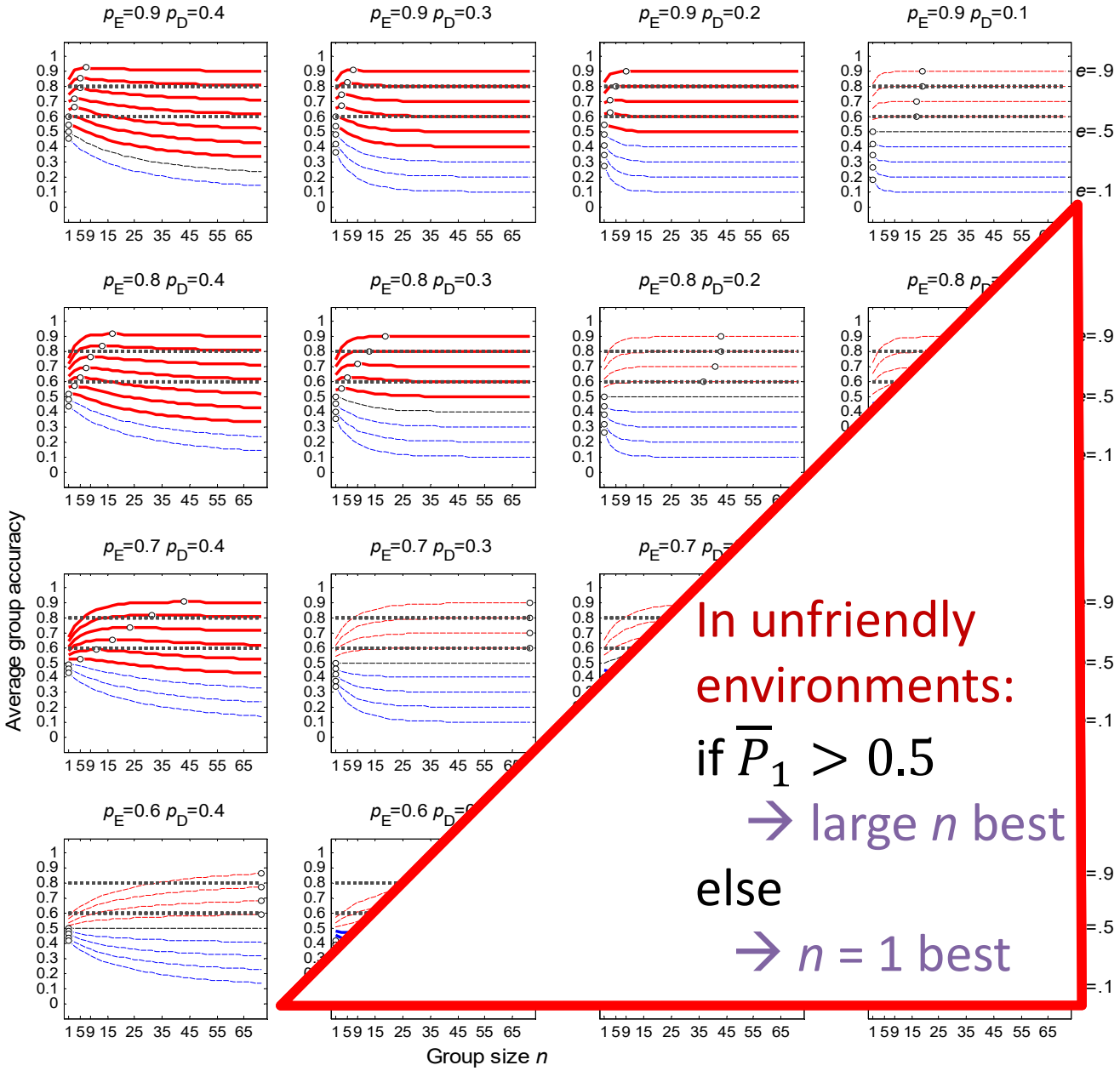
All possible task environments



All possible task environments



All possible task environments



More formally

Average group accuracy across many tasks:

$$\overline{P}_n = e\, P_{\mathrm{E},n} + (1-e) P_{\mathrm{D},n}$$

$\bar{P}_n$ → average accuracy of group of size *n* across tasks

e → proportion of easy tasks that group needs to solve

$P_{\mathrm{E},n}$ ($P_{\mathrm{D},n}$) → accuracy of group of size *n* on easy (difficult) tasks

