Social/Prisoner's dilemma

Payoff matrix:

$$\mathbf{A} = \begin{array}{c} \\ D \\ C \end{array}\begin{array}{c} \begin{array}{cc} D & C \end{array} \\ \begin{pmatrix} (P,P) & (T,S) \\ (S,T) & (R,R) \end{pmatrix} \end{array},$$

two strategies: C and D

Rank order of payoffs: $T > R > P > S$; NE: DD

For repeated games we also assume that $2R>T+S$

P: **P**unishment for mutual defection

R: **R**eward for cooperation

T: **T**emptation to choose defection

S: **S**ucker's payoff

Payoff matrix:

P: **P**unishment for mutual defection

R: **R**eward for cooperation

T: **T**emptation to choose defection

S: **S**ucker's payoff

$$\mathbf{A} = \begin{matrix} & D & C \\ D & (P,P) & (T,S) \\ C & (S,T) & (R,R) \end{matrix},$$

two strategies: C and D

Rank order of payoffs: $T > R > P > S$; NE: DD

For repeated games we also assume that $2R>T+S$

Real life situations:

- disarmament negotions
- choice between fair and unfair behaviour
- to do your job correctly and thoroughly or not
- to learn all skills required to do my job perfectly or not
- to protect the environment or not
- to bring alcohol to a party or not
- to get vaccinated or not
- to use the turn signal or not
- to help tourists or foreigners or not

The curious nature of the **Prisoner's dilemma** was discovered by Merrill Flood (1950).

The story and the name come from Albert W. Tucker.

Many questions: Are we rational? Are we selfish?

The dilemma contradicts the 'invisible hand' argument of Adam Smith…

Can we apply game theory to describe economic, human behaviour, etc? ...

Experiments (with Melvin Dresher at RAND Corporation):

The price of used cars

Price = (buying price + selling price)/2

Asymmetric PD game played 100 times in succession

the players kept a log of their comments

General belief (**Folk Theorem**):

Repetition introduces further possibilities that resolve the dilemma.

Computer tournament conducted by Robert Axelrod (1979-1981)

N players play a repeated PD game against each other (round robin)

After M rounds the winner is the one who has the highest average payoff ($M \to \infty$)

Payoff matrix:

$$\mathbf{A} = \begin{matrix} & D & C \\ D & (1,1) & (5,0) \\ C & (0,5) & (3,3) \end{matrix}$$

The players know all previous decisions.

Goal: to develop an algorithm that achieves the highest payoff (against its opponents).

Axelrod nominated the random strategy: $\mathbf{s}^T = (0.5, 0.5)$, and declared that each player will also play a game against itself.

The average payoff was determined by averaging the results between rounds 200 and 300, while the repeated game was actually longer (N=600 rounds). This way the undesired effects of the **shadow of the future,** which dictates that players should choose D in the last round, which reduces the number of meaningful iterations to (N–1) making D the best choice in step (N–1) too, etc. In conclusion, rational players should all choose D in a finite game, and thus remain stuck in the single-shot game's dilemma.

Possible strategies for the repeated game:

AllD: always choose *D* (unconditional defector, bad guy, …)

AllC: always choose *C* (good guy or sucker)

Random: choose *D* or *C* with probabilities *q* and (1–*q*)

TFT (Tit-for-tat): choose *C* first, then respond in kind to the coplayer's last move

Suspicious TFT: first choose *D*, then respond in kind

Generous TFT: TFT, but choose *C* with probability *q* instead of retaliating with *D*

WSLS (win-stay-lose-shift): first choose *C* or *D*, then switch if payoff is smaller than a predetermined aspiration level (when $U_x < a$)

Stochastic reactive strategies: Choose *C* or *D* with probabilities that depend on the previous decision of the coplayer

Stochastic reactive strategies with longer memory

Etc.

Home assignment 3.1 How does the repeated PD game play out for the following strategy pairs: AllD–AllC, AllD–TFT, TFT–WSLS?

Axelrod's tournament: 14 strategies were nominated

The winner: TFT (the simplest/shortest algorithm)

nominated by Anatol Rapoport

Second tournament: with 50 new strategies

Most of the new strategies would have won the first tournament.

Nonetheless, the winner was again TFT.

Axelrod's advice:

Don't be envious!

Don't be the first to defect!

Reciprocate both cooperation and defection!

Don't be too clever!

The moral: PD games should be transformed into repeated PD games with an uncertain endpoint and use TFT.

Axelrod systematically investigated the competition between the nominated strategies.

Conclusions:

Knowing the strategies of the opponents, we can develop a strategy that beats TFT, otherwise we should follow TFT.

TFT plays a key role in the maintenance of cooperation if the players are allowed to adopt the strategy of a better performing player

Hamilton suggested the following:

After each round, make the worst-performing player adopt the winning strategy.

This a simple realization of Darwinian selection.

The winner of this evolutionary competition was also TFT.

In the absence of TFT, AllD became the winner.

Shortcomings:

- The strategy set was not complete.

Consequently, the validity of the conclusions is questionable.

-TFT fails under noisy conditions.

What happens between TFT players if mistakes can happen with a low probability?

A possible series of choices:

error error error error

↓ ↓ ↓ ↓

C C C C C C D C D C D C D C D C D D D D D D D D D C D C D C D C C C C C C

C C C C C C C D C D C D C D C D C D D D D D D D D D C D C D C D C C C C C

The mistakes significantly decrease the average payoff.

- If $2R<T+S$, then it is even better to pick *C–D* and *D–C* pairs alternately.

25th anniversary of the Axelrod tournament

The computer tournament was repeated.

TFT was beaten by a strategy that received support from a team.

During an initial period the team members identified each other and later they chose *C* against the „leader" who exploited them by choosing *D* against the team members. Otherwise, the „leader" used TFT against all other players.

The average payoff of the team was lower than TFT's.

Real example: bicycle racing, long-distance running, ...

The advantage of a team can be compensated by the participation of other teams.

Home assignment 3.2: Show that WSLS can win in a population consisting of WSLS, AllC and TFT players! What happens if AllD players are present, too?

Experiments

Games provide opportunities to investigate animal and human behaviour.

First experiment by Merril Flood and Melvin Dresher (~1950 at RAND Corporation)

details in Poundstone: Prisoner's Dilemma (Anchor Books, 1992)

Rapoport and Chammah: Prisoner's Dilemma (1965, Univ. of Michigan Press)

1) Question to new colleague who bought a car from a leaving colleague:

How have you shared the profit of the second-hand car dealer?

(this is basically the „Ultimatum game")

2) An iterated two-player 2x2 game

the players' reactions were recorded (see Poundstone's book)

mutual cooperation with a frequency of 75 %

the game was complicated (asymmetries in payoffs)

3) A series of experimental studies was conducted at Ohio State University

(sponsored by the Air Force)

Animal experiments,

Milinski [*Nature* 325 (1987) 433]

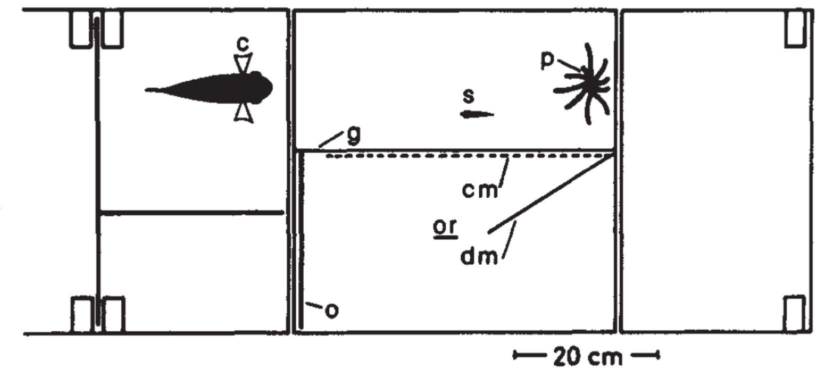


Fig. 1 Plan of the experimental equipment. The glass tanks shown contain the cichlid (c) and the plant (p), respectively. A typical position of a test stickleback (s) is also shown, as are the positions of the cooperating mirror (cm) and the defecting mirror (dm); 0, opaque screen; g, glass partition. Only one of the two mirrors was present in each experiment.

Sticklebacks occasionally check whether their predator is hungry.

If it's not, then they can safely continue eating.

In the experiment the predator was in a separate tank.

The coplayer's behaviour was imitated by a mirror.

Conclusions: The prey fish apply a strategy reminiscent of „Tit-for-tat", animals learn.

These experiments were later repeated in a more realistic environment using groups of fish.

Conclusions: Sticklebacks recognize each other and form permanent partnerships for inspecting their predator.

Examples of cooperation among animals

Possibilities for collaboration:

Feeding each other
(bacteria, vampire bats, etc.)
Warming each other
(sheep, etc.)
Guarding each other
(fish school, herding or flocking, etc.)
Hunting
(pack of wolves or lions)

Raising offspring

Specialization, division of work
(ants, bees, multicellular biological systems)

Symbiotism among plants and animals

Musk oxes against wolves

Ants form a bridge

Monkeys

Ultimatum game [Güth et al. *J. Econ. Behav. Organ.* 3 (1982) 367]

Two players are to share €100. The first (proposer) suggests a division the second (acceptor) either accepts or not. If the second player accepts the offer, then the sum is divided as it was proposed, otherwise they both receive nothing.

Game theory's advice: offer just €1 to the second player, who should accept it, as it is still more than nothing.

Experimental results:

- A large percentage of people (~ 50%) suggest fraternal division.

- The average offer is ~30–40%

- Low (<10%) and high (>60%) offers were rare.

- If the proposed offer amounted to less than 20%,

then 80% of the second players rejected the offer (punished the proposer).

This experiment was repeated worldwide (with people with different customs and cultural backgrounds, and from different walks of life) with the same results.

Explanation: It is the result of a long evolutionary process that developed over a long time.

Dictator game (“similar” to the Ultimatum game)

A sum (€100) is simply divided according to the proposal of the first (dictator) player. It is not a real game as the second player cannot affect its outcome.

Results: the average offered portion is about 20%.

sometimes the proposer suggested fraternal sharing.

sometimes the proposer kept the whole sum.

Trust game [Berg et al. *Games Econ. Behav.* 10 (1995) 122]

The first player decides what portion of $10 to invest.
The triple of the invested amount is given to the second player,
who decides what portion of the profit to pay back to the first player.

Results: the average investment is ~$5 (50%).

5 of 32 players invested the whole sum

while half of the second players rewarded nothing or very little

the average investment exceeded the average payback.

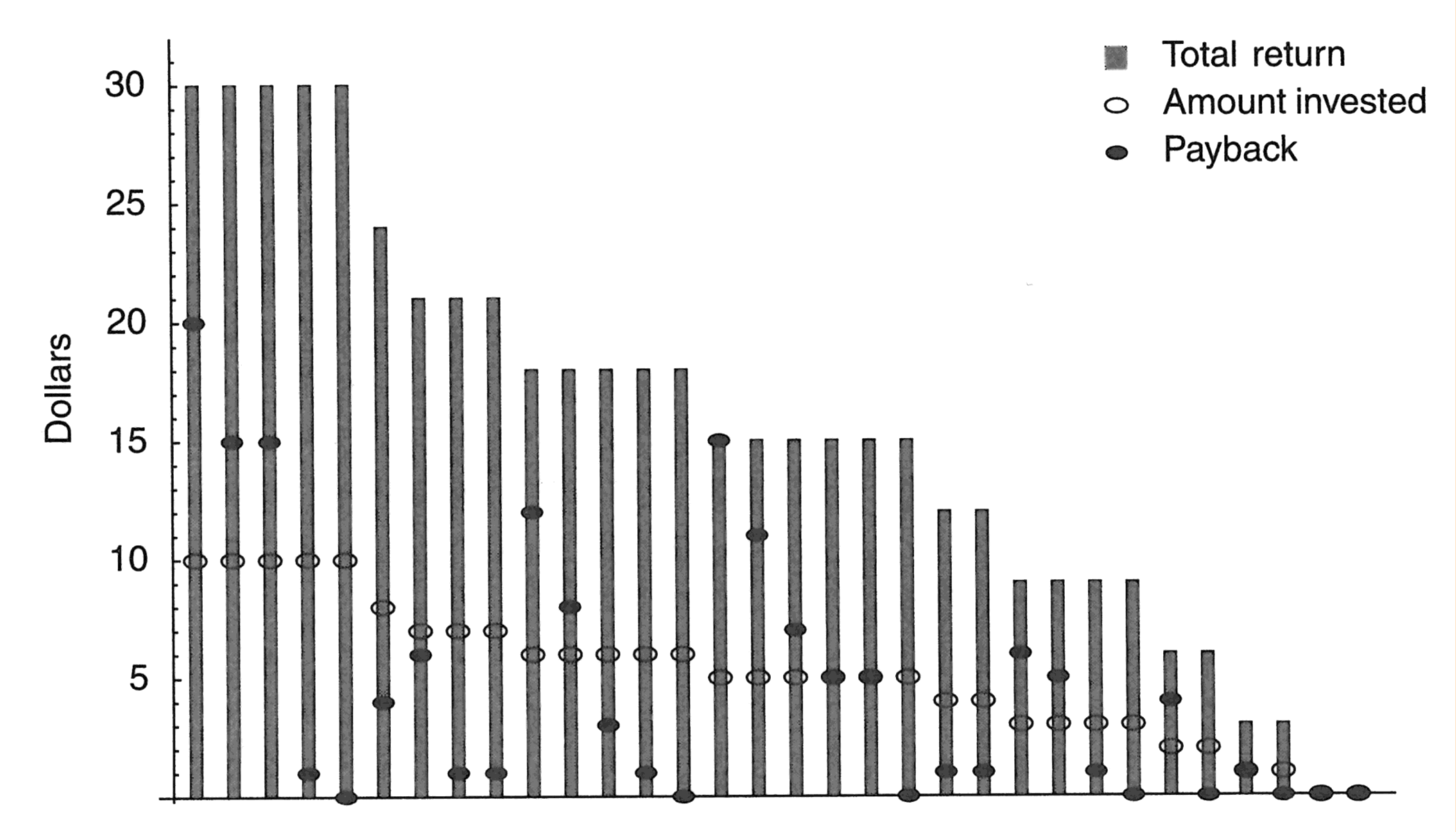
large fluctuations in human behaviour

Literature: Camerer, Behavioral Game Theory, (Princeton Univ. Press, 2003)

Experiments by Berg et al.

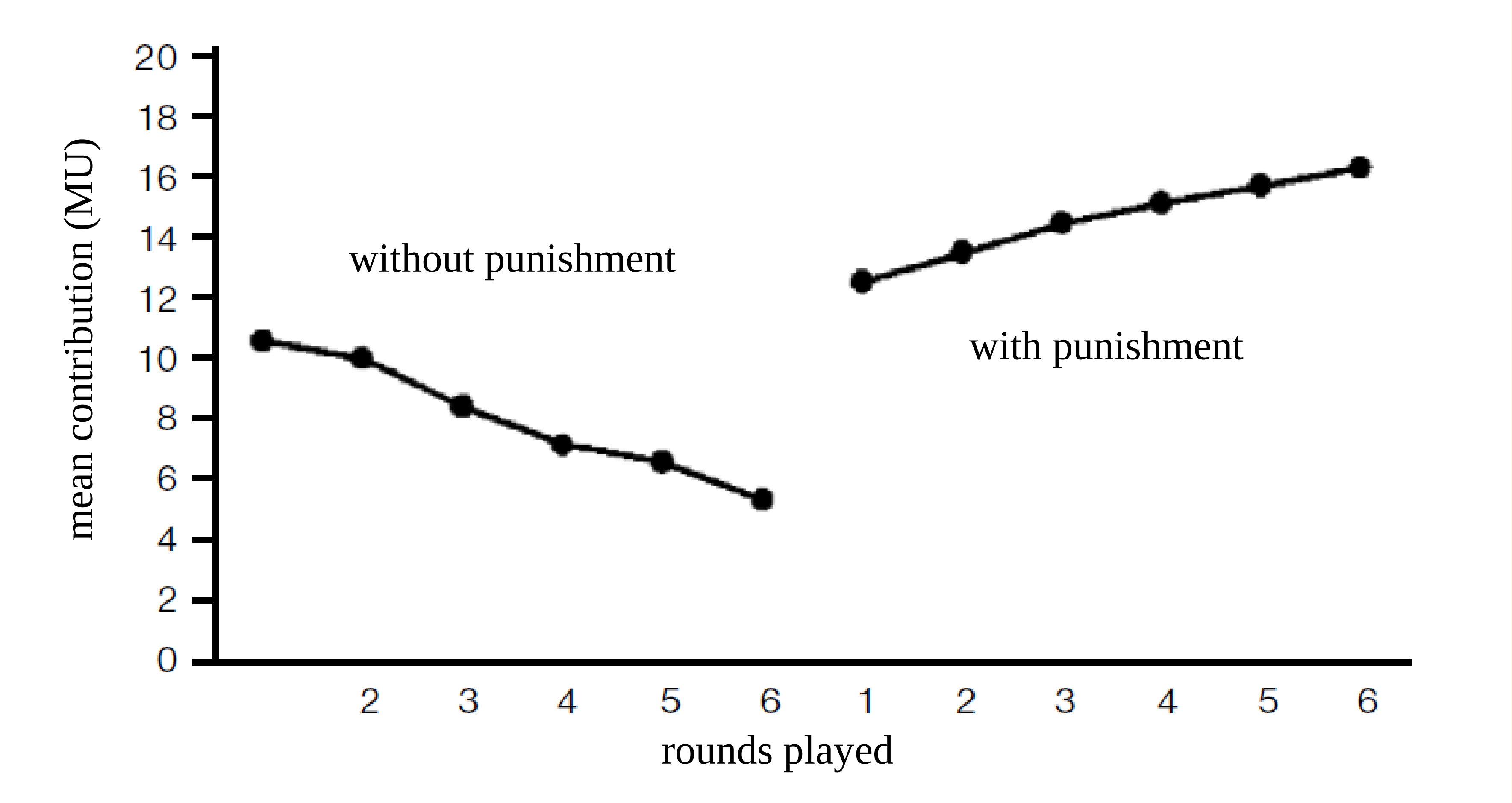
32 pairs of players

Ordered according to the investments



Public goods game:

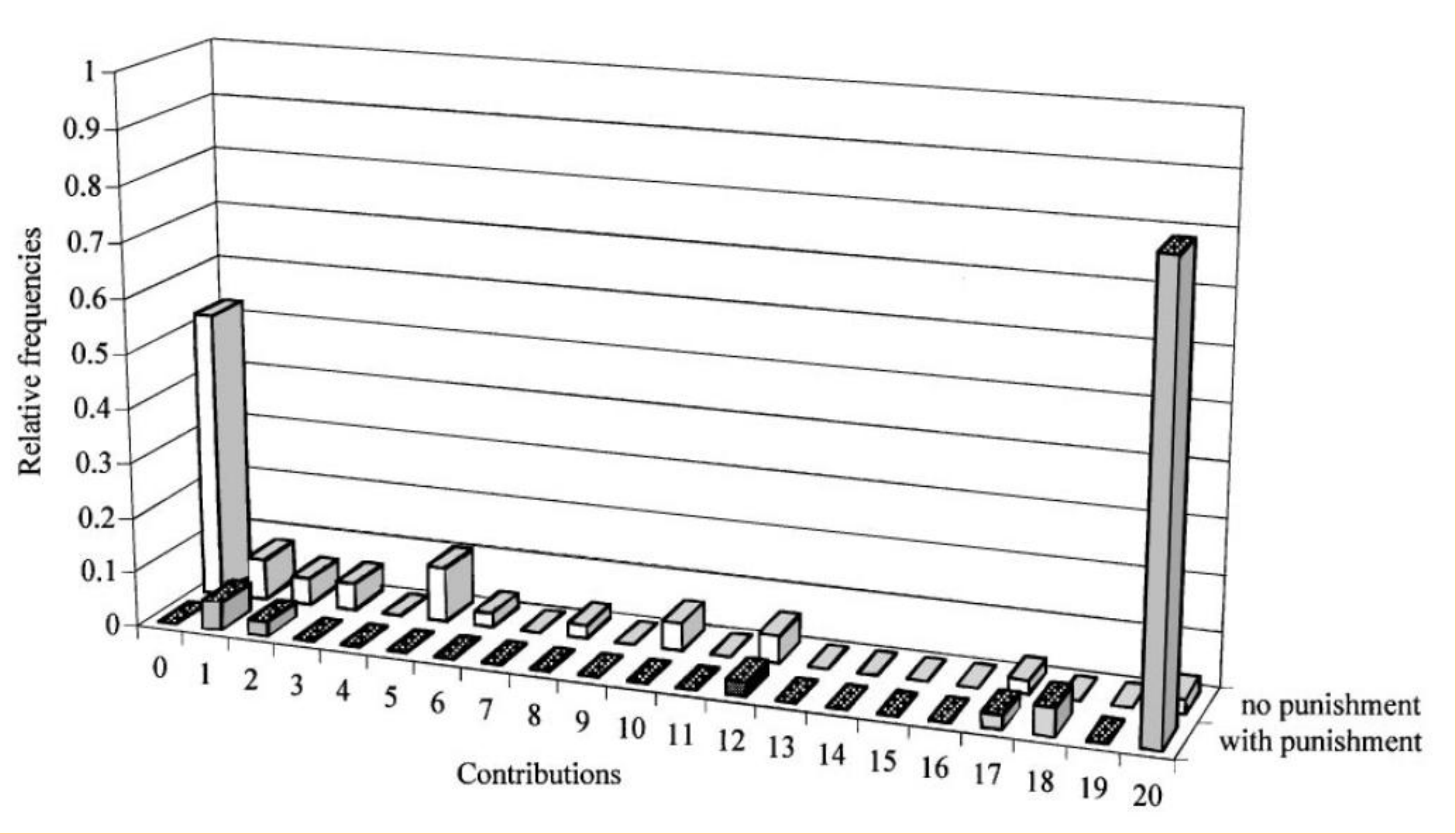
Students played public goods games. The players were divided into four-person groups after each round. After the 6th round the possibility of punishment was introduced. Each cooperator could impose a punishment (at a cost of decreasing their own income) reducing the defectors' income by a fine (higher than the cost of punishment). It is an altruistic form of punishment as the player groups were rearranged after each round. [Fehr and Gächter, *Nature* 415 (2002) 137]



Fehr and Gächter [*Eur. Econ. Rev.* 42 (1998) 845–859]:

investment between 0 and 20 in a public goods game with or without punishment

Quantitative comparison of the frequency of a given investment in the final round:



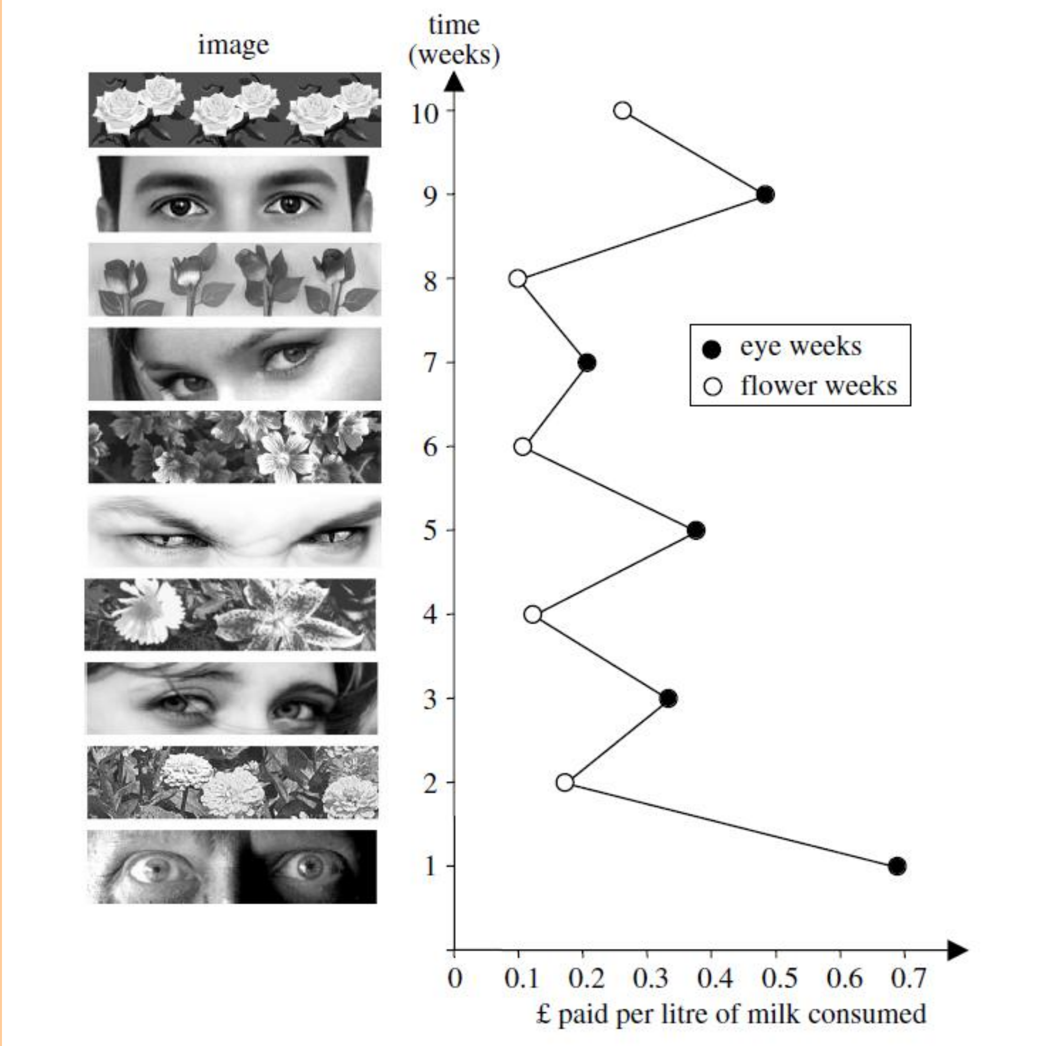
Being watched

Bateson et al., [*Biology Letters* 2 (2006) 412–414.]

The members of a department had the option to pay for tea, coffee, and milk via an honesty box. A notice explaining this arrangement was displayed at eye height.

The notice also featured a banner above the prices that alternated each week between a pair of watching eyes and flowers.

Results:



Conclusion:

Our cooperative behaviour can be affected by weak social cues.

Brain research

Using tomography the neurobiological activity of the players' brain was investigated while playing (ultimatum) games.

First results: - brain reaction is fast (no thinking) and comes from the ancient region and sometimes similar to anger/fury

- brain rewards the punisher by releasing hormones that promote positive feelings
- brain reaction depends on the personal profile

Takahashi et al., *PNAS* 109 (2012) 4281; Gabay et al., *Neurosci. Biobehav. Rev.* 47 (2014) 549

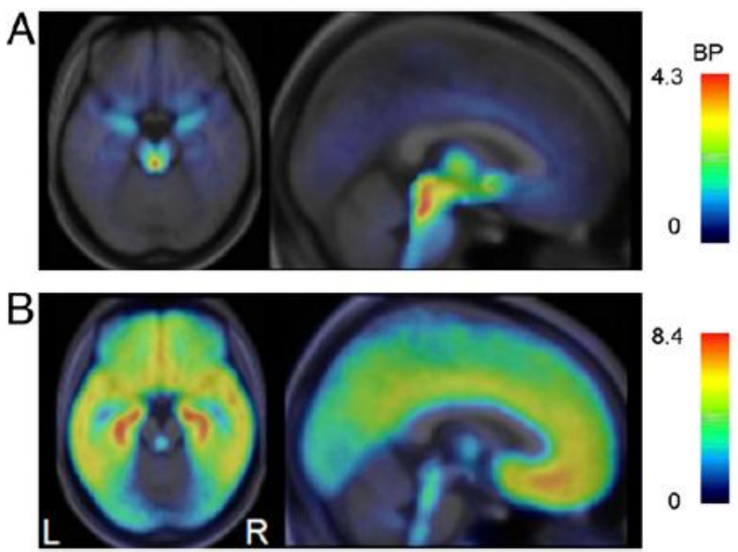


Fig. 1. Maps of 5-HTT and 5-HT1A receptor BP, averaged across participants. (*A*) [^{11}C]DASB image fused with MRI. (*B*) [^{11}C]WAY100635 image fused with MRI. Bar indicates the range of BP. In *Left* to *Right* columns, axial and sagittal planes of the brain are displayed. R, right; L, left.

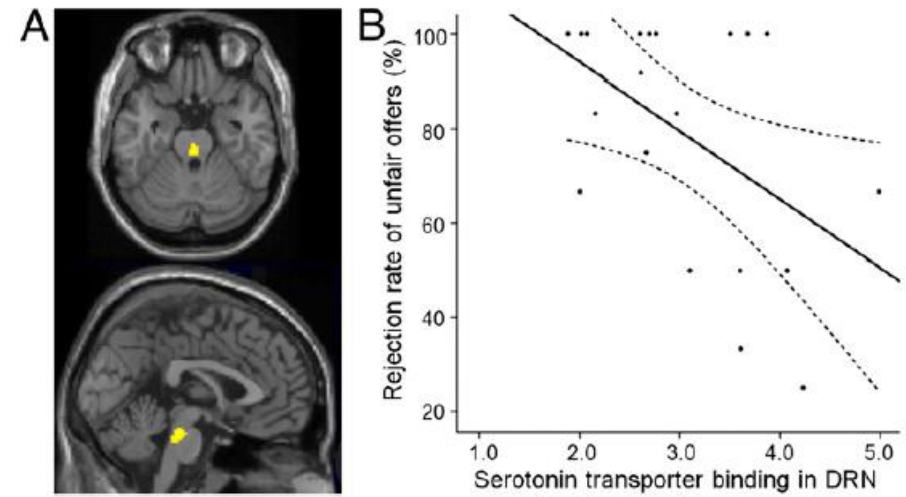


Fig. 2. Correlation between rejection rate of unfair offers in UG and 5-HTT binding in DRN. (*A*) SPM image showing regions of negative correlation between rejection rate of unfair offers and 5-HTT binding in DRN. (*B*) Plots and regression line of correlation between rejection rate of unfair offers and 5-HTT binding in DRN ($R = -0.50$, $P = 0.026$). Dashed lines are 95% confidence interval boundaries.

Home assignments

3.1. How does the repeated PD game play out for the following strategy pairs: AllD-AllC, AllD-TFT, TFT-WSLS?

3.2. Show that WSLS can win in a population consisting of WSLS, AllC and TFT players! What happens if AllD players are present too?