doi: 10.1038/nature06723

Winners don't punish

Supplementary Information

Anna Dreber[1,2*], David G. Rand[1,3*], Drew Fudenberg[4] & Martin A. Nowak[1,5,6†]
[1]Program for Evolutionary Dynamics, [3]Department of Systems Biology, [4]Department of Economics, [5]Department of Mathematics, [6]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [2]Stockholm School of Economics, Department of Economics, Stockholm, Sweden.
**These authors contributed equally to this work, [†]Corresponding author*

1. Supporting Figures and Data

Figure S1 shows the payoff matrices for each experiment. In all four payoff matrices the strategy Grim – start playing C and play C unless D has been played in the past – is a subgame-perfect equilibrium with the specified continuation probability of ¾.

C1	C	D
C	1	-2
D	3	0

T1	C	D	P
C	1	-2	-5
D	3	0	-3
P	1	-2	-5

C2	C	D
C	2	-2
D	4	0

T2	C	D	P
C	2	-2	-5
D	4	0	-3
P	2	-2	-5

Figure S1. Payoff matrices for each experiment. The row player's payoff is shown. Each game unit is worth \$0.10.

Each pairing of participants was drawn at random from the entire group. Therefore, not all interactions are independent, because some interactions share the same player. For this reason, we have not conducted our statistical analysis at the level of interactions (N between 293 and 324, depending on the session), but at the level of subjects (N between 22 and 30, depending on the session).

We have used quantile regression as opposed to ordinary least squares (OLS) regression in all of our correlation analyses. Quantile regression has been shown to perform better than ordinary least squares (OLS) for data with non-Gaussian error distributions[1-3]. Average payoff per decision is likely to have a non-Gaussian error distribution, as subjects using different strategies will presumably have payoffs centered around different values. Additionally, quantile regression is more robust than OLS to the presence of outliers[1-3]. For both of these reasons, a quantile regression is more appropriate here than an ordinary least squares regression for our data. Nonetheless, OLS regression with robust standard error gives similar results to quantile regression for our data. Regressing average payoff per decision against punishment use is significant, and gives a similar slope to quantile regression (T1: slope coefficient = -0.038, $p<0.001$; T2: slope coefficient = -0.031, $p<0.001$). Regressing average payoff per decision against probability to punish in response to defection is significant in T1 (slope coefficient = -.730, $p<0.001$), and significant in T2 (slope coefficient = -1.211, $p=0.04$) with the exclusion of one outlier, who has a probability to punish in response to defection more than 3 standard deviations greater than the mean. Quantile regression is less sensitive to outliers than OLS, and so does not necessitate the exclusion of any data.

Our random partner-matching method does not prevent cyclic interactions, such as A playing with B, then B playing with C, and then A playing with C. To assess whether these cycles affect our conclusion, we have examined the effect of ignoring such interactions. Excluding all interactions between A and C such that A played with B, then B played with C, and then A played with C (~55% of decisions in C1 and T1, and ~66% of interactions in C2 and T2), we still find a significant negative correlation between punishment use and average payoff per round (Quantile regression; T1 slope = -0.068, t = -2.95, p = 0.006; T2, slope = -0.086, t = -2.26, p = 0.034).

As described in the main text, the option for costly punishment significantly increases the average cooperation frequency (Fisher's Exact Test, two-tailed; C1 vs T1: $p < 0.001$; C2 vs T2: $p < 0.001$), as shown in Fig. S2A. But the average payoff per round is not significantly different between each control and its corresponding treatment (Mann-Whitney test; C1 vs T1: $z = 1.043$, $p = 0.30$; C2 vs T2: $z = -0.231$, $p = 0.82$), as shown in Fig. S2B. Therefore, punishment does not provide any advantage for the group. Additionally, the variance in payoffs is larger with punishment than without (C1, std = 14.4; T1, std = 21.5; C2, std = 20.8; T2, std =25.0).

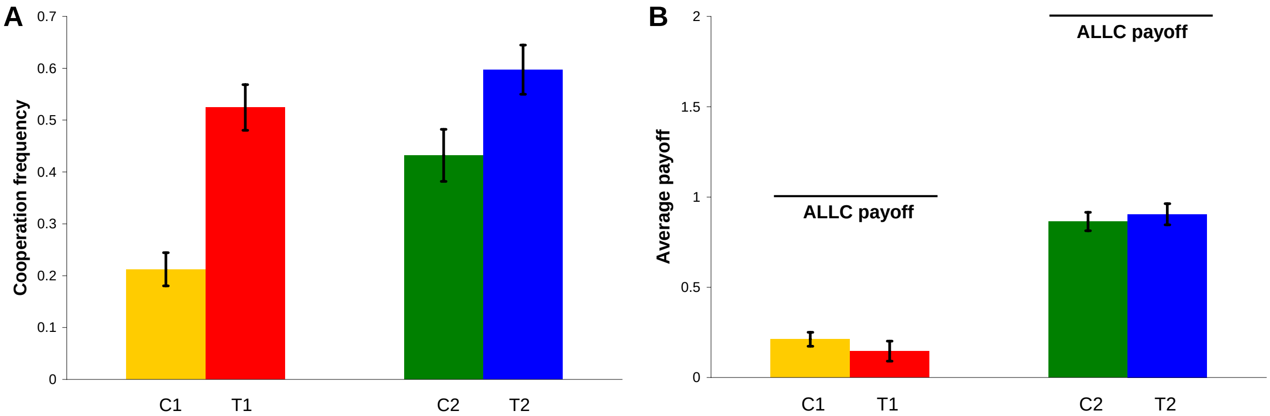

Figure S2. Cooperation frequency (A) and average payoff per round (B) in each session. Error bars represent standard error of the mean. There is significantly more cooperation in each treatment than in the corresponding control, but no significant difference in the average payoff. All control and treatment payoffs are significantly lower than the optimally cooperative payoff for ALLC play.

In our control experiment C1 (with a b/c ratio of 2), unprovoked defection increases over the course of the session (Quantile regression, unprovoked defections occurrence against interaction number; slope = 0.5, $t = 5.06$, $p < 0.001$). This means people 'learn' to defect. In the control experiment C2 (with a b/c ratio of 3), unprovoked defection decreases over the course of the session (Slope = -0.1875, $t = -2.65$, $p = 0.014$); hence, people 'learn' to cooperate. Interestingly, in both treatments, T1 and T2, there is no significant change of unprovoked defection over the course of the session (T1: $t = 0$, $p = 1.00$; T2: $t = -1.12$, $p = 0.27$). Hence, the threat of punishment seems to reduce unprovoked defection over time when comparing C1 with T1, but not when comparing C2 with T2.

In the control treatments, the correlation between average payoff and cooperation use varies between sessions, as can be seen in Fig. S3. In C1, where the benefit-to-cost ratio is 2, a significant negative correlation exists between average payoff and cooperation use (Fig. S3A; Quantile regression; $t = -2.61$, $p = 0.015$, slope $= -0.50$). In C2, where the benefit-to-cost ratio is 3, no such correlation exists (Fig. S3B; Quantile regression; $t = 0.77$, $p = 0.452$).

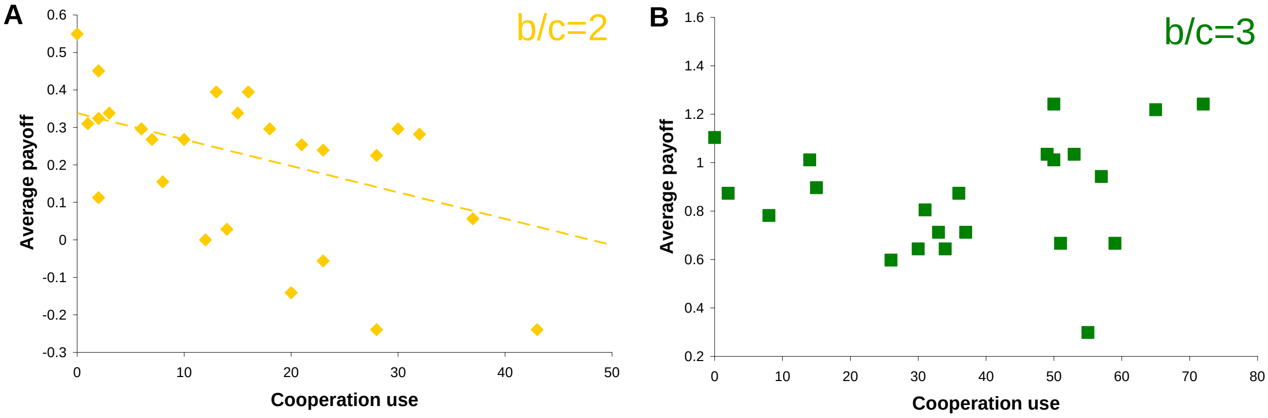

Figure S3. In the control sessions, the benefit-to-cost ratio affects the correlation between cooperation use and average payoff. A: In C1, where the benefit-to-cost ratio is 2, there is a negative correlation between cooperation use and average payoff. B: In C2, where the benefit-to-cost ratio is 3, no such correlation exists.

In addition to the negative correlation between average payoff and punishment use, and between average payoff and probability to use punishment in response to defection, there is further evidence that punishment use is the main determinant of payoff in the treatments. As shown in Fig. S4, punishers get lower payoffs than non-punishers, despite being equally cooperative. There is no significant difference in the frequency of cooperation between punishers and non-punishers (Fig. S4A; Mann-Whitney test; T1, $z = -0.392$, $p = 0.69$; T2, $z = 0.900$, $p = 0.37$). Punishers have significantly lower average payoffs than non-punishers (Fig. S4B; Mann-Whitney test; T1, $z = 3.262$ $p = 0.001$; T2, $z = 2.502$, $p = 0.012$). Both of these results are robust to the inclusion of players who might have "trembled" and punished only once. The fact that punishers are as cooperative as non-punishers refutes the possibility that the real difference driving payoffs is cooperation. It could be thought that non-punishers got higher payoffs because they were defectors whereas the punishers were cooperators, but this is not the case. This analysis further demonstrates that the key difference between high and low earners is the use of punishment.

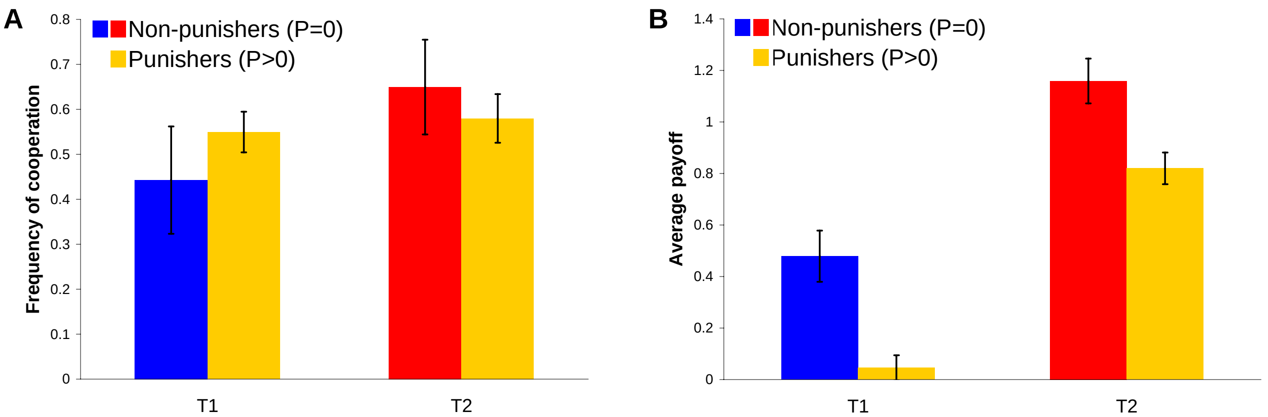

Figure S4. Punishers get lower payoffs than non-punishers, despite being equally cooperative. Error bars represent standard error of the mean. A: There is no significant difference in the frequency of cooperation between punishers and non-punishers. B: Punishers have significantly lower payoffs than non-punishers. Both of these results are robust to the inclusion of players who might have "trembled" and punished only once.

To assess the effect of experience, we have examined the last ¼ of interactions. In T1 these are interactions 15-21. In T2 these are interactions 20-27. In both treatments, there is still a strong negative correlation between average payoff and punishment use, when considering only the final ¼ of interactions (Quantile regression; T1: slope = -0.167, t = -4.85, p<0.001; T2 slope = -0.170, t = -4.29, p<0.001).

Therefore, we conclude that the benefits of punishment are not increasing with experience in our experiment. Even in the last ¼ of interactions, it is the case that winners don't punish.

In the traditional approach, the 'punishment' for non-cooperative behavior is defection[4-10]. Tit-for-tat, for example, cooperates when the co-player has cooperated and defects when the co-player has defected. The proposal of strong reciprocity[11-12] is to use costly punishment, P, instead of defection, D, in response to a co-player's defection. Our data show that such behavior is maladaptive: winners use classical tit-for-tat like behavior [5,7], while losers use costly punishment.

Average move frequency as a function of round in the treatment sessions is show in Fig. S5. Cooperation use decreases over the course of an interaction, while defection use increases. Although this may appear to be an effect of players inappropriately anticipating the game's end despite the constant probability of continuation each round, this is not necessarily the case. This same pattern could be explained by a constant probability to defect coupled with the tit-for-tat style response to defection.

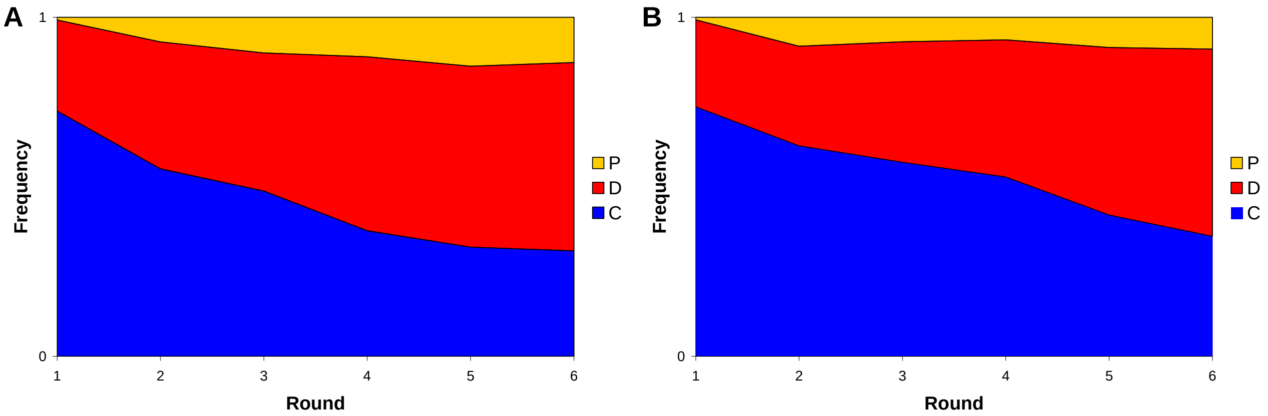

Figure S5. Average frequency of cooperation (blue), defection (red), and punishment (yellow) over the course of an interaction, for sessions T1 (A) and T2 (B). As the number of rounds increases, cooperation decreases and defection increases.

As can be seen in Fig. S6, punishment use is almost non-existent in the first round, as one would expect. In T1, punishment use increases over the course of an interaction. In T2, punishment use is essentially constant for rounds 2 to 6.

The round in which punishment is first used in a given interaction is shown in Fig. S6. Consistent with Fig. S5, punishment is rare in the first round. The first use of punishment is most likely to occur in the second round.

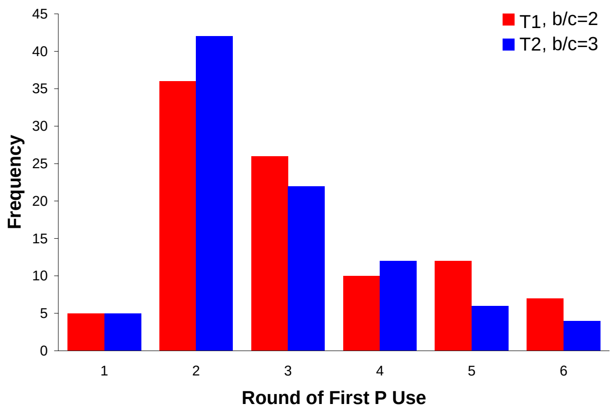

Figure S6. Histogram of rounds in which punishment is first used. Most often, punishment is first used during round 2, in response to the action taken by the other player on round 1.

1. Koenker, R. & Bassett, G. Regression Quantiles, *Econometrica*, **46**, 33-50 (1978).

2. Koenker, R. & Hallock, K. Quantile Regression, *Journal of Economic Perspectives*, **15**, 143-156 (2001).

3. Cade, B. & Noon, B.. A Gentle Introduction to Quantile Regression for Ecologists. *Frontiers in Ecology and the Environment*, **1**, 412-420 (2003).

4. Trivers, R. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35-57 (1971).

5. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390-1396 (1981).

6. Fudenberg, D. and Maskin, E. Evolution and cooperation in noisy repeated games. *Am. Econ. Rev.* **80**, 274-279 (1990).

7. Nowak, M.A. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250-253 (1992).

8. Binmore, K. & Samuelson, L. Evolutionary stability in repeated games played by finite automata. *J. of Econ. Theory* **57**, 278-305 (1992).

9. Colman, A. M. *Game Theory and its Applications in the Social and Biological Sciences* (Routledge, New York, 1995).

10. Nowak, M.A. Five rules for the evolution of cooperation. *Science* **314**, 1560-1563 (2006).

11. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785-791 (2003).

12. Boyd, R., Gintis, H., Bowles, S. & Richerson, P.J. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA* **100**, 3531-3535 (2003).

2. Experimental Instructions (for session T1)

Thank you for participating in this experiment.

Please read the following instructions very carefully. If you have any questions, please do not hesitate to ask us. Aside from this, no communication is allowed during the experiment.

In this experiment about decision making, you have been randomly matched with another person in the room. Neither of you will ever know the identity of the other. Everyone will receive a fixed amount of $15 for participating in the experiment. In addition, you will be able to earn more money based on the decisions you make in the experiment. The fixed amount and the money that you earn will be paid to you in cash immediately after the experiment is over.

You will interact several times with several different people. **Based on the choices made by you and the other participants** over the course of these interactions, you will receive between **$0 and $25, in addition to the $15 show-up amount.**

The Interaction:

There are **three possible options** available to **both you and the other person** in every round of the experiment: **A, B or C**. **Throughout the experiment, the person who makes a decision will consider him/herself as 'You' and consider the other person as 'The other person'**.

The payoffs of the options (**in units**)

Option	You will get	The other person will get
A:	−1	+2
B:	+1	−1
C:	−1	−4

1 unit = $0.10

If **you choose A** then **you** will get **−1 units**, whereas the **other person** will get +**2 units**.

If **you choose B** then **you** will get +**1 units**, whereas the **other person** will get **−1 units**.

If **you choose C** then **you** will get **−1 unit**, whereas the **other person** will get **−4 units**.

An experiment round is composed of two steps:

Step 1:

Both you and the other person begin by choosing one of these three options: A, B or C. There is a time limit on each decision. If you take more than 25 seconds a random choice will be picked for you, so **it is very important that you not take longer than 25 seconds.**

Step 2:

You and the other person are presented with each other's choice. Your score for round 1 will be calculated and presented to you on your computer screen. Your score in every round of the experiment is the sum of your payoff from your chosen option and of your payoff from the other person's chosen option. **Your score each round is thus determined by both your decision and the other person's decision, from step 1 and step 2.** ***See the examples below for clarification***.

The number of rounds in an interaction is determined by a random mechanism. The probability that there will be another round is ¾. **Therefore, each pair will interact another round with probability 3/4.**

Your behavior will have no effect on the number of rounds. Every round will follow the same pattern of two steps. The total scores will be calculated when the interaction is finished. Thereafter, you will be **anonymously and randomly matched with another student** and will **repeat the same task again**. This change of person that you are interacting with will occur several times.

The score (number of units) that you have at the end of these interactions will determine how much money you earned in total. Therefore, the additional money you and the other persons each earn depends on which options you both choose. However, **the final scores of the other participants do not matter for your final score.**

Examples:

The payoffs of the options (**in units**)

Option	You will get	The other person will get
A:	−1	+2
B:	+1	−1
C:	−1	−4

If you **both choose A** then each of you will get **+1**
(-1 from yourself, +2 from the other = +1 total)

If you **both choose B** then each of you will get **0**
(+1 from yourself, -1 from the other = 0 total)

If you **both choose C** then each of you will get **-5**
(-1 from yourself, -4 from the other = -5 total)

If **person 1 chooses A**, and **person 2 chooses B** then **person 1** gets **–2** (-1 from person 1, -1 from person 2) and **person 2** gets **+3** (+2 from person 1, +1 from person 2).

If **person 1 chooses C**, and **person 2 chooses A** then **person 1** gets +**1** (-1 from person 1, +2 from person 2) and **person 2** gets **-5** (-4 from person 1, -1 from person 2).

If **person 1 chooses B**, and **person 2 chooses C** then **person 1** gets **–3** (+1 from person 1, -4 from person 2) and **person 2** gets **–2** (-1 from person 1, -1 from person 2).

Earning additional money:

In addition to the $15 show-up fee, you will begin the experiment with an additional $5. This is the base line, which corresponds to 0 game units.

Based on your decisions in this experiment, units will be added or subtracted from this initial amount. At the end of all the interactions, your total monetary payoff will be computed to determine the amount of money earned.

If you have a total score of **0** after completing all the interactions, you will have earned the additional **$5** in the experiment.

If you have a total score **above 0**, the exchange rate will be **1 unit = $0.10**. The maximum amount that you can earn will be **$25**, however, and this is rather unlikely to happen.

If you have a total score of **less than 0**, the exchange rate will be **1 unit = $0.10**, such that negative units will be withdrawn from the initial $5. However, you cannot lose more than the initial $5, so you will always walk away here with at least the **$15** show-up fee.

1 unit = $0.10