Friendly environment:

$$\overline{p}_{\mathrm{E}} - 0.5 > 0.5 - \overline{p}_{\mathrm{D}} \quad \rightarrow \quad \overline{p}_{\mathrm{E}} + \overline{p}_{\mathrm{D}} > 1$$

$\overline{p}_{\mathrm{E}}$ ($\overline{p}_{\mathrm{D}}$) → individual accuracy on easy (difficult) tasks

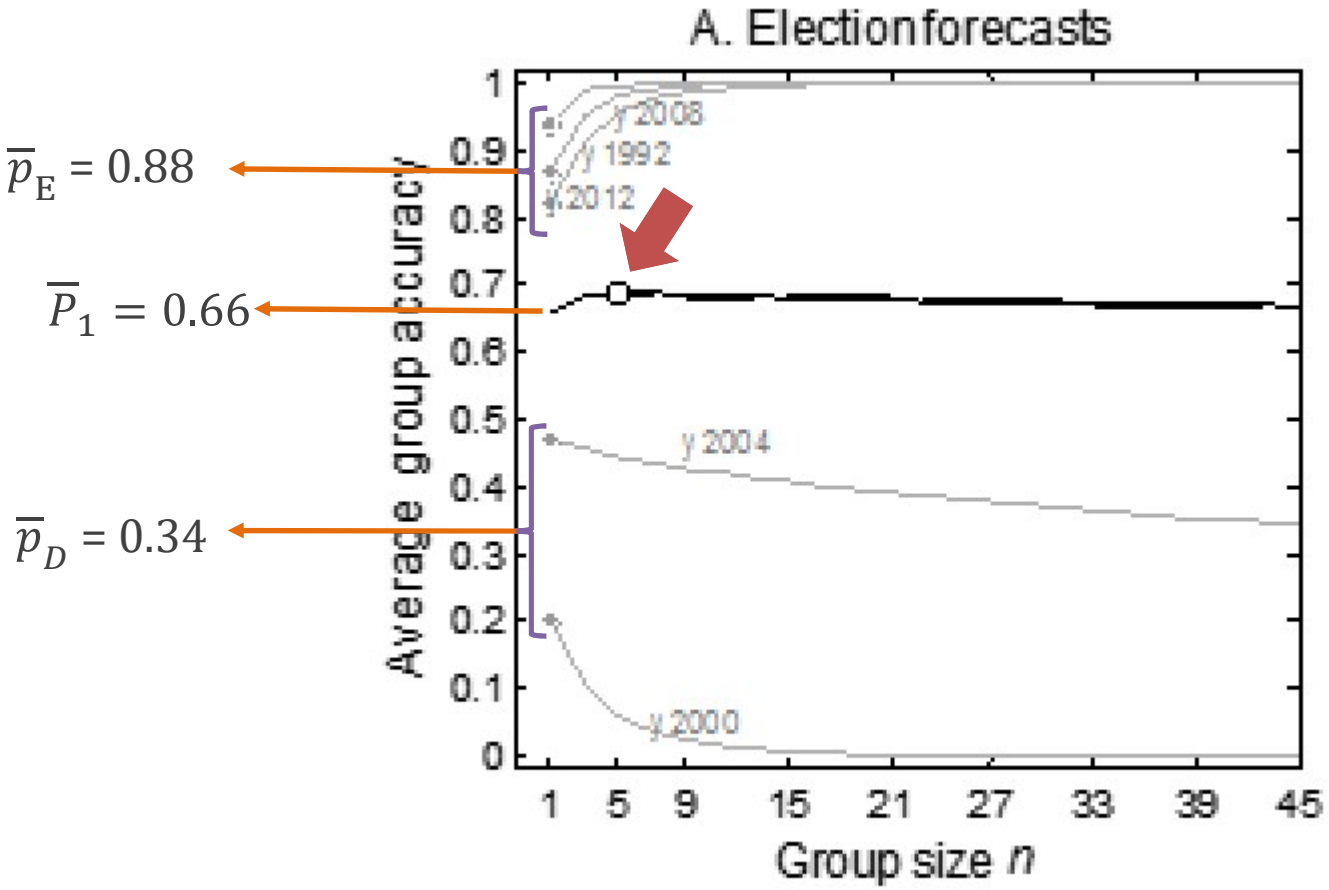
Moderate group sizes better whenever

friendly environment and $\overline{P}_1 > 0.5$

Is the real world friendly or unfriendly?

Real-world task environments: Elections

(Graefe, 2014)



→ $\overline{p}_E + \overline{p}_D > 1$ and $\overline{P}_1 > 0.5$

(friendly environment, and average expert more accurate than chance across tasks)

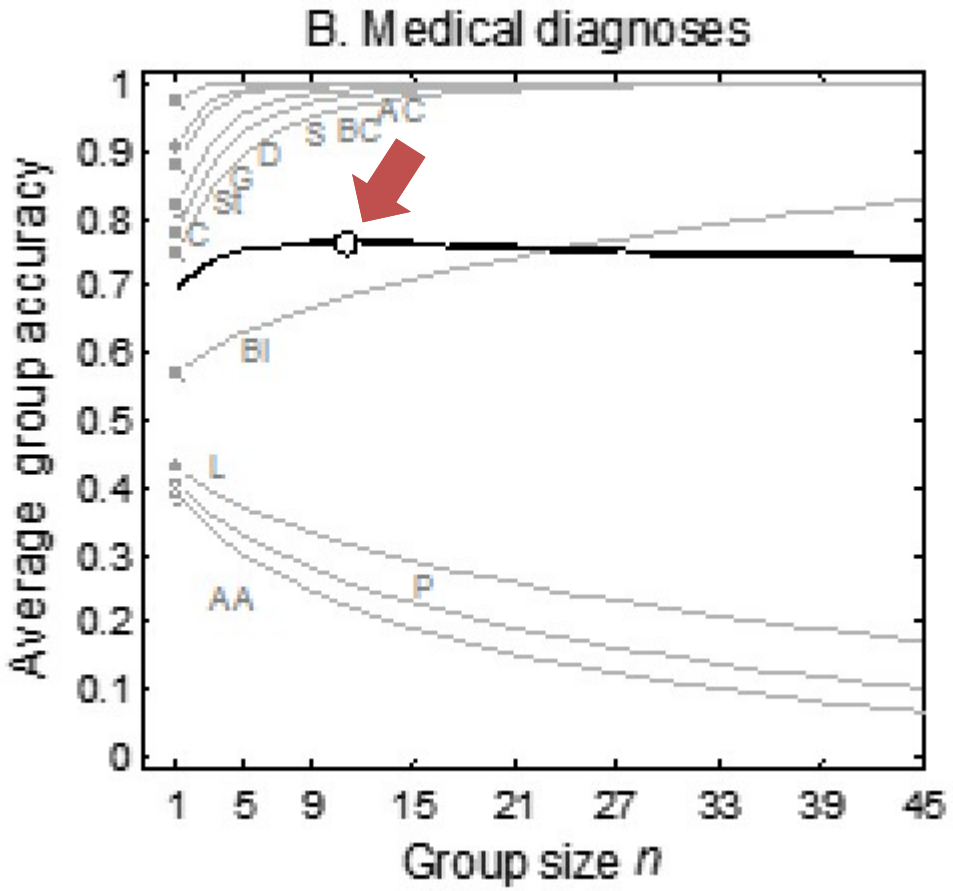
Real-world task environments: Medicine

(Schiff et al, 2009)

$\overline{p}_{\mathrm{E}} = 0.81$

$\overline{P}_1 = 0.70$

$\overline{p}_D = 0.41$



→ $\overline{p}_{\mathrm{E}} + \overline{p}_{\mathrm{D}} > 1$ and $\overline{P}_1 > 0.5$

(friendly environment, and average expert more accurate than chance across tasks)

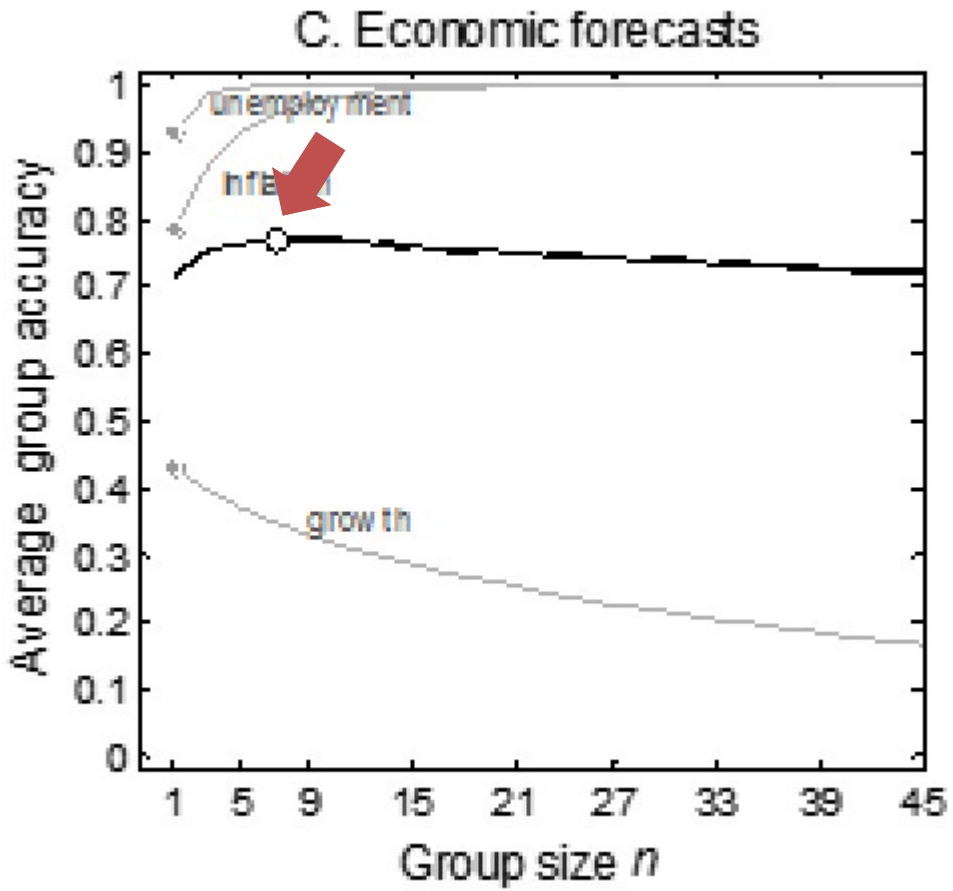
Real-world task environments: Economics

(Hilsenrath & Peterson, 2013)

$\overline{p}_{\mathrm{E}} = 0.86$

$\overline{P}_1 = 0.71$

$\overline{p}_D = 0.43$



→ $\overline{p}_{\mathrm{E}} + \overline{p}_{\mathrm{D}} > 1$ and $\overline{P}_1 > 0.5$

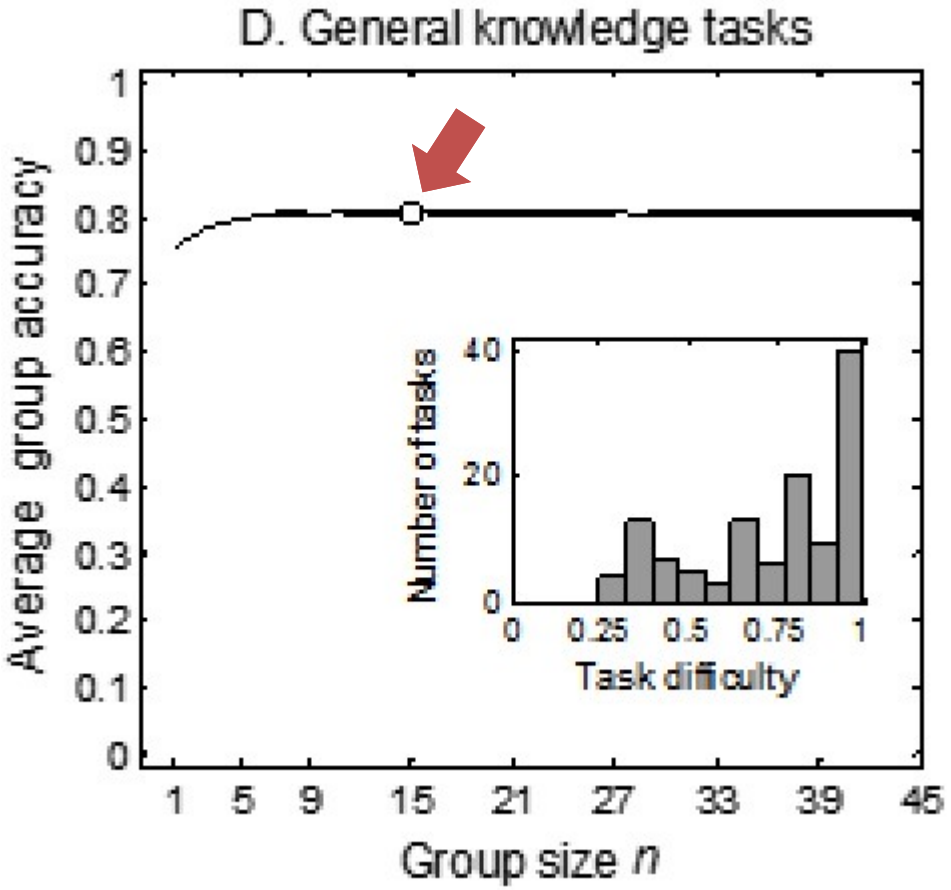
(friendly environment, and average expert more accurate than chance across tasks)

Real-world task environments: General knowledge (Juslin, 1997)

$\overline{p}_{\mathrm{E}} = 0.86$

$\overline{P}_1 = 0.76$

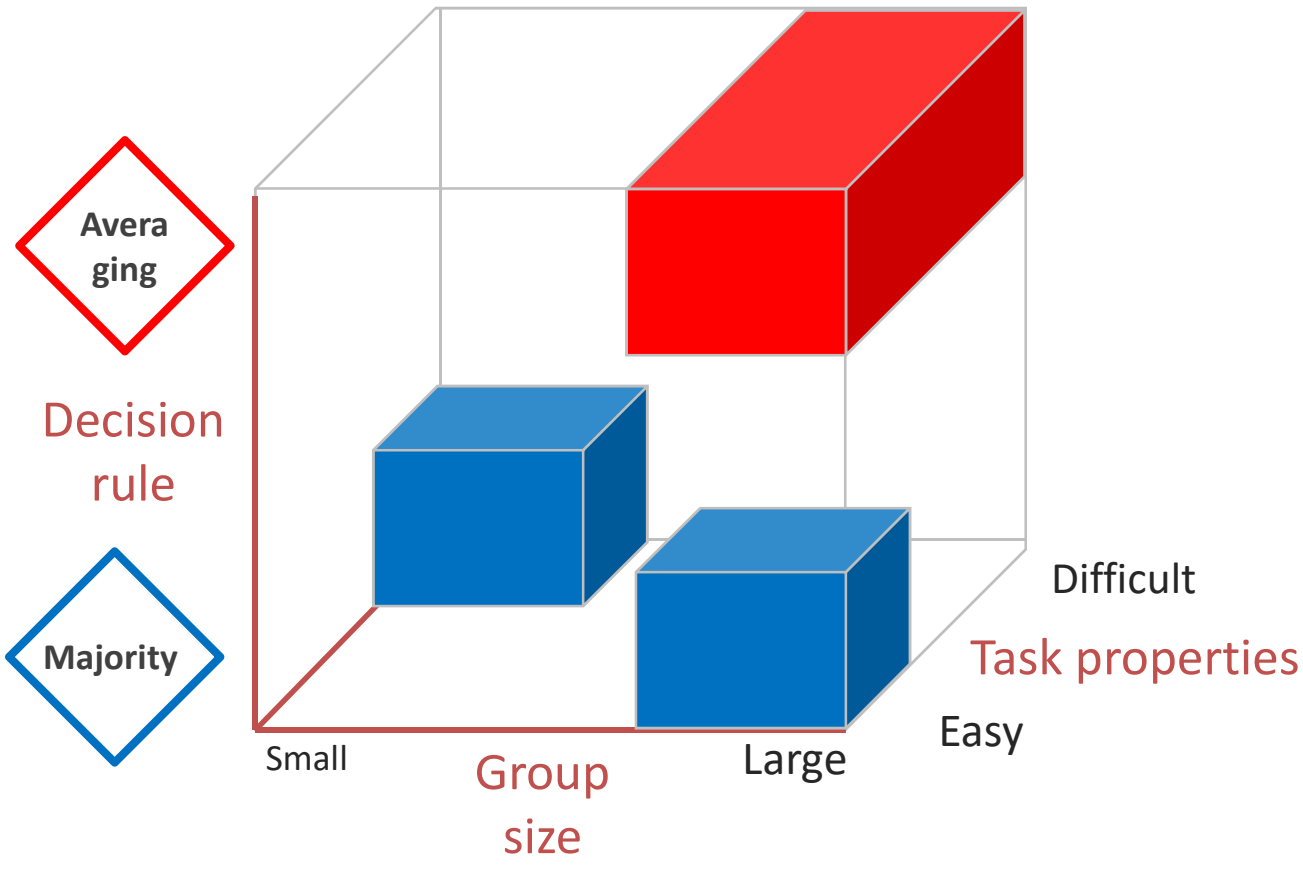
$\overline{p}_D = 0.38$



→ $\overline{p}_{\mathrm{E}} + \overline{p}_{\mathrm{D}} > 1$ and $\overline{P}_1 > 0.5$

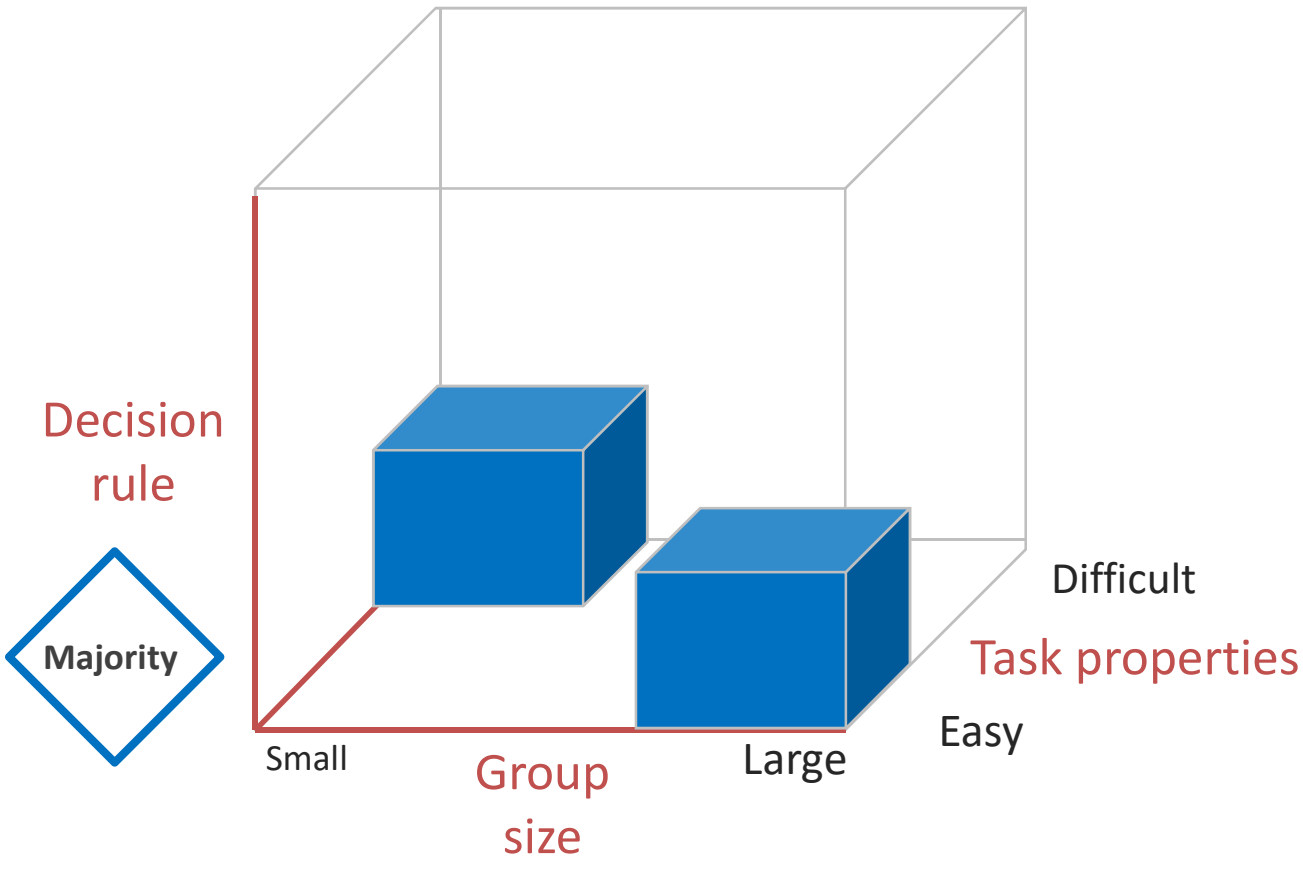
(friendly environment, and average expert more accurate than chance across tasks)

