



## Research

**Cite this article:** Krams I, Kokko H, Vrublevska J, Āboliņš-Ābols M, Krama T, Rantala MJ. 2013 The excuse principle can maintain cooperation through forgivable defection in the Prisoner's Dilemma game. *Proc R Soc B* 280: 20131475. <http://dx.doi.org/10.1098/rspb.2013.1475>

Received: 7 June 2013

Accepted: 26 June 2013

### Subject Areas:

behaviour, ecology, evolution

### Keywords:

cooperation, reciprocal altruism, tit-for-tat, birds, Prisoner's Dilemma

### Author for correspondence:

Indrikis Krams

e-mail: [indrikis.krams@ut.ee](mailto:indrikis.krams@ut.ee)

# The excuse principle can maintain cooperation through forgivable defection in the Prisoner's Dilemma game

Indrikis Krams<sup>1,3</sup>, Hanna Kokko<sup>2</sup>, Jolanta Vrublevska<sup>3</sup>, Mikus Āboliņš-Ābols<sup>4</sup>, Tatjana Krama<sup>3</sup> and Markus J. Rantala<sup>5</sup>

<sup>1</sup>Institute of Ecology and Earth Sciences, University of Tartu, 51014 Tartu, Estonia

<sup>2</sup>Department of Ecology, Evolution and Genetics, Research School of Biology, Australian National University, Canberra, Australian Capital Territory 0200, Australia

<sup>3</sup>Institute of Systematic Biology, University of Daugavpils, Daugavpils 5401, Latvia

<sup>4</sup>Department of Biology, Indiana University, Bloomington, IN 47405, USA

<sup>5</sup>Department of Biology, University of Turku, 20024 Turku, Finland

Reciprocal altruism describes a situation in which an organism acts in a manner that temporarily reduces its fitness while increasing another organism's fitness, but there is an ultimate fitness benefit based on an expectation that the other organism will act in a similar manner at a later time. It creates the obvious dilemma in which there is always a short-term benefit to cheating, therefore cooperating individuals must avoid being exploited by non-cooperating cheaters. This is achieved by following various decision rules, usually variants of the tit-for-tat (TFT) strategy. The strength of TFT, however, is also its weakness—mistakes in implementation or interpretation of moves, or the inability to cooperate, lead to a permanent breakdown in cooperation. We show that pied flycatchers (*Ficedula hypoleuca*) use a TFT with an embedded 'excuse principle' to forgive the neighbours that were perceived as unable to cooperate during mobbing of predators. The excuse principle dramatically increases the stability of TFT-like behavioural strategies within the Prisoner's Dilemma game.

## 1. Introduction

Cooperation is widespread on many levels of biological organizations, and reciprocity is often suggested as an explanation of cooperation between non-kin. However, misunderstandings of intent, and mistakes in perