Reward, punishment and the evolution of cooperation

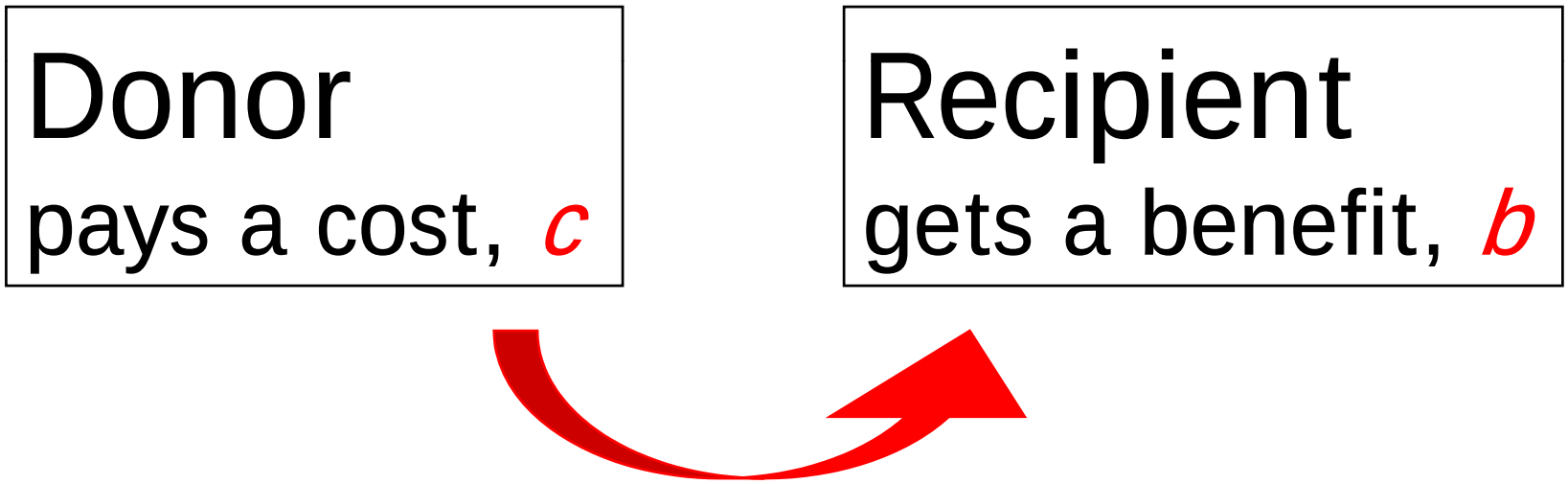
David G. Rand

Program for Evolutionary Dynamics
Psychology Department
Berkman Center for Internet & Society

Harvard University

www.DaveRand.org

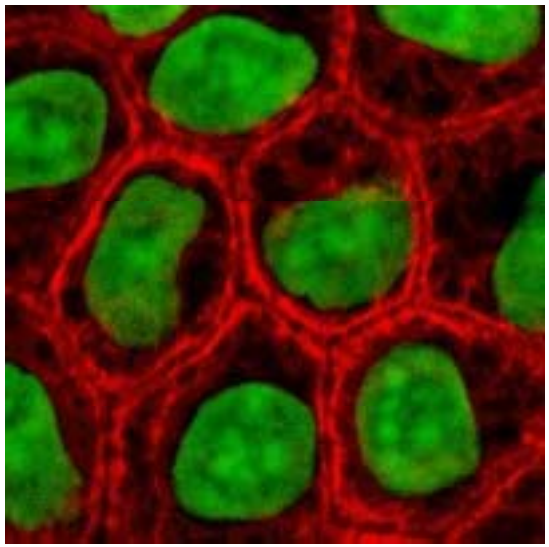
What is cooperation?



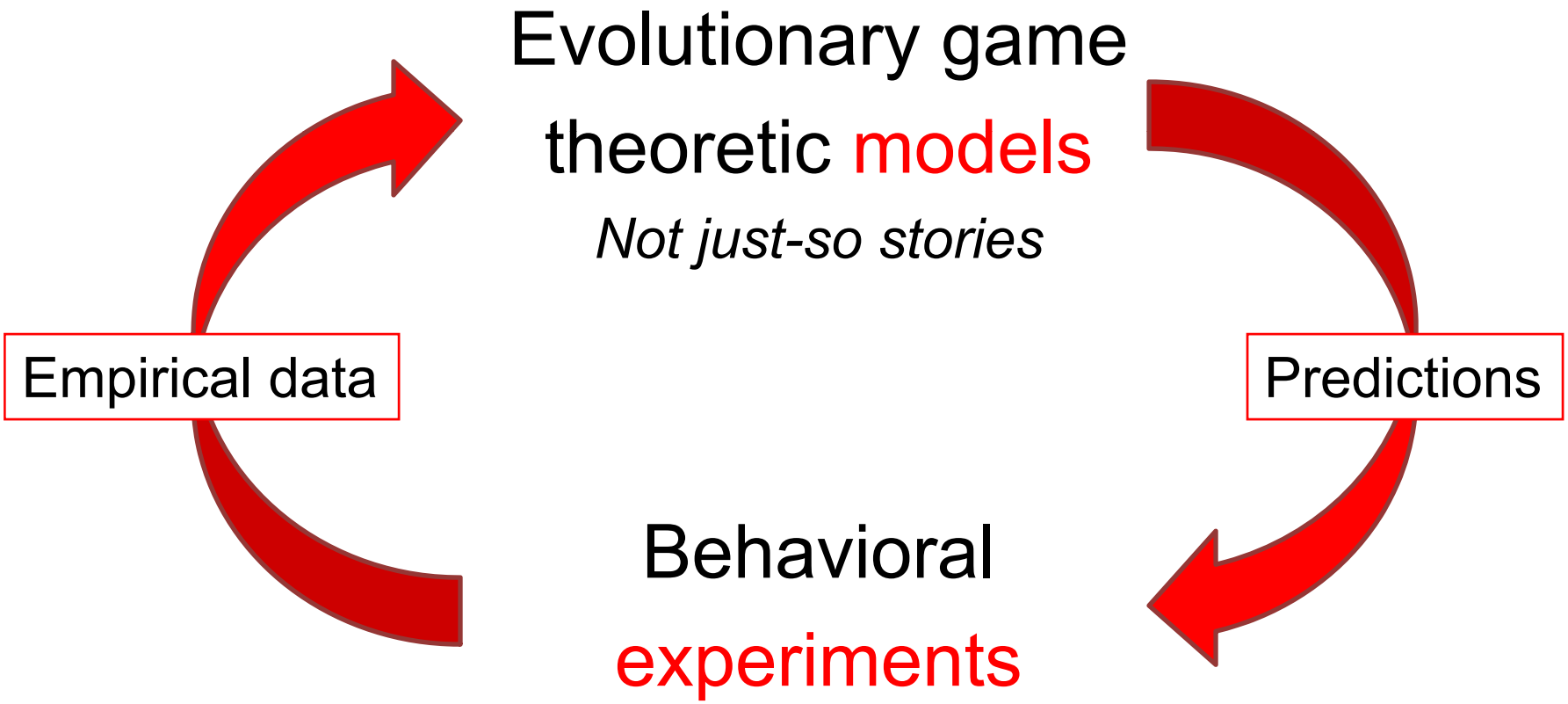
The puzzle of cooperation

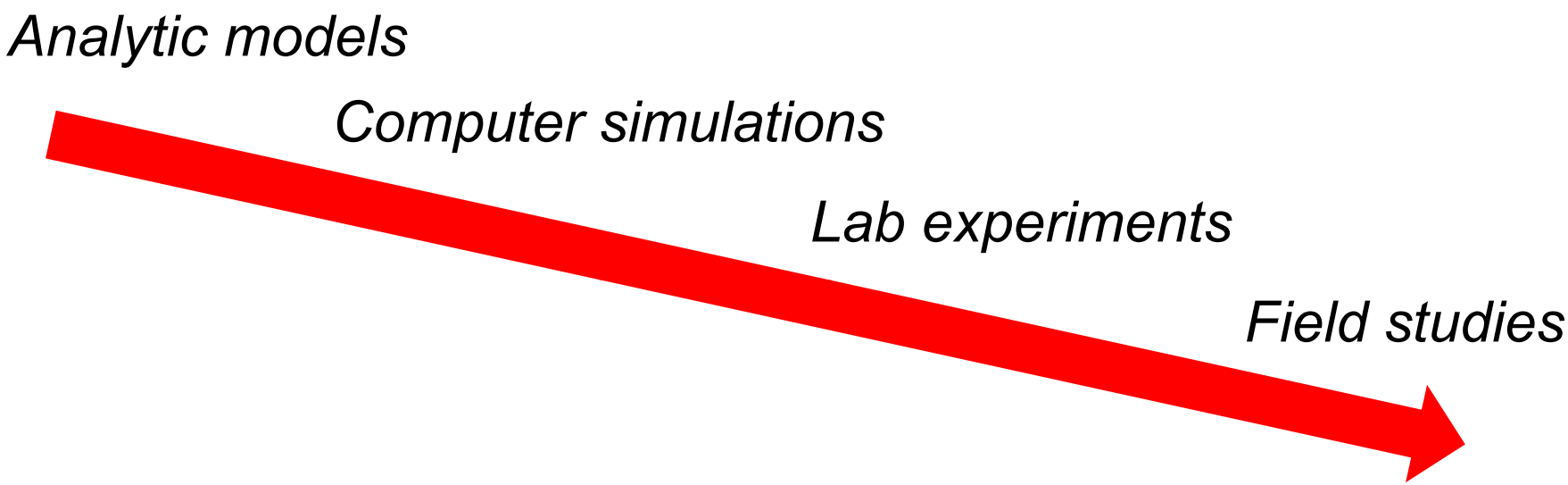
Evolution = Survival of the fittest

Yet cooperation is common



Integrated approach to the evolution of [human] cooperation





Behavior = snapshot of evolutionary process

Anonymous

Incentivized

Full information / no deception

Amazon Mechanical Turk

Online labor markets make experiments fast and cheap [and easy for theorists]

Short tasks (<5 min) for little money (<1$)

→ Performance-dependent payments

Quantitative replication of lab behavior

Horton Rand Zeckhauser, *Experimental Economics*, 2011 **[PDF]**

Evolutionary perspective

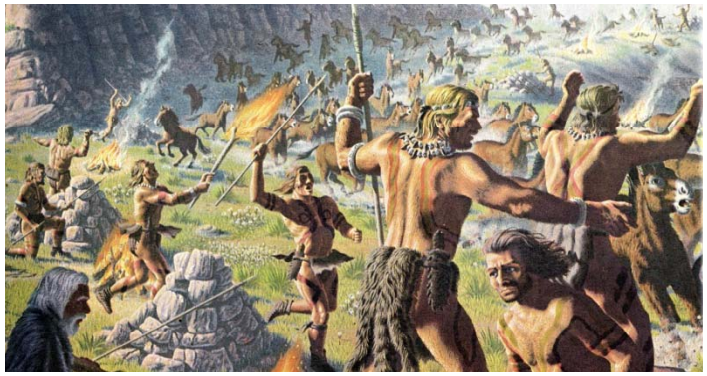
Human psychology did not develop in lab

→ Potential for mismatch

→ Interesting, but must be careful when interpreting

Explaining 'irrationality'?

Genetics vs cultural



Reciprocity

Repeated interactions

Conditional cooperation based on past actions

Direct reciprocity: Grim, TFT etc

Allows the evolution of cooperation

Costly punishment

TFT ‘punishes’ defection (**D**) with defection (**D**)

Costly punishment is a new proposal

P: you pay *A* to make other lose *B*

Costly punishment stabilizes cooperation in non-repeated games (1-shot or fixed length/name shuffled)

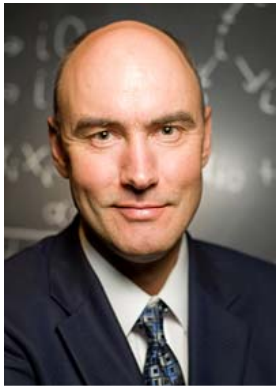
Yamagishi 1986, Ostrom et al 1994 ,
Fehr & Gachter 2000, 2002

Evolution of punishment

"Punishment promotes cooperation – therefore we evolved the taste for punishment"

But what about truly repeated games?

Is it a good idea to punish defection in a repeated Prisoner's Dilemma?



Martin Nowak

Winners don't punish

Anna Dreber

Repeated 2-player Prisoner's Dilemma

→ Control (C/D) vs treatment (C/D/P)

Pairs plays random # of PD rounds

→ 3/4 continuation probability

→ No info about partner's previous games

104 subjects, avg of 24 pairings, 79 PDs

Dreber Rand Fudenberg Nowak, *Nature,* 2008 **[PDF]**

Definitions

C: you pay 1, other gets 2

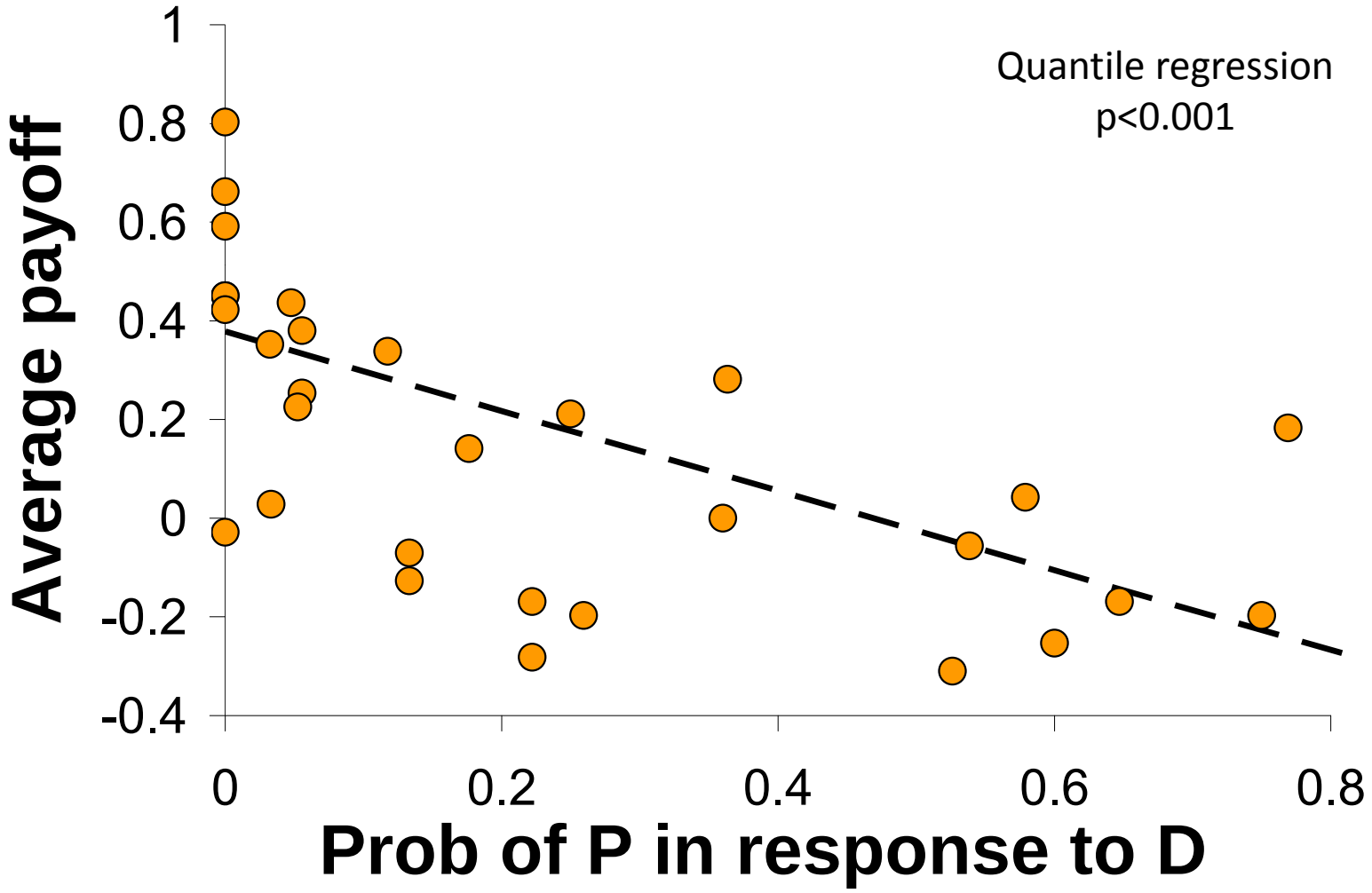
D: you gain 1, other loses 1

P: you lose 1, other loses 4

	C	D	P
C	1,1	-2,3	-5,1
D	3,-2	0	-3,-2
P	1,-5	-2,-3	-5,-5

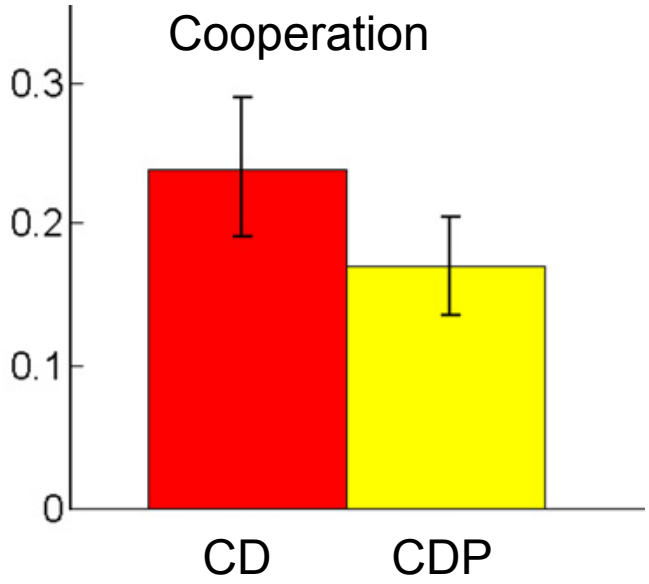
Punishers don’t win

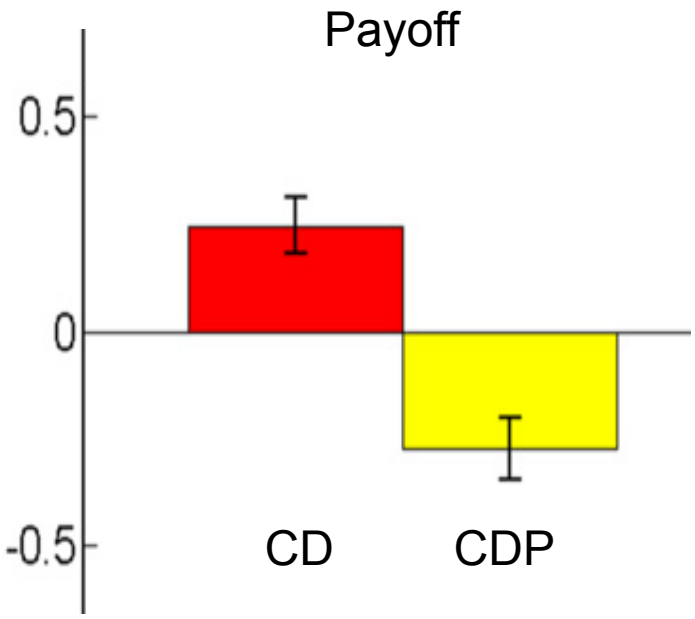
[And no group-level payoff benefit]



Cross-cultural replication

Study replicated in Beijing – punishment is even worse *Wu et al PNAS 2009*





Evolutionary dynamics

Experiments suggest punishment is maladaptive in repeated PD

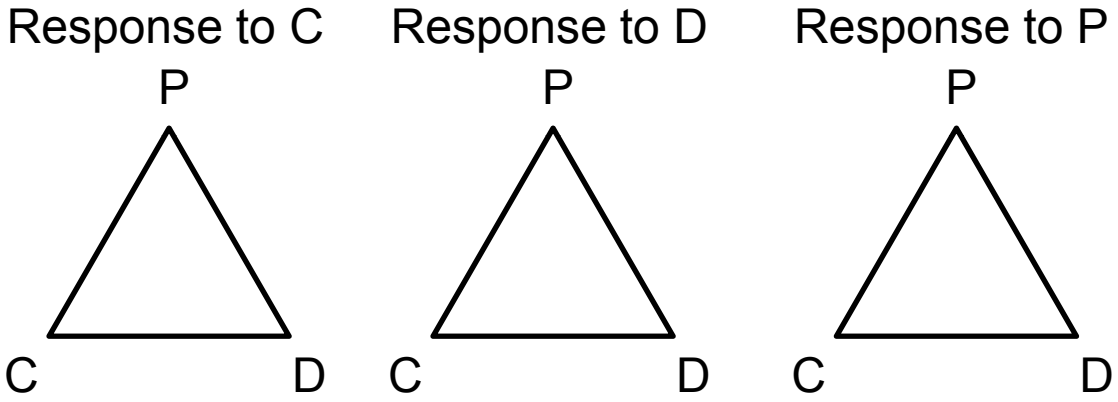
Use evolutionary dynamics to ask what strategies evolve

Does natural selection favor the use of costly punishment after the opponent defects?

Direct reciprocity with costly punishment

Hisashi Ohtsuki

Reactive first order strategies



	C	D	P
Repsonse to C	p1	q1	1-p1-q1
Response to D	p2	q2	1-p2-q2
Response to P	p3	q3	1-p3-q3

Rand Ohtsuki Nowak, *J Theoretical Biology*, 2009 **[PDF]**

Equilibrium analysis

Cooperative punishers can be Nash eq.

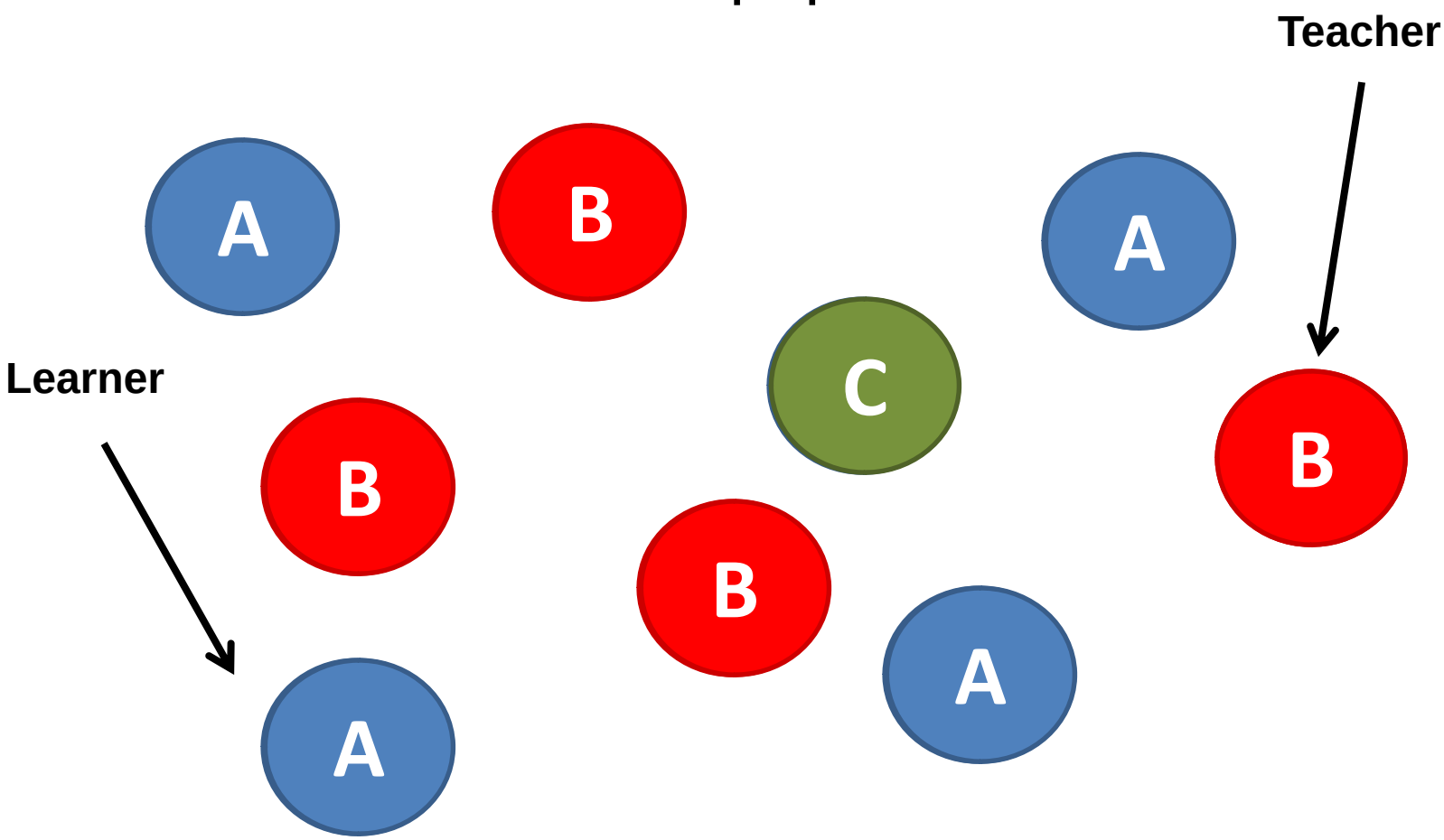
If cost of P ≤ cost of C

For experimental params, there are cooperative punisher Nash strategies

In response to P, cooperative Nash always play C

Stochastic game dynamics

Pairwise comparison process
Well-mixed population



Decision rule

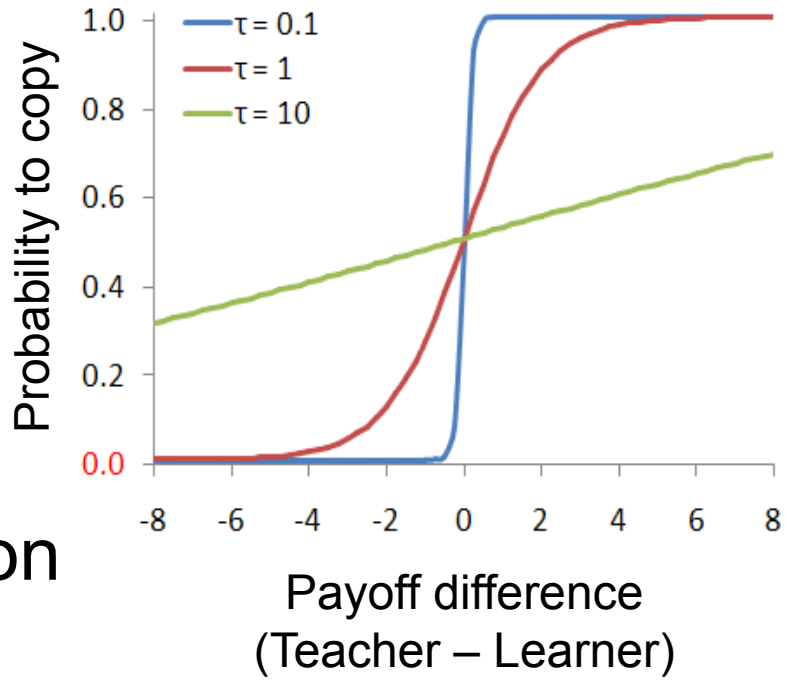
Learner copies Teacher with probability P

$$P = \frac{1}{1 + e^{-(\pi(T) - \pi(L))/\tau}}$$

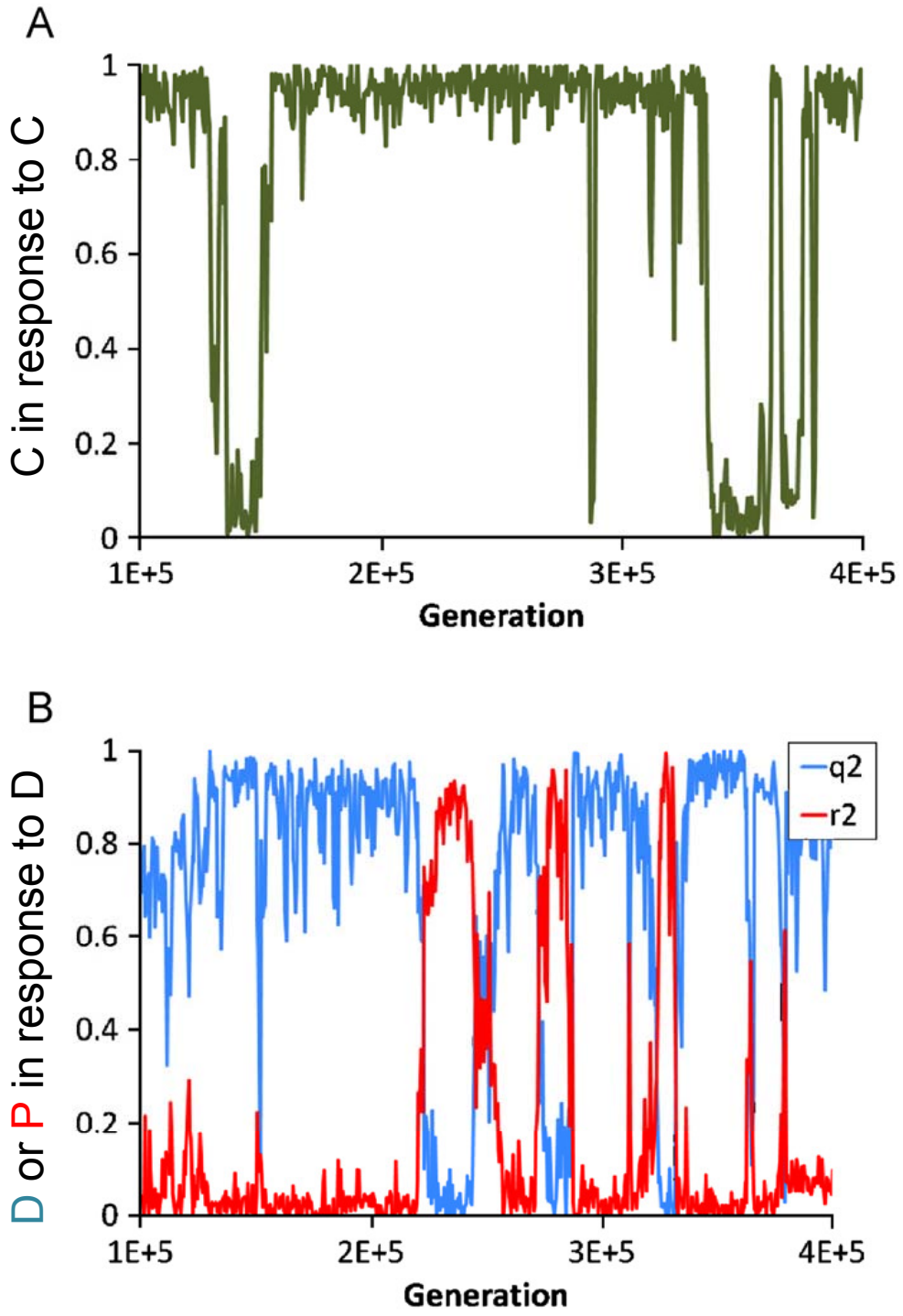
π(T) = Teacher payoff

π(L) = Learner payoff

τ = temperature of selection (noise in learning)



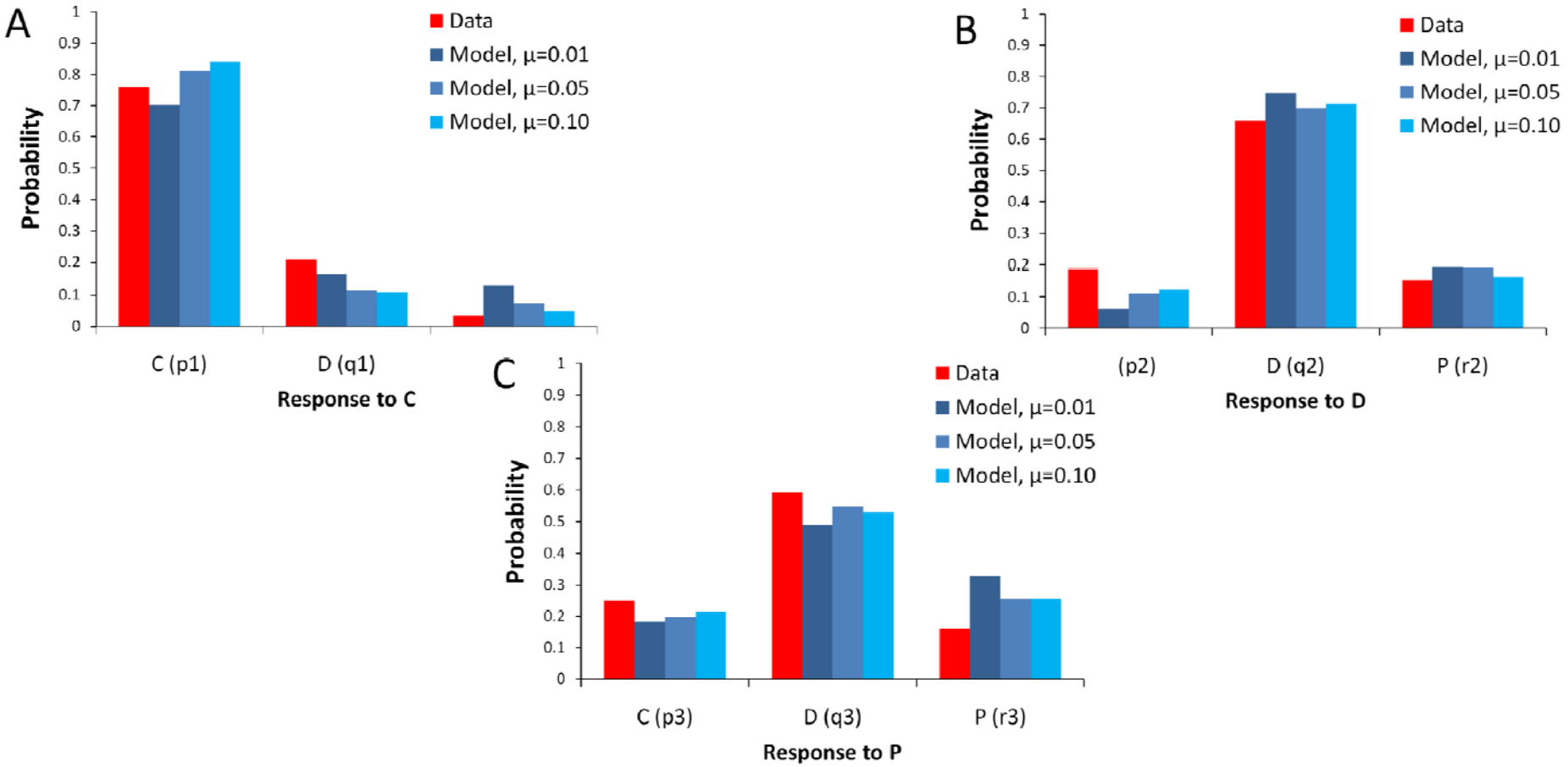
Or with probability *u,* mutation occurs


A
C in response to C
1
0.8
0.6
0.4
0.2
0
1E+5
2E+5
3E+5
4E+5
Generation
B
D or P in response to D
q2
r2
1
0.8
0.6
0.4
0.2
0
1E+5
2E+5
3E+5
4E+5
Generation

Fit simulation model to data (1 free param $\tau = 0.8$)

→ Nash predicts only C in response to P

Punishment disfavored over wide parameter range



Why punishment loses

Nash calc: never actually use P

→ P is OK

Evolutionary model: punish mutants/poor learners

→ P is costly

Same for Ultimatum, Centipede games

Evolution vs. rational choice

Nash equilibrium analysis not in agreement with behavioral data

Evolutionary model reproduces experimental behavior

Stochastic evolutionary dynamics may underlie development of strategies

From individuals to groups

Choose how much to contribute to a common pool

All contributions are multiplied by a factor and split evenly by everyone (regardless of contribution)

Cooperation breaks down in the lab

Punishment & public goods

Costly punishment stabilizes contribution

Yamagishi 1986, Ostrom et al 1994 , Fehr & Gachter 2000, 2002

Punishment better than [denial of] reward

Sutter et al 2006, Sefton et al 2007

Previous studies focused on end-game effects

Let's look at truly repeated games

Positive interactions promote public cooperation

Repeated 4-player public goods game, 192 subjects

20 unit endowment, 1.6x contribution multiplier

Four treatments. PGG followed by

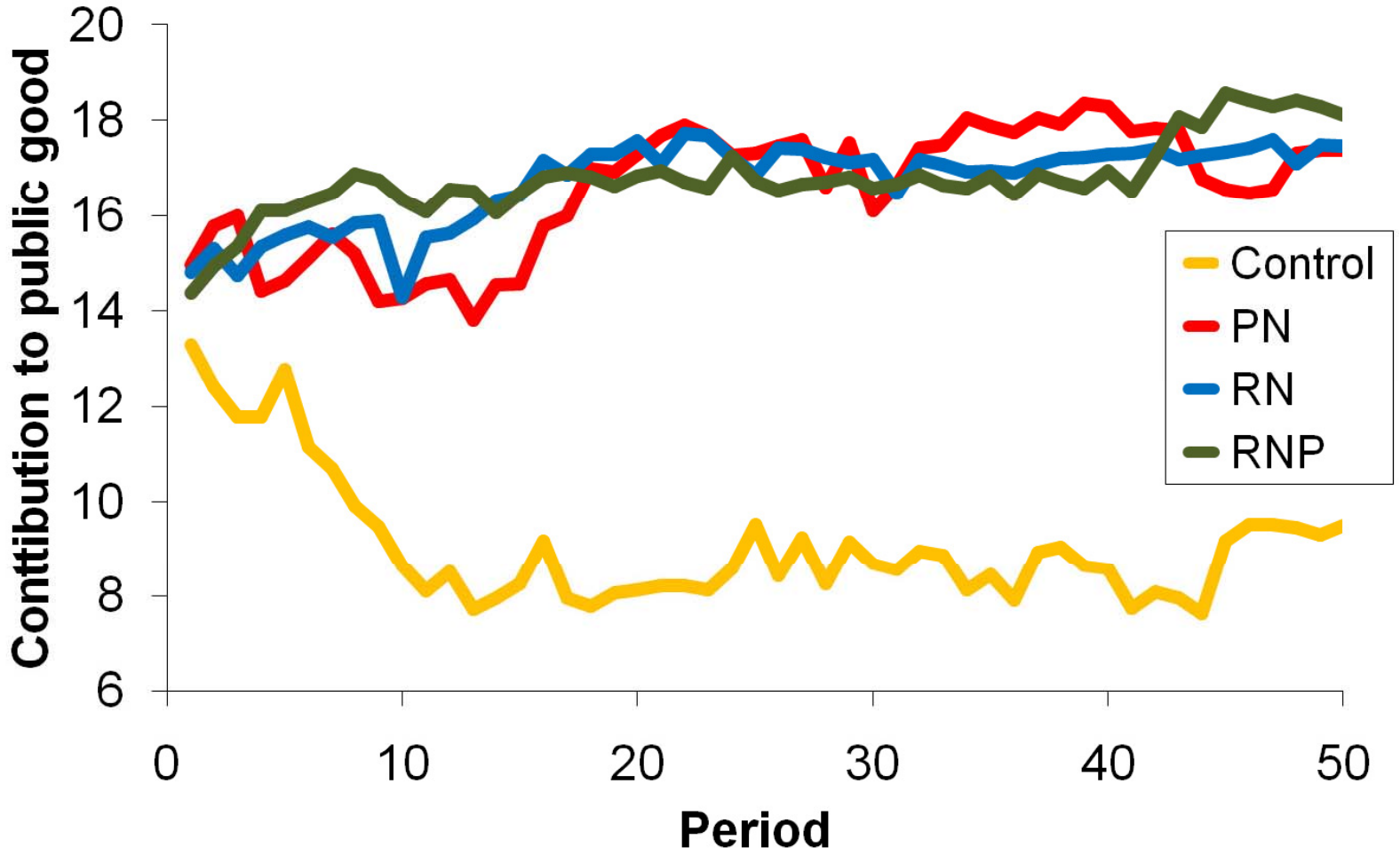
- Control: Nothing.
- PN: Punish (-4 for you, -12 for other) or no action
- RN: Reward (-4 for you, +12 for other) or no action
- RNP: Reward, no action, or punish

Game length unknown to participants (50 rounds)

Rand Dreber Ellingsen Fudenberg Nowak, *Science,* 2009 **[PDF]**

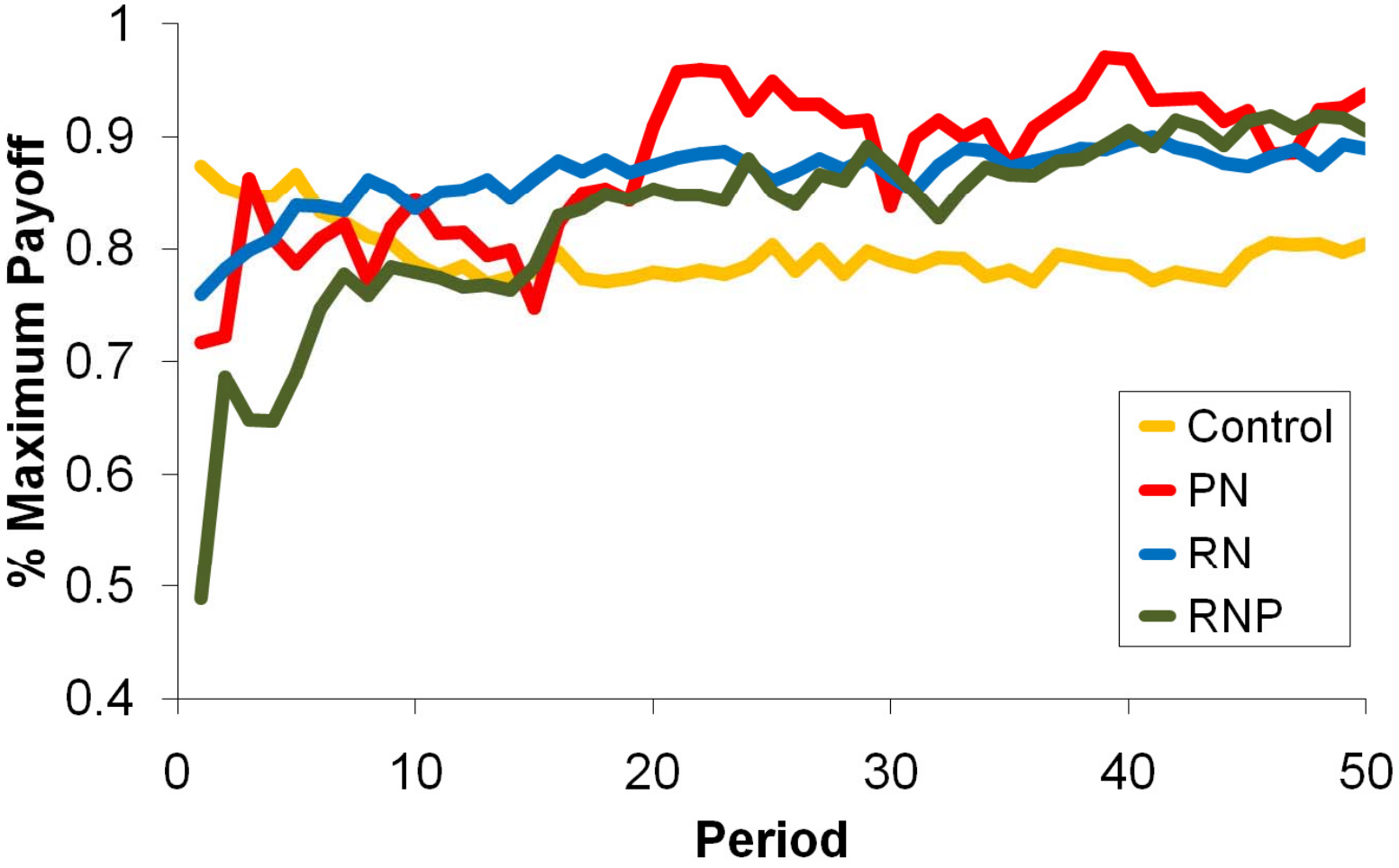
Effect on contributions:

All 3 treatments equally effective



Effect on % of max possible payoff

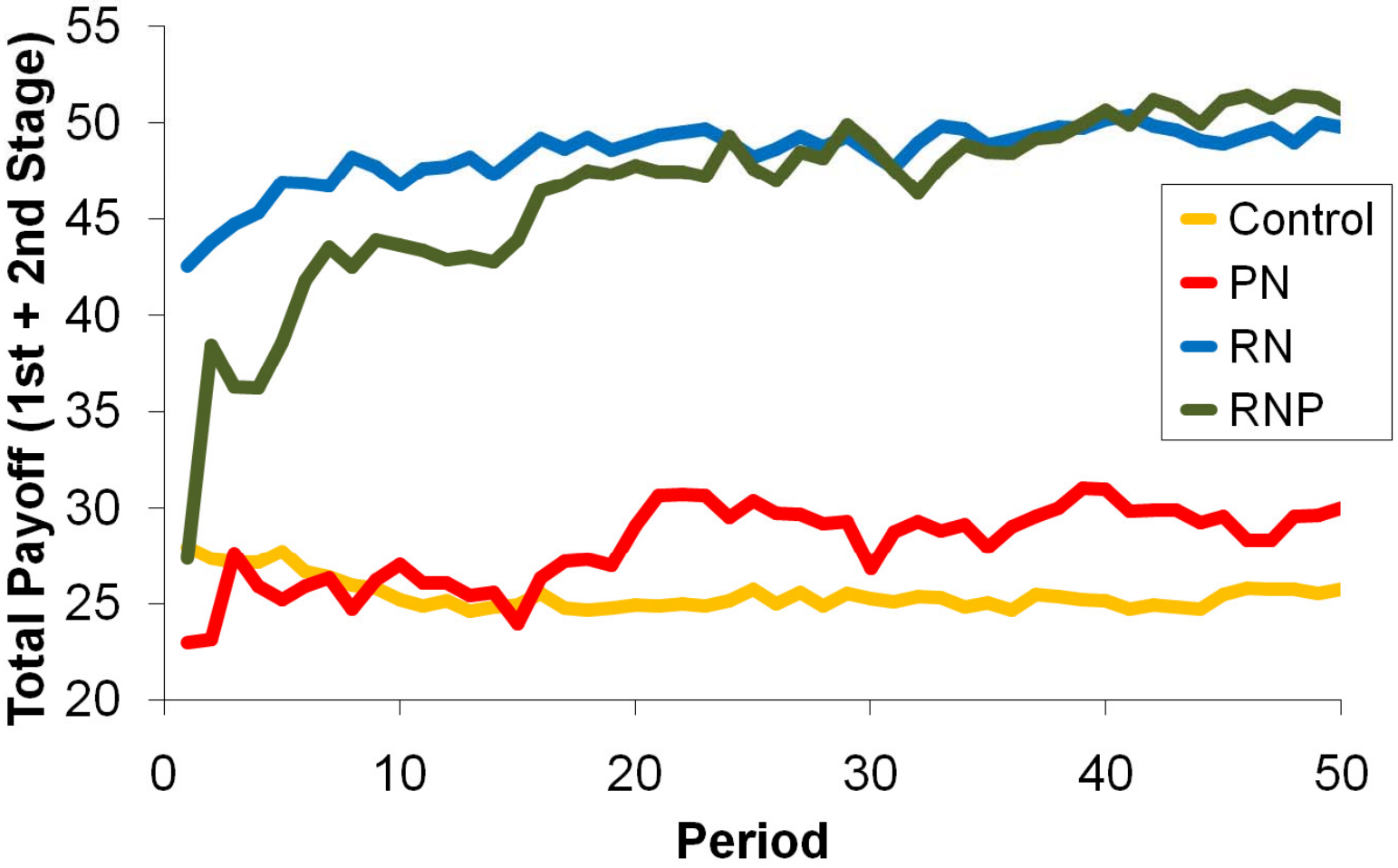
All 3 treatments equally effective



Effect on total payoff:

Reward out-performs punishment

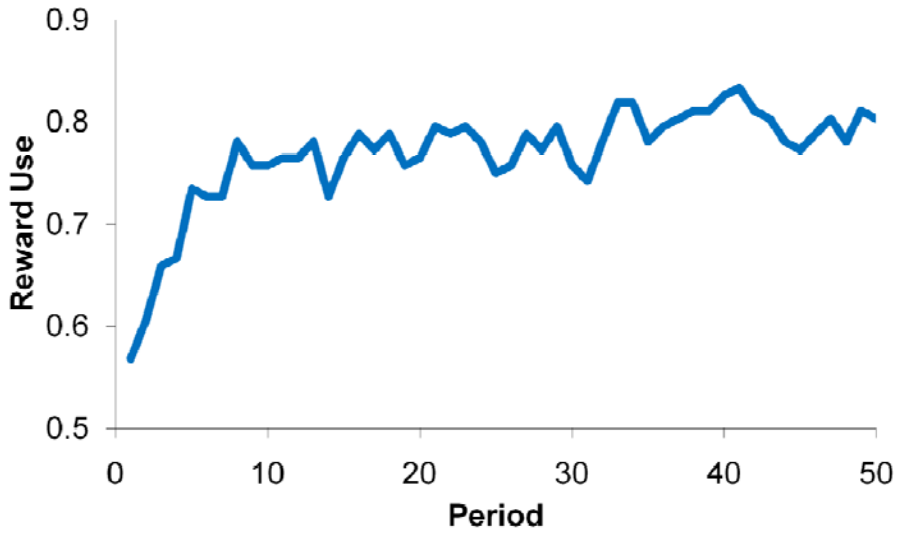
Equal % possible payoff → better actual payoff



Reward use is stable

Previous experiments: reward use decays

But persistent identities & shadow of future maintain rewarding



Availability of rewards

Life Is full of chances to help each other (or not) in a non-zero way (PD)

Public life and private life are coupled

Denying future rewards (like TFT) is a non-destructive way to “punish”

Direct vs indirect reciprocity

Similar results for games with reputation

→ Experimental *Milinski et al., Science, 2002; Ule et al., Science 2009*

→ Theoretical *Ohtsuki et al., Nature, 2009*

Cross-cultural differences

Benedikt Herrmann

EU/ERC

PGG + reward or punishment in Romania

Control same as US

No effect of RN or PN on contributions

→ Dramatic decrease in payoffs!

Why?

→ PN: Anti-social punishment

→ RN: Pairwise rewarding, ignoring group

Anti-social punishment

Inconsistent with rational self-interest

Contrary to most standard preference models

In opposition to concept of ‘altruistic’ punishment

Evolution of anti-social punishment

Excluded *a priori* from previous models

Does punishment promote cooperation when anti-social punishment is allowed?

Can we explain the evolution of anti-social punishment?

Local interaction & competition

‘Viability updating’

Mixing population with limited interaction

→ defectors that always punish

Lattice

→ defectors that punish cooperators
(bad to punish own offspring)

Rand Armao Nakamaru Ohtsuki, *JTB,* 2010 **[PDF]**

Summary

People have a taste for punishment

But not necessarily because of cooperation

Punishment can ‘self-interested’

Antisocial punishment and retaliation are dangerous

Society is best built on positive interactions

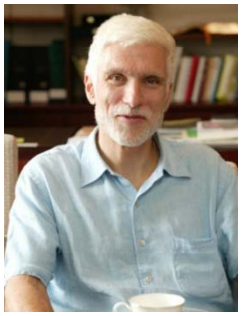
Collaborators

Anna Dreber

SIFR

Tore Ellingsen

Stockholm School of Economics



Drew Fudenberg

Harvard Economics

Benedikt Herrmann

EU/ERC

Hisashi Ohtsuki

Tokyo Inst.Tech.

Martin Nowak

Harvard Math & Biology

Corina Tarnita

Harvard Society of Fellows

1. Dreber A, Rand DG, Fudenberg D, Nowak MA (2008) Winners don't punish. *Nature* **452**, 348-351.
2. Rand DG, Dreber A, Ellingsen T, Fudenberg D, Nowak MA (2009) Positive interactions promote public cooperation. *Science* **325**, 1272-1275.
3. Rand DG, Ohtsuki H, Nowak MA (2009) Direct reciprocity and costly punishment: Generous tit-for-tat prevails. *J Theor Biol* **256**, 45-57.
4. Rand DG, Armao J, Nakamaru M, Ohtsuki H (2010) Anti-social punishment can prevent the co-evolution of punishment and cooperation. *J Theor Biol* **265**, 624-632

www.DaveRand.org