The Long-Run Benefits of Punishment

Simon Gächter,* Elke Renner, Martin Sefton

We investigated experimentally whether costly punishment (*1*, *2*) to enforce socially beneficial cooperation can improve group welfare. This question is motivated by recent evolutionary (group selection) models of altruistic cooperation and punishment (*2*, *3*). In these models, costly punishment sustains costly cooperation. Once cooperation is established, the costs of punishment are low because punishment is rarely needed. Thus, the evolutionary pressure against punishers at that stage will be weak and can be overcome by the group average benefits from cooperation (which may come from an increased likelihood of winning intergroup contests). Recent research challenged these models because in many experiments the incurred costs of punishment outweighed the gains from increased cooperation: Punishment in these experiments was detrimental, not beneficial (*4*–*8*). For instance, a recent study (*6*) reported cooperation experiments with and without punishment conducted in 16 participant pools around the world. With the exception of three participant pools, the average payoff in experiments with punishment opportunities was lower than the average without punishment; and in those three participant pools with higher payoffs the increase was very small.

The evidence that punishment is detrimental stems from short experiments (typically 10 periods or less), and in many of them payoffs improve over time (*4*–*7*). Thus, the time horizon of these experiments may be too short to fully reveal the effects of punishment. By contrast, the evolutionary models make predictions about ancestral groups that interacted frequently over very long periods.

We examined whether the duration of interaction affects the efficacy of punishment by running public goods experiments with punishment (the P experiments) and by varying the time horizon: The experiment either lasted 10 periods (labeled P10) or 50 periods (P50). We also ran experiments with no punishment opportunities (N10 and N50) because a long time horizon might also increase cooperation without punishment and therefore reduce the scope for punishment to be beneficial. Furthermore, in line with frequent assumptions of the group selection models, our design attempted to replicate situations in our ancestral past in which small groups had to solve public goods problems with a limited set of people who interacted over a long period of time with each other.

Participants ($n = 207$) played the public goods experiments in groups of three and knew that the group membership would stay constant for the announced duration. Participants had an endowment of 20 tokens that they could either keep or contribute to a public good. Each token kept yielded one money unit (MU) for that subject, and each token invested yielded 0.5 MUs for each group member. One unit of punishment cost the punisher one MU and reduced the punished group members' earnings by three MUs (*9*).

On average, cooperation was significantly higher in both P experiments than in the respective N experiment. Per-period contributions were 3.6 tokens higher in P10 than N10 ($P = 0.0343$) (*10*) and 9.6 tokens higher in P50 than N50 ($P = 0.0000$). Per-period contributions were substantially higher in P50 than in P10 (by 4.9 tokens, $P = 0.0027$). By contrast, per-period contributions in N50 were slightly lower than in N10 (by 0.9 tokens, $P = 0.1201$).

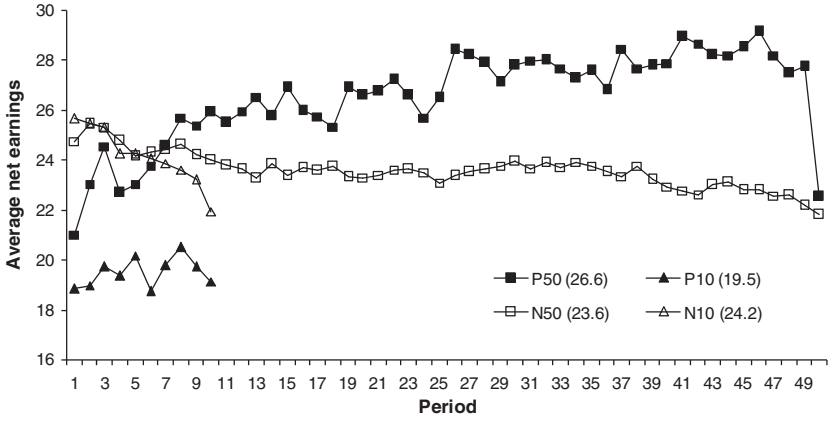

Fig. 1. Average net earnings in public goods experiments with punishment opportunities (P experiments) and with no punishment opportunities (N experiments). The experiments lasted either 10 periods (P10 and N10) or 50 periods (P50 and N50). The numbers in parentheses are the average earnings across all periods (N10 and N50, earnings from the public goods contribution stage; P10 and P50, earnings from the public goods contribution stage minus punishment expenditures minus received punishment).

The presence of a punishment option decreased the average net earnings in P10 compared with those of N10 (by 4.68 MUs per period, $P = 0.0329$). By contrast, average net earnings were significantly higher in P50 than in N50 (by 2.98 MUs per period, $P = 0.0065$) (Fig. 1). In the 10-period experiments, most groups in the N experiment did better than most groups in the P experiment, whereas in the 50-period experiments the opposite held.

The prospect of a longer duration of interaction influenced behavior already in early periods. Contributions were significantly higher, and incurred punishment costs were significantly lower, in the first 10 periods of P50 than in P10. As a result net earnings were significantly higher already in the first 10 periods of P50 compared with P10, where punishment was detrimental in all rounds. Also apparent in Fig. 1 is a prominent drop in net earnings in the last period of P50. The fixed termination period of the experiment is arguably not a feature of real interactions, and the beneficial effect of punishment would be even greater if the final period were excluded from the analysis (*9*).

Overall, our experiments show that punishment not only increases cooperation, it also makes groups and individuals better off in the long run because the costs of punishment become negligible and are outweighed by the increased gains from cooperation. These results support group selection models of cooperation and punishment (*2*, *3*), which require that punishment increases not only cooperation but also group average payoffs.

References and Notes

1. K. Sigmund, *Trends Ecol. Evol.* **22**, 593 (2007).
2. R. Boyd, H. Gintis, S. Bowles, P. J. Richerson, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3531 (2003).
3. S. Bowles, *Microeconomics: Behavior, Institutions, and Evolution* (Princeton Univ. Press, Princeton, NJ, 2003).
4. E. Fehr, S. Gächter, *Nature* **415**, 137 (2002).
5. Ö. Gürerk, B. Irlenbusch, B. Rockenbach, *Science* **312**, 108 (2006).
6. B. Herrmann, C. Thöni, S. Gächter, *Science* **319**, 1362 (2008).
7. M. Egas, A. Riedl, *Proc. R. Soc. Lond. B. Biol. Sci.* **275**, 871 (2008).
8. A. Dreber, D. G. Rand, D. Fudenberg, M. A. Nowak, *Nature* **452**, 348 (2008).
9. Materials and methods are available as supporting material on *Science* Online.
10. All statistical tests are two-sided Mann-Whitney tests with the group averages over all periods as the independent observations.
11. This work was supported by the University of Nottingham and the British Academy.

Supporting Online Material

www.sciencemag.org/cgi/content/full/322/5907/1510/DC1
Materials and Methods
SOM Text
Fig. S1
Table S1
References and Notes

18 August 2008; accepted 16 September 2008
10.1126/science.1164744

Centre for Decision Research and Experimental Economics, University of Nottingham, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, UK.

*To whom correspondence should be addressed. E-mail: simon.gaechter@nottingham.ac.uk