LETTERS

Parochial altruism in humans

Helen Bernhard[1], Urs Fischbacher[1] & Ernst Fehr[1,2]

Social norms and the associated altruistic behaviours are decisive for the evolution of human cooperation[1–9] and the maintenance of social order[10], and they affect family life, politics[11] and economic interactions[12]. However, as altruistic norm compliance and norm enforcement often emerge in the context of inter-group conflicts[13,14], they are likely to be shaped by parochialism[15]—a preference for favouring the members of one's ethnic, racial or language group. We have conducted punishment experiments[16], which allow 'impartial' observers to punish norm violators, with indigenous groups in Papua New Guinea. Here we show that these experiments confirm the prediction of parochialism. We found that punishers protect ingroup victims—who suffer from a norm violation—much more than they do outgroup victims, regardless of the norm violator's group affiliation. Norm violators also expect that punishers will be lenient if the latter belong to their social group. As a consequence, norm violations occur more often if the punisher and the norm violator belong to the same group. Our results are puzzling for evolutionary multi-level selection theories based on selective group extinction[2–5] as well as for theories of individual selection[17–19]; they also indicate the need to explicitly examine the interactions between individuals stemming from different groups in evolutionary models.

The human species is unique in the extent to which it regulates social life with normative obligations that constrain selfish behaviour[1,20]. Social norms such as food sharing, or those related to cooperative hunting and participation in warfare, shaped human life throughout important evolutionary phases[1,21–23]. It is therefore likely that the existence of these norms had a deep impact on the properties of human altruism, because norm obedience and norm enforcement involve important altruistic behaviours. The fact that social norms are group level phenomena[24] suggests that parochial social instincts[15,25]—which we define as preferences for favouring the members of one's own social group—may have shaped human altruism in decisive ways. Norms emerge through interactions in groups and apply to interactions within groups; group members enforce them, and they often arise in the context of inter-group conflicts[13,14]. Normative obligations are thus likely to apply only to ingroup members; people who do not belong to the group neither obey the norm nor benefit from the altruistic behaviours the norm enforces.

In this paper, we study the potentially parochial nature of altruistic norm enforcement by conducting third-party punishment experiments[16] with members of two small, distinct, cohesive and non-hostile indigenous groups in the Western Highlands of Papua New Guinea (PNG)—the Wolimbka and the Ngenika. Centralized institutions for the enforcement of legal rules are largely absent in PNG, meaning that social norms almost exclusively regulate social life. In addition, PNG societies more closely resemble the human societies under which our social instincts evolved than the modern, complex societies in which most people at present live. PNG is therefore an ideal environment for studying the parochial nature of human altruism.

The use of cohesive, indigenous groups from simple societies distinguishes our study from minimal group experiments[26,27] based on artificially created laboratory groups of students. Researchers have found some forms of ingroup favouritism in simple allocation experiments based on the minimal group paradigm. However, the ingroup favouritism observed in these experiments[26,27] does not reveal parochial tendencies in human altruism because the subjects only distributed resources between two other subjects. Subjects thus bore no cost, regardless of how they allocated the available resources between the other two subjects. Moreover, subsequent research[28,29] has shown that costless ingroup favouritism in these experiments is due to the expectation that ingroup members will receive some reciprocation from other ingroup members. Thus, what looked like a preference for ingroup favouritism was in fact based on the expectation of ingroup reciprocity[29].

In our study, we conducted an anonymous, one-shot, third-party punishment experiment[16] involving a dictator (player A), a recipient (player B) and a third party (player C). Player A receives an endowment of 10 Kina (K), which is equivalent to a high daily labourer's wage in the informal sector of the economy. Player B, the recipient, has no endowment and player C, the third party, receives K5. First, A can transfer any amount between K0 and K10 to B. Then C is informed about A's transfer and has the opportunity to punish A's action by spending K0, K1 or K2 on punishment. Every Kina spent on punishment reduces A's income by K3. After player A and C had chosen their actions, we also elicited their expectations about how the dictators would be punished at the three transfer levels K0, K5 and K10.

Our experiment is designed to capture the altruistic enforcement of egalitarian sharing norms that ethnographic studies have documented[22,30,31]. Such sharing norms are beneficial for the group because they insure group members against the uncertainties in individual food acquisition success. If an egalitarian sharing norm exists, we should observe both that dictators transfer money to the recipients and that third parties exhibit altruistic punishment[7] for transfer levels below the equal split. As we wanted to examine the parochial behavioural patterns, we allocated each subject in our study to one of the following four treatment conditions. (1) All three players in the game are from the same tribe (treatment ABC). (2) Only players B and C are from the same tribe, while A is an outgroup member (BC). (3) Only players A and B are from the same tribe (AB). (4) Only players A and C are from the same tribe (AC). The decision-makers in all four treatments were informed about the other two players' group affiliation.

Current evolutionary models based on the idea that human altruism evolved through the selective (cultural or biological) extinction of groups in inter-group conflicts[2–5] predict the following punishment pattern. No punishment should be observed in treatment AC if A does not share, because there is no obligation to share with a recipient B from the outgroup. Sharing with an outgroup member would only help the competing group, possibly at the

[1]University of Zurich, Institute for Empirical Research in Economics, Blumlisalpstrasse 10, CH-8006 Zurich, Switzerland. [2]Collegium Helveticum, Schmelzbergstrasse 25, CH-8092 Zurich, Switzerland.

© 2006 Nature Publishing Group

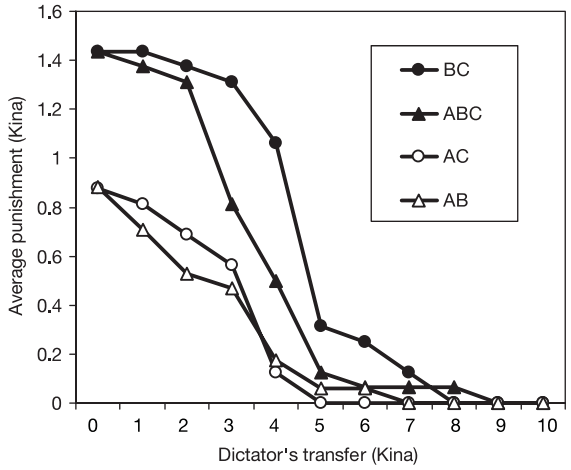

Figure 1 | Punishment pattern across conditions in the third-party punishment game. The third party spends money (Kina) on punishing the dictator for the violation of an egalitarian sharing norm in all conditions. Punishment expenditures by third parties increase if dictators reduce transfers below the egalitarian sharing norm; punishment is highest in all conditions if the dictator does not share at all. Punishment expenditures are much higher in the ABC condition, where dictator (A), recipient (B) and third party (C) belong to the same group, than in the AB condition (where only the dictator and the recipient belong to the same group) and the AC condition (where only the dictator and the third party belong to the same group). Contrary to predictions, punishment is also much higher in the BC condition (where the recipient and the third party belong to the same group) than in the AB and the AC condition.

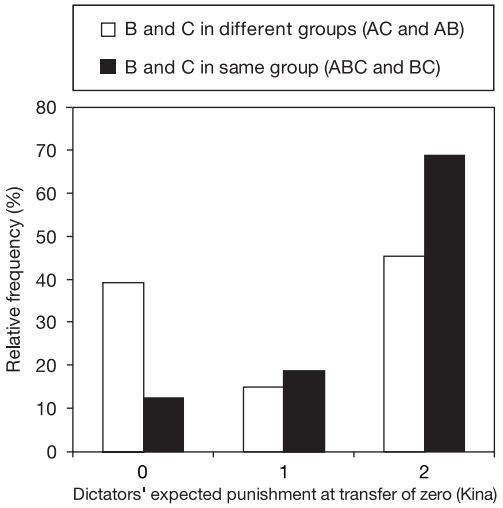

Figure 2 | The punishment threat for the lowest transfer level as perceived by the dictators. The punishment threat subjectively perceived is important, because higher expectations about punishment induce more compliance with the sharing norm. In those conditions where the recipient (B) and the third party (C) belong to the same group (that is, in conditions ABC and BC), dictators perceive a higher punishment threat than in conditions where B and C belong to different groups (that is, in conditions AB and AC).

expense of the ingroup. The same punishment prediction is made for condition AB, where the third party is an outgroup member; in this case A violates a sharing norm if she or he does not share with B, but player C—being an outgroup member—has no obligation to punish a norm violation within the other group. Likewise, the group competition theories predict zero punishment in condition BC because A is an outgroup member who is not obliged to obey a sharing norm in interactions with an ingroup member. Thus, the ABC treatment is the only condition in which evolution should have favoured altruistic punishment, because punishment in ABC sustains group norms that enhance a group's ability to compete with other groups.

We observe that the punishment pattern is qualitatively similar in all four treatment conditions (Fig. 1). There is little punishment for transfers at and above the egalitarian level, while sharing decisions that give the dictator a larger share of the 'pie' are more heavily punished the more the dictator deviates from the equal split. For example, even in the conditions with relatively low punishment levels (AB and AC), 58% of the third parties punish if the dictator transfers nothing, but only 3% punish at K5 and nobody punishes at transfers higher than K6. This finding suggests the existence of an egalitarian sharing norm in all four conditions, and not just in the ABC condition—a fact that is puzzling in view of the predictions above. The third parties' and the dictators' beliefs further support this interpretation. Regardless of the treatment condition, individuals in both roles believe that a transfer of K0 will be punished severely, while transfers of K5 or K10 will not be punished.

Although punishment of low transfers is not zero in the AB and the AC conditions, the size of the punishment across the ABC, the AC and the AB condition obeys the order that selective extinction models[2–5] predict: punishment is much higher when all three players belong to the same group (ABC) compared to the AC and the AB treatments, where punishment is roughly the same (Fig. 1). The difference between ABC and the other two treatments is highly significant (ordered probit regression, $P = 0.007$, two sided, $N = 539$), while the difference between the AC and the AB conditions is negligible and insignificant (ordered probit regression, $P = 0.947$, two sided, $N = 539$; see also statistical methods in Supplementary Information).

Punishment in the BC condition is also much higher than in the AC and the AB conditions (Fig. 1; ordered probit regression, $P < 0.001$, two sided, $N = 539$). Controlling for the transfer level, the probability that the third party punishes a violation of the sharing norm is 30.4 percentage points higher in the BC condition than in the AC and AB conditions. In fact, punishment in BC is even significantly higher than in ABC (Fig. 1; ordered probit regression, $P = 0.022$, two-sided, $N = 352$). The difference between these two conditions is particularly large at transfer levels of K3 and K4. For example, the average punishment at a transfer level of K4 is more than twice as high in BC than in ABC. Thus, third parties are more lenient if the norm violator belongs to their group.

These findings imply that regardless of the dictator's group affiliation, punishment is much higher if the recipient and the third party belong to the same group. Thus, to the extent to which third-party punishment deters potential norm violators, the victims of potential norm violations are much more protected by the threat of third-party punishment if the third party belongs to the victim's group. The dictators' expectations also support this deterrence effect of third-party punishment. The dictators expect significantly more punishment at K0 in those conditions where B and C belong to the same group (Fig. 2; ordered probit regression, $P = 0.019$, two sided, $N = 65$), and the expected punishment level at K0 has a significantly positive effect on transfer levels. A one-unit increase in expected punishment raises the transfer level by 1.43 units (tobit regression, $P < 0.001$, two-sided, $N = 65$). Thus, both actual punishment and punishment expectations suggest that victims of norm violations are much more protected if the victim and the third party belong to the same group.

How do these punishment patterns, together with the parochial pattern of voluntary norm compliance, shape the transfer levels? We find that the transfers are higher in those conditions where A and B belong to the same group (tobit regression, $P = 0.018$, two-sided, $N = 65$). The dictators even transfer more money in these conditions if we control for their punishment expectations (tobit regression, $P = 0.086$, two-sided, $N = 65$). Thus, dictators who expect the same

© 2006 Nature Publishing Group

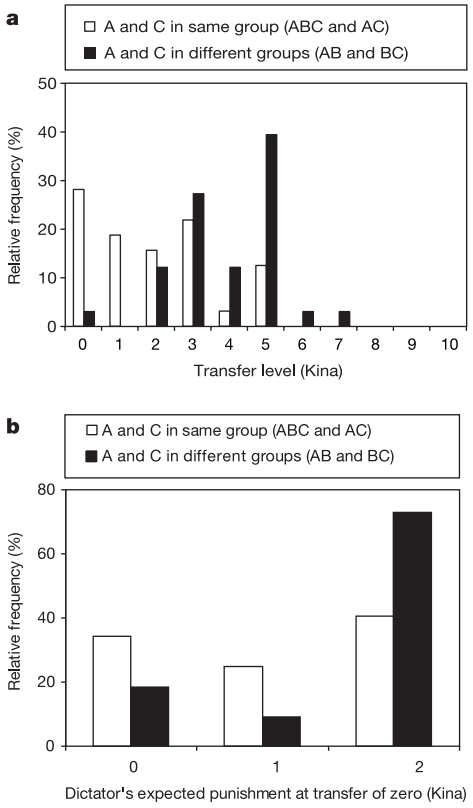

Figure 3 | Dictators' transfers and perceived punishment threats if the third party belongs to the dictator's group. **a**, Relative frequency of dictators' transfer decisions. Dictators comply much less with the egalitarian sharing norm if the third party belongs to their group. **b**, Relative frequency of dictators' perceived punishment threats in the case of a transfer of zero. Dictators believe that they will be punished much less if the third party belongs to their group.

level of punishment give more in those conditions where A and B belong to the same group, suggesting a higher degree of voluntary norm compliance in these treatments (see also statistical methods in Supplementary Information).

Player A's reluctance to share if B does not belong to his group is strongly reinforced if the third party comes from A's group. In this case the dictators' transfers are substantially lower on average compared to the conditions where A and C belong to different groups (Fig. 3a; tobit regression, $P < 0.001$, two-sided, $N = 65$). Moreover, the punishment threat subjectively perceived is significantly lower if the third party belongs to the dictator's group (ordered probit regression, $P = 0.022$, two sided, $N = 65$). For example, only 41% of the dictators expect the maximal punishment at a transfer of K0 if the third party is from their group, but 73% of the dictators expect the maximal punishment if the third party is from the other group (Fig. 3b). Thus, dictators expect that 'their' third parties will be lenient, inducing them to transfer little to the recipient.

Although our findings are puzzling from the viewpoint of current selective extinction models[2–5], they also suggest how the models could be extended in order to explain the full empirical pattern. First, these models focus on norm enforcement within groups for the purpose of winning inter-group conflicts while neglecting the potential benefits from cooperative inter-group interactions. This approach makes it difficult to understand when hostility characterizes inter-group reactions and when cooperative norms govern them. The current models therefore have problems in explaining our first main result—the existence of egalitarian sharing norms in all four conditions—but a suitably extended model, which explicitly formalizes individual strategies in inter-group encounters, may be able to capture this fact. Second, the lack of explicit modelling of individual inter-group encounters makes it also difficult to understand why—regardless of the norm violator's group affiliation—punishment is so high in those conditions where the third party and the recipient (ABC and BC) belong to the same group. Punishing outsiders who harm an ingroup victim increases the general security of all ingroup members by preventing attacks by outgroup members. If a group has a reputation for punishing individual attacks by outgroup members, the latter are deterred from such attacks and all ingroup members enjoy more protection. Thus, taking the problem of group reputation into account could possibly explain the high punishment in both the BC and the ABC conditions.

Current individual selection models[17–19] also cannot readily explain the full pattern of punishment behaviours. Models based on repeated interactions or reputation formation[9,18,19] seem to predict that punishment will be highest in the ABC condition, because the third party's reputational benefits from punishing are most favourable in this condition: protecting an ingroup victim may yield a future ally and punishing an ingroup violator lessens the likelihood of being cheated in future interactions with the norm violator. Kin selection theory[17] is also not fully satisfactory, because the average genetic relatedness between two randomly selected adult tribe members in tribes of several hundred people is rather low, due to migration and marriages with outgroup members. It is therefore difficult to see why kin selection should have favoured a sharing norm that applies to all tribe members alike—in particular in situations such as our experiment, where the dictator's fitness loss roughly equals the recipient's fitness gain. Perhaps, if a sharing norm already exists for other reasons, kin selection might have favoured a lower punishment of ingroup members, which could explain the lower punishment in ABC compared to the BC condition.

Thus, if future research confirms the robustness of our results, the parochial patterns of human altruism constitute a challenge for existing evolutionary theories. Currently, no single theory seems to be able to explain the entire pattern of parochialism across treatments, providing an opportunity for developing new theories or modifying existing ones.

METHODS

Subjects. A total of 195 members—aged 17 to 60—of two small-scale societies in the Western Highlands of PNG (the Wolimbka and the Ngenika) participated in a third-party punishment game, permitting us to conduct 65 games with three players each. We conducted 17 games in the AB treatment and 16 in each of the other treatments.

Tribal warfare is a frequent event in PNG, but the Ngenikas and the Wolimkas never conducted tribal warfare with each other within the memorized history of the older members of the two tribes. They are neutral towards each other and do not exchange any gifts or goods, except in the rare case of inter-tribe marriage. Due to the absence of any hostilities between the two tribes, finding parochialism across these two tribes makes our results even stronger.

Experimental procedures. Game instructions and procedures are based on the work of Henrich *et al.*[32]. In each experimental session, 18 participants first received some preliminary verbal instructions as a group. We ensured that the participants did not communicate about the game before the experiment. Each participant received a show-up fee of K3 and drew a number at the beginning. One at a time, in the order of the pulled numbers, the subjects then came into a separate room to participate in the experiment. The game was then explained to them verbally in much detail. Participants who failed to understand the instructions were dismissed from the experiment but could keep their show-up fee.

We elicited player C's punishment decision with the strategy method (see Supplementary Information). This means that player C indicated how much he or she is willing to spend on punishment for each of player A's feasible transfers. Player C made this decision before (s)he knew the dictator's actual transfer level. Since we collected 11 punishment decisions from each player C—one punishment decision for each feasible transfer level—we always controlled for repeated measurement in the statistical analysis of punishment decisions.

Statistical methods. The punishment decisions were examined with ordered probit regressions. We controlled for repeated individual measurements and for individuals' transfer levels in all regressions. Treatment effects were measured by dummy variables that take on a value of one if the observation comes from the

© 2006 Nature Publishing Group

treatment of interest. Otherwise the dummy variable is zero. The dictators' transfer decisions were analysed with tobit regressions that consider the transfer level as a function of the expected punishment and the treatment dummies of interest.

Received 18 April; accepted 15 June 2006.

1. Sober, E. & Wilson, D. S. *Unto Others—The Evolution and Psychology of Unselfish Behavior* (Harvard Univ. Press, Cambridge, Massachusetts, 1998).
2. Gintis, H. Strong reciprocity and human sociality. *J. Theor. Biol.* **206,** 169–179 (2000).
3. Henrich, J. & Boyd, R. Why people punish defectors–weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *J. Theor. Biol.* **208,** 79–89 (2001).
4. Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* **100,** 3531–3535 (2003).
5. Bowles, S., Choi, J.-K. & Astrid, H. The coevolution of individual behaviors and group level institutions. *J. Theor. Biol.* **223,** 135–147 (2003).
6. Sigmund, K., Hauert, C. & Nowak, M. A. Reward and punishment. *Proc. Natl Acad. Sci. USA* **98,** 10757–10762 (2001).
7. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415,** 137–140 (2002).
8. Milinski, M., Semmann, D. & Krambeck, H. J. Reputation helps solve the 'tragedy of the commons'. *Nature* **415,** 424–426 (2002).
9. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437,** 1291–1298 (2005).
10. Elster, J. *The Cement of Society—A Study of Social Order* (Cambridge Univ. Press, Cambridge, UK, 1989).
11. Lindbeck, A., Nyberg, S. & Weibull, J. W. Social norms and economic incentives in the welfare state. *Q. J. Econ.* **114,** 1–35 (1999).
12. Stutzer, A. & Lalive, R. The role of social work norms in job searching and subjective well-being. *J. Eur. Econ. Assoc.* **2,** 696–719 (2004).
13. Otterbein, K. F. *The Evolution of War: A Cross-Cultural Study* (Human Relations Area Files Press, New Haven, 1985).
14. Bornstein, G. Intergroup conflict: Individual, group, and collective interests. *Pers. Soc. Psychol. Rev.* **7,** 129–145 (2003).
15. Bowles, S. & Choi, J.-K. in *Cultural Diversity versus Economic Solidarity* (ed. Parijs, P. V.) 81–97 (Franqui Scientific Library, Brussels, 2004).
16. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **25,** 63–87 (2004).
17. Hamilton, W. D. Genetical evolution of social behaviour I & II. *J. Theor. Biol.* **7,** 1–25, 26–52 (1964).
18. Trivers, R. L. Evolution of reciprocal altruism. *Q. Rev. Biol.* **46,** 35–57 (1971).
19. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **393,** 573–577 (1998).
20. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425,** 785–791 (2003).
21. Kaplan, H. & Gurven, M. in *Moral Sentiments and Material Interests—The Foundations of Cooperation in Economic Life* (eds Gintis, H., Bowles, S. & Boyd, R.) 75–113 (MIT Press, Cambridge, Massachusetts, 2005).
22. Boehm, C. Egalitarian behavior and reverse dominance hierarchy. *Curr. Anthropol.* **34,** 227–254 (1993).
23. Wiessner, P. Norm enforcement among the Ju/'hoansi bushmen: A case of strong reciprocity? *Hum. Nat.* **16,** 115–145 (2005).
24. Henrich, J. *et al.* In search of homo economicus: behavioural experiments in 15 small-scale societies. *Am. Econ. Rev.* **91,** 73–78 (2001).
25. Shinada, M., Yamagishi, T. & Ohmura, Y. False friends are worse than bitter enemies: "Altruistic" punishment of in-group members. *Evol. Hum. Behav.* **25,** 379–393 (2004).
26. Tajfel, H., Billig, M., Bundy, R. & Flament, C. Social categorization in intergroup behaviour. *Eur. J. Soc. Psychol.* **1,** 149–178 (1971).
27. Brewer, M. In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychol. Bull.* **86,** 307–324 (1979).
28. Rabbie, J. M., Schot, J. C. & Visser, L. Social identity theory: A conceptual and empirical critique from the perspective of a behavioural interaction model. *Eur. J. Soc. Psychol.* **19,** 171–202 (1989).
29. Yamagishi, T., Jin, N. & Kiyonari, T. in *Advances in Group Processes* (eds Thye, S. R., Lawler, E. J., Macy, M. W. & Walker, H. A.) 161–197 (JAI Press, Stamford, 1999).
30. Kaplan, H. & Hill, K. Food Sharing among Ache foragers: Tests of explanatory hypotheses. *Curr. Anthropol.* **26,** 223–246 (1985).
31. Wiessner, P. & Tumu, A. *Historical Vines: Enga Networks of Exchange, Ritual, and Warfare in Papua New Guinea* (Smithsonian Institution Press, Washington DC, 1998).
32. Henrich, J *et al.* Costly punishment across human societies. *Science* (in the press).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This paper is part of the Research Priority Program 'Foundations of Human Social Behaviour–Altruism versus Egoism' at the University of Zurich. Financial support from the MacArthur Foundation (Network on Economic Environments and the Evolution of Individual Preferences and Social Norms) and the Cogito Foundation is also acknowledged.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.F. (efehr@iew.unizh.ch) and U.F. (fiba@iew.unizh.ch).

© 2006 Nature Publishing Group