Please cite this article in press as: Bshary and Bshary, Self-Serving Punishment of a Common Enemy Creates a Public Good in Reef Fishes, Current Biology (2010), doi:10.1016/j.cub.2010.10.027

Current Biology *20*, 1–4, November 23, 2010 ©2010 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2010.10.027

Report

Self-Serving Punishment of a Common Enemy Creates a Public Good in Reef Fishes

Andrea Bshary[1,*] and Redouan Bshary[1]
[1]University of Neuchâtel, Institute of Biology, Rue Emile-Argand 11, 2009 Neuchâtel, Switzerland

Summary

A key challenge for evolutionary biologists is to determine conditions under which individuals benefit from a contribution to public goods [1, 2]. For humans, it has been observed that punishment of free riders may promote contributions [3, 4], but the conditions that lead to stable cooperation based on punishment remain hotly debated [5–8]. Here we present empirical evidence that public goods may emerge as a by-product of self-serving punishment in interactions between coral reef fishes and parasitic saber-tooth blennies that stealthily attack their fish victims from behind to take a bite [9]. We first show that chasing the blenny functions as punishment [10], because it decreases the probability of future attacks. We then provide evidence that in female scalefin anthias, a shoaling species, punishment creates a public good because it increases the probability that the parasite switches to another species for the next attack. A final experiment suggests that punishment is nevertheless self-serving because blennies appear to be able to discriminate between look-alike punishers and nonpunishers. Thus, individuals that do contribute to the public good may risk being identified by the parasite as easy targets for future attacks.

Results and Discussion

Cooperation in groups is difficult to explain because individual contributions are shared among group members, a condition that is often vulnerable to free riding [1, 2]. Free riders benefit from the contribution of others but do not contribute themselves, and they therefore gain higher payoffs than contributors do. Two main solutions have emerged in studies on humans. First, contributions to public goods can be stable if players may gain a positive reputation from contributing, which increases the probability of receiving help in other situations [11–13]. Second, humans may achieve stable cooperation in n-player games if players are allowed to punish cheaters [4]. Empirical evidence for the success of punishment even in one-off interactions has led to a heated theoretical discussion about which conditions such punishment may evolve from. Models of “cultural group selection” [5, 6] propose that fast social learning within groups stabilizes punishment against the problem of so-called second-order free riding (contributing to public goods but failing to pay for the punishment of free riders). Others have pointed out that punishment in one-shot games can only evolve as a result of indirect fitness benefits [7, 8].

For a better understanding of the conditions that lead to stable punishment, it seems to be of paramount importance to identify more natural conditions in which punishment occurs and to identify the consequences of punishment on an individual’s fitness. For example, under natural conditions it seems more likely that individuals interact repeatedly. Repeated interactions may create conditions in which, at least in theory, individuals may gain direct benefits from punishment, even if it creates a public good. The empirical evidence from experiments on humans is mixed in the sense that in some studies, punishing individuals gain more money than the average nonpunishing individual, whereas other studies find the opposite [14–19]. The cited studies differ from each other with respect to many parameters such as group size, information about the behavior of others, payoff structure, strategic options, and ethnic membership, making it difficult to draw general conclusions. Here we explore the possibility that punishment yields direct benefits in an animal system in which field observations suggest that the likelihood of punishment depends on both a repeated game structure and group size [9]. In our study system, saber-tooth blennies stealthily attack other reef fish species from behind to take a bite of mucus, scales, and/or tissue [20]. The blennies occupy small territories, and resident victim species that repeatedly interact with the same blenny often chase blennies in response to bites [9]. In contrast, visiting species that may simply swim beyond a blenny’s range typically never chase blennies but swim off [9], suggesting that an enforced repeated game structure is essential for the emergence of residents chasing blennies.

We tested first whether chasing a blenny functions as punishment sensu Clutton-Brock and Parker [10]: the immediate reduction in payoffs due to the costs of chasing the blenny might be offset by the blenny seeking alternative victims in the future. If that was the case, we asked what the consequences of punishment on the parasite’s behavior might be if the punisher was a member of a shoaling species. Many resident victim species live in aggregations, raising the question of whether conspecifics may also benefit from the chasing done by a group member. This would be the case if punishment increases the probability that a blenny switches to a different victim species for its future attacks. We addressed this question on one particular shoaling resident species, scalefin anthias *Pseudanthias squamipinnis*. Individuals of this species frequently chase blennies [9], an observation that cannot be explained with kin selection as a potential mechanism because it has been demonstrated that anthias groups lack any kin structure [21]. Therefore, punishment has to increase the direct fitness of punishers even in this shoaling species. One possible explanation is that blennies select for stable punishment because they discriminate between look-alikes, where one individual punishes and another one does not. We tested this hypothesis experimentally.

Does Aggression Function as Punishment?

To test whether aggression reduces the probability of being selected again as a victim, we let 18 blennies feed off of two differently colored Plexiglas plates: one was invariably retrieved in response to the blenny taking a bite, whereas the other one invariably chased the blenny in response to taking a bite. After training, three naive persons presented

*Correspondence: andrea.bshary@unine.ch

independently of each other both plates simultaneously to the blenny at equal distance and scored what the blenny chose, without any further chasing. Blennies attacked the fleeing plate significantly more often than the chasing plate, no matter whether the chasing had been done by hand or in a more standardized way with a machine (sign test: handheld plate: n = 8 individuals, x = 0, p = 0.008; machine used to standardize chasing: n = 10 individuals, x = 1, p = 0.02).

These data demonstrate that chasing may function as punishment: the apparent reduction in immediate payoffs of both actor (through energetic expenditure) and recipient (through energetic expenditure and risk of injury) yields future benefits to the actor, because the act alters future behavior of the recipient. In our case, the blenny will preferentially seek other victims. Clear demonstrations of punishment have hitherto been rare in animals (but see [22–24]) because it is often difficult to exclude the fact that immediate benefits drive the behavior. In our case, aggression does not yield immediate benefits because blennies almost always bite only once during an attack and retreat even without any reaction of their victim [9].

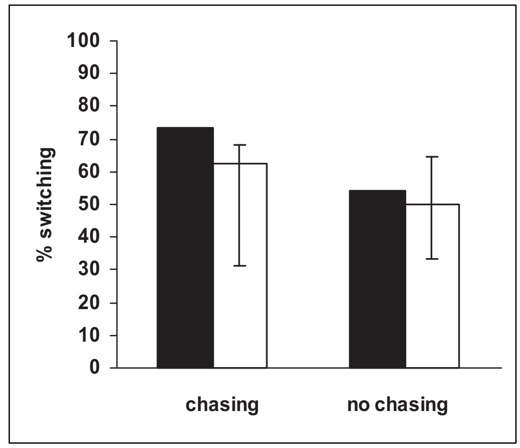

Figure 1. Punishment Creates a Public Good

The probabilities that blennies switch to another victim species for their next attack, depending on whether a female anthias chased or did not chase the blenny in the previous interaction. Black: data on one blenny studied in detail for 16 hr. White: data on 17 blennies, median and interquartiles of individual mean values.

Does Punishment Create a Public Good in Locally Abundant Species?

To test whether chasing the blenny can create a public good in shoaling fishes, we observed 17 blennies in the field. We concentrated our observations on the females of a locally abundant species, scalefin anthias, *P. squamipinnis*, to get a sufficient amount of data of one species. We compared the percentage of the blennies switching to another species if they were aggressively chased with the percentage of switching without any aggressive response of the victim.

Overall, chasing the blenny increased the probability that it switched to another species for its next attack both in a detailed sample on one blenny (X^2 test, n = 256 interactions following a female anthias-blenny interaction, $X^2 = 9.2$, df = 1, p = 0.0024, Figure 1) and when we compared mean switching probabilities of several blennies after being aggressed and after not being aggressed by a female anthias (Wilcoxon signed ranks test, n = 17 blennies, z = −2.059, p = 0.039, Figure 1).

These data suggest that all members of a shoal of look-alikes profit from the chasing done by an individual fish. Punishment by one individual thus appears to create a public good in shoaling species. In line with this assessment, a between-species comparison revealed a slightly negative correlation between local abundance and likelihood of chasing saber-tooth blennies [9]. Nevertheless, the creation of a public good does not lead to a breakdown of punishment, neither in scalefin anthias nor in other resident aggregating species [9]. Therefore, the question arises of what factors may cause that punishment to remain self-serving, even if it additionally creates a public good. One possibility is that the punisher gets a larger proportion of the benefits than other group members [25]. This logic has been proposed in the framework of "group augmentation" [26]. This concept proposes that individuals may benefit from helping other group members because their own reproductive success is linked to the survival of these group members. The concept was developed to explain apparently unconditional individual contributions to vigilance and pup feeding in meercats [27, 28]. Indeed, individuals contribute most to group benefitting activities when they are satiated, and they therefore benefit disproportionally from activities other than foraging. More generally, it has been noted that both by-product mutualism [29], in which an individual shows a self-serving behavior that benefits others as a by-product, and pseudoreciprocity [30], in which an individual's investment causes self-serving responses that benefit the investor as a by-product, may explain cooperation in situations in which more than one individual benefits from helping [31, 32].

Could Blenny Foraging Decisions Select for Stable Punishment Even in Large Groups?

In a laboratory experiment, we tested whether blennies are able to identify free riders. We used four plates (see Figure S1A available online) to present look-alike pairs (mimicking conspecifics) and a different-looking individual (mimicking an allospecific) in a counterbalanced design across individual blennies. Each blenny was confronted with a 100% punishing plate, a look-alike 0% punishing "free riding" plate, and a different-looking plate that punished a foraging bite of the blenny with 50% probability. In each trial, two plates were presented simultaneously, and the blenny was allowed to take a single bite. Depending on its choice, this would either lead to a punishment action followed by removal of both plates or to immediate removal of both plates. The three possible combinations in which two plates could be presented pairwise were counterbalanced across trials. Our construction, to which the lever of the plate was attached, allowed a standardized movement of the plate of fixed direction and distance to punish a blenny (Figure S1B).

Under these conditions, blennies generally preferred the nonpunishing plate over its look-alike 100% punishing plate (Wilcoxon signed rank test, n = 8, Z = −2.38, p = 0.017). Three out of eight blennies met our learning criteria for a significant preference to attack the nonpunishing plate when presented together with the 100% punishing look-alike plate within a maximum of 80 joint presentations (Figure 2). In addition, we found that blennies generally preferred the two look-alike plates over the different-looking plate. They significantly preferred the nonpunishing plate over the 50% punishing plate (Wilcoxon signed rank test, n = 8, Z = −2.52, p = 0.012, Figure 3), and there was a nonsignificant tendency that the

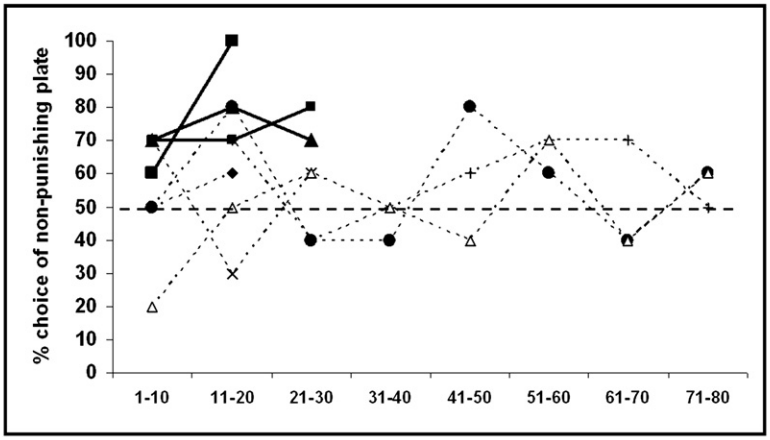

Figure 2. Blennies May Discriminate between Punishers and Free Riders

The probabilities for eight blennies to take a bite from the nonpunishing plate when paired with the look-alike 100% punishing plate over a maximum of eight series of ten trials each. Bold lines: individual blennies that fulfilled our criteria for significant preference for the nonpunishing plate. Dashed line: expected value if blennies do not discriminate between plates. For the experimental setup of experiment 2, see also Figure S1.

blennies preferred the 100% punishing plate over the 50% punishing plate (Wilcoxon signed rank test, n = 8, Z = −1.82 p = 0.069, Figure 3).

The results demonstrate that at least some blennies may be able to learn to discriminate between punishers and nonpunishers and then selectively bite nonpunishers. Their foraging decision rules would then select against nonpunishing individuals and hence select against free riders. In addition, the results suggest that blennies generally prefer more-abundant victim types over rarer ones. In this respect, the blennies act like many predators: it is textbook knowledge that predators may select prey disproportionally to its abundance either because predators develop search images or because they prefer to frequent the habitat of abundant prey species [33]. The blennies' preference for abundant victim types could strongly interfere with their willingness to switch from frequently punishing but abundant species to rarely punishing species that occur at low densities. Nevertheless, we note that there are far more individual fish in the reef than the number of plates we used in our experiment. Thus, the challenge blennies face under natural conditions appears to be more cognitively demanding. On the other hand, it might still be easier for a blenny to distinguish between fish than between plates. In addition, in nature, blennies could simply stay close to a nonpunishing individual for repeated attacks and thereby select against free riders, whereas such a blenny strategy was excluded in our experiment because the plates were removed after each attack.

In summary, our various results demonstrate that aggression by victims functions as punishment because it increases the probability that saber-tooth blennies select alternative victims during future attacks. In this respect, our experiments in which we used Plexiglas plates seem to corroborate well with our field observations. Punishment is likely to produce a public good in shoaling species because blennies are likely to switch to other victim species. In our victim study species, there is no kin structure [21] that could explain why individuals punish, though they provide a public good [7, 8]. More generally, kin structures will rarely exist in reef fish species, because the vast majority of them are open-water spawners with pelagic egg and larval stages, which creates a mixed population structure [34, 35]. Therefore, direct benefits of the observed punishment are needed to explain the persistence of the public good. The reason why contributions to the public good appear to be self-serving is that the common enemy selects for the contribution in our system. Free riders risk being identified as easy targets for future attacks by the blenny, and avoiding this risk may compensate for the costs of chasing, thereby selecting against free riding. Strictly speaking, contributions to the public good are therefore not a case of cooperation, because individuals are selected to punish completely independently of the positive effects their behavior might have on conspecifics [31]. Nevertheless, by-product effects may be important starting points for the evolution of more sophisticated helping behaviors [36].

Our study system differs in a key factor from standard game theoretic experiments on human groups: in each round, the

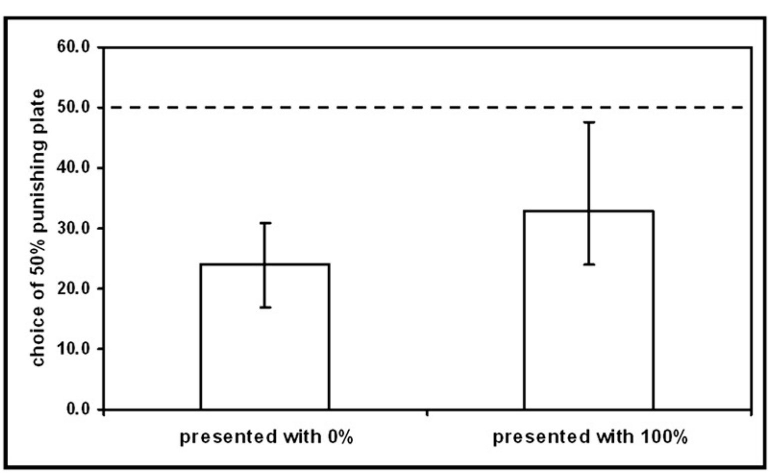

Figure 3. Blennies Prefer Abundant Victim Types

The percentages with which blennies took a bite from the plate that punished with 50% probability when presented with a different-looking plate that either punished with 0% probability or with 100% probability. Median and interquartiles of mean values for eight different blennies are presented. Dashed line: expected value if blennies do not discriminate between plates.

blenny (the cheater) only interacts with a single individual of its own choice, whereas in many human experiments, all group members interact simultaneously [3, 4, 14, 16–18, 37]. As a consequence, it is always clear which individual of a shoal has to punish the blenny for its cheating. A victim cannot expect others to do the punishment, because they did not have a negative experience. We think that such conditions often apply to humans as well, in which punishment is a self-serving response to being cheated while benefiting the community as well. A person whose house gets broken into or who gets attacked by robbers will have to take action (call the police/fight back) even though all the neighbors may profit from this. This self-serving scenario might explain why humans "erroneously" punish free riders in a public goods game in which, in each round, a single individual is randomly selected to obtain the opportunity to punish other group members, even though punishment is costly, because group composition changes every round [38]

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures and one figure and can be found with this article online at doi:10.1016/j.cub.2010.10.027.

Acknowledgments

We thank the Egyptian Environmental Affairs Agency for the permission to conduct research in the park and the park rangers and Ingo Riepl for local support. We further thank the staff of the Lizard Island Research Station and Lexa Grutter for their support and friendship. We are grateful to Aurore Malapert for contributing to the field data collection and to Tom Sherratt and two anonymous referees for comments and corrections of the English. The study was financed by a Natural Environment Research Council fellowship and a grant by the Swiss Science Foundation to R.B.

Received: September 3, 2010
Revised: October 8, 2010
Accepted: October 11, 2010
Published online: November 4, 2010

References

1. Hardin, G. (1968). The tragedy of the commons. Science *162*, 1243–1248.
2. Ostrom, E. (1990). Governing the Commons: The Evolution of Institutions for Collective Action (New York: Cambridge University Press).
3. Fehr, E., and Gächter, S. (2000). Cooperation and punishment in public goods experiments. Am. Econ. Rev. *90*, 980–994.
4. Fehr, E., and Gächter, S. (2002). Altruistic punishment in humans. Nature *415*, 137–140.
5. Boyd, R., Gintis, H., Bowles, S., and Richerson, P.J. (2003). The evolution of altruistic punishment. Proc. Natl. Acad. Sci. USA *100*, 3531–3535.
6. Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2003). Explaining altruistic behavior in humans. Evol. Hum. Behav. *24*, 153–172.
7. Gardner, A., and West, S.A. (2004). Cooperation and punishment, especially in humans. Am. Nat. *164*, 753–764.
8. Lehmann, L., Rousset, F., Roze, D., and Keller, L. (2007). Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. Am. Nat. *170*, 21–36.
9. Bshary, A., and Bshary, R. (2010). Interactions between sabre-tooth blennies and their reef fish victims: Effects of enforced repeated game structure and local abundance on victim aggression. Ethology *116*, 681–690.
10. Clutton-Brock, T.H., and Parker, G.A. (1995). Punishment in animal societies. Nature *373*, 209–216.
11. Alexander, R.D. (1987). The Biology of Moral Systems (New York: Aldine de Gruyter).
12. Nowak, M.A., and Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. Nature *393*, 573–577.
13. Milinski, M., Semmann, D., and Krambeck, H.-J. (2002). Reputation helps solve the 'tragedy of the commons'. Nature *415*, 424–426.
14. Rockenbach, B., and Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. Nature *444*, 718–723.
15. Dreber, A., Rand, D.G., Fudenberg, D., and Nowak, M.A. (2008). Winners don't punish. Nature *452*, 348–351.
16. Gächter, S., Renner, E., and Sefton, M. (2008). The long-run benefits of punishment. Science *322*, 1510–1510.
17. Gächter, S., Herrmann, B., and Thöni, C. (2010). Culture and cooperation. Philos. Trans. R. Soc. Lond. B Biol. Sci. *365*, 2651–2661.
18. Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? J. Public Econ. *92*, 91–112.
19. Carpenter, J.P. (2007). Punishing free-riders: How group size effects mutual monitoring and the provision of public goods. Games Econ. Behav. *60*, 31–51.
20. Smith-Vaniz, W.F. (1976). The Saber-Toothed Blennies, Tribe Nemophini (Pisces: Blenniidae). Monograph 19 (Philadelphia: The Academy of Natural Sciences of Philadelphia).
21. Avise, J.C., and Shapiro, D.Y. (1986). Evaluating kinship of newly settled juveniles within social groups of the coral reef fish *Anthias squamipinnis*. Evolution *40*, 1051–1059.
22. Bshary, R., and Grutter, A.S. (2005). Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. Biol. Lett. *1*, 396–399.
23. Bshary, R., Grutter, A.S., Willener, A.S.T., and Leimar, O. (2008). Pairs of cooperating cleaner fish provide better service quality than singletons. Nature *455*, 964–966.
24. Raihani, N.J., Grutter, A.S., and Bshary, R. (2010). Punishers benefit from third-party punishment in fish. Science *327*, 171.
25. Sherratt, T.N., Roberts, G., and Kassen, R. (2009). Evolutionary stable investment in products that confer both an individual benefit and a public good. Front. Biosci. *1*, 4557–4564.
26. Kokko, H., Johnstone, R.A., and Clutton-Brock, T.H. (2001). The evolution of cooperative breeding through group augmentation. Proc. Biol. Sci. *268*, 187–196.
27. Clutton-Brock, T.H., O'Riain, M.J., Brotherton, P.N.M., Gaynor, D., Kansky, R., Griffin, A.S., and Manser, M. (1999). Selfish sentinels in cooperative mammals. Science *284*, 1640–1644.
28. Clutton-Brock, T.H., Brotherton, P.N.M., O'Riain, M.J., Griffin, A.S., Gaynor, D., Sharpe, L., Kansky, R., Manser, M.B., and McIlrath, G.M. (2000). Individual contributions to babysitting in a cooperative mongoose, *Suricata suricatta*. Proc. Biol. Sci. *267*, 301–305.
29. Brown, J.L. (1983). Cooperation: A biologist's dilemma. In Advances in the Study of Behaviour, J.S. Rosenblatt, ed. (New York: Academic Press), pp. 1–37.
30. Connor, R.C. (1986). Pseudo-reciprocity: Investing in mutualism. Anim. Behav. *34*, 1562–1566.
31. West, S.A., Griffin, A.S., and Gardner, A. (2007). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. J. Evol. Biol. *20*, 415–432.
32. Bshary, R. (2010). Cooperation between unrelated individuals: A game theoretic approach. In Animal Behaviour: Evolution and Mechanisms, P. Kappeler, ed. (Heidelberg, Germany: Springer), pp. 213–240.
33. Krebs, J.R., and Davies, N.B. (1993). An Introduction to Behavioural Ecology, Third Edition (Oxford: Blackwell).
34. Cowen, R.K. (2002). Larval dispersal and retention and consequences for population connectivity. In Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem, P.F. Sale, ed. (London: Academic Press), pp. 149–170.
35. Leis, J.M., and McCormick, M.I. (2002). The biology, behaviour, and ecology of the pelagic, larval stage of coral reef fishes. In Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem, P.F. Sale, ed. (London: Academic Press), pp. 149–170.
36. Leimar, O., and Hammerstein, P. (2010). Cooperation for direct fitness benefits. Philos. Trans. R. Soc. Lond. B Biol. Sci. *365*, 2619–2626.
37. West, S.A., Gardner, A., Shuker, D.M., Reynolds, T., Burton-Chellow, M., Sykes, E.M., Guinnee, M.A., and Griffin, A.S. (2006). Cooperation and the scale of competition in humans. Curr. Biol. *16*, 1103–1106.
38. O'Gorman, R., Henrich, J., and Van Vugt, M. (2009). Constraining free-riding in public goods games: Designated solitary punishers can sustain human cooperation. Proc. Biol. Sci. *276*, 323–329.