Best group size depends on decision strategy and task properties

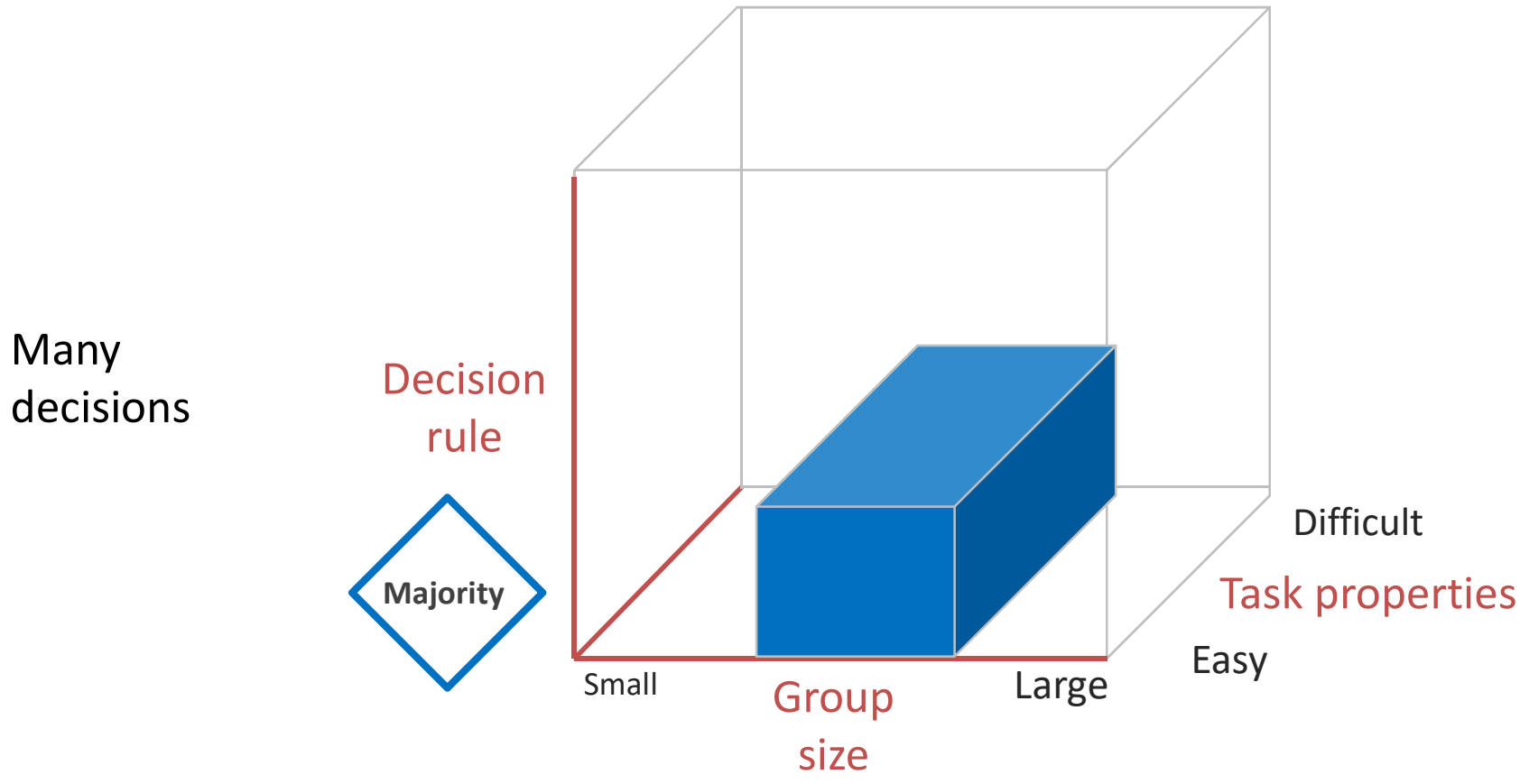


Best group size depends on decision strategy and task properties

Single decision



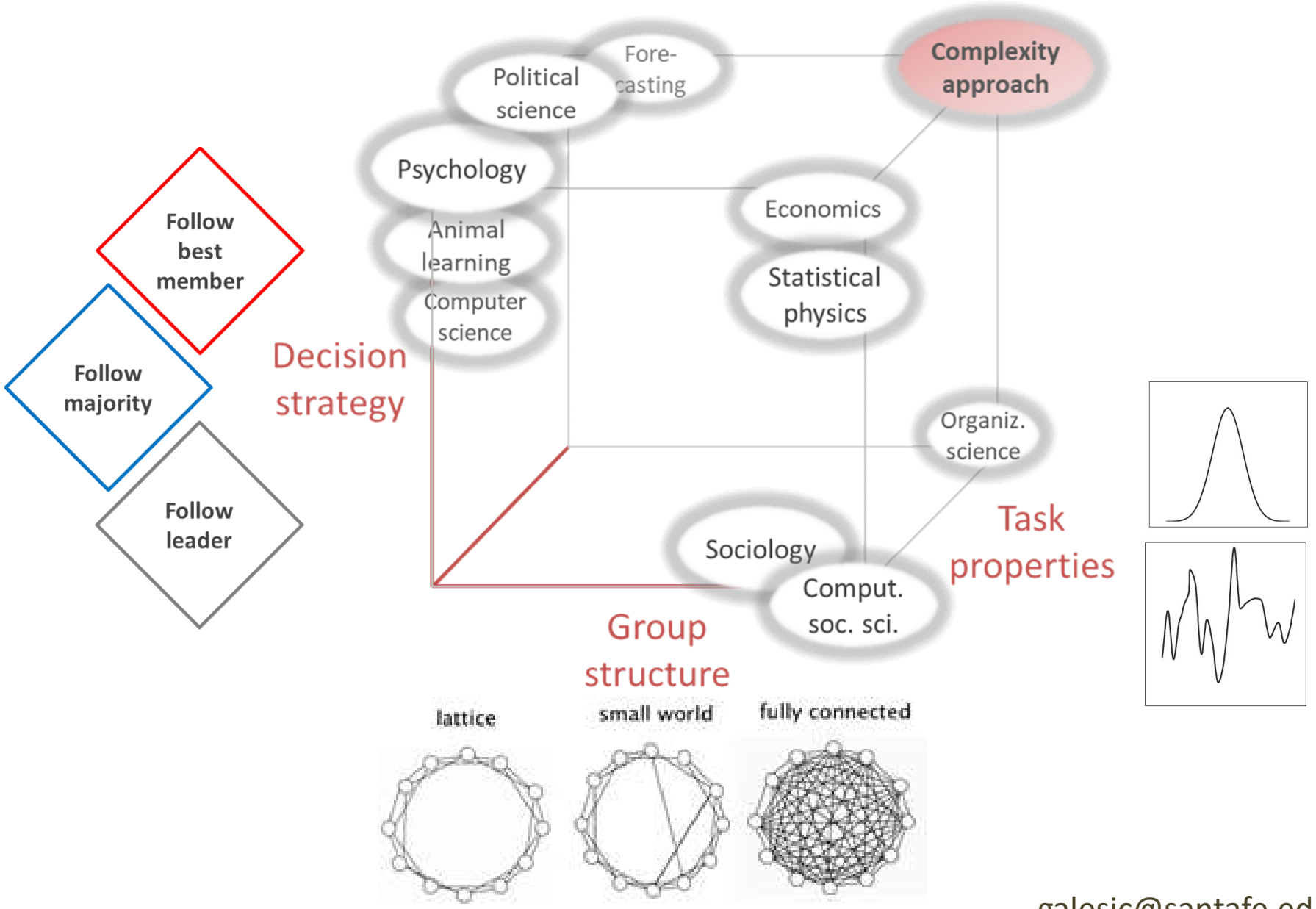
Best group size depends on decision strategy and task properties



Implications

- When designing committees, larger is not always better
- If there is a chance of surprising mistakes, 5-9 member-team can be as good as or better than larger teams

In sum: Collective problem solving as a complex system



galesic@santafe.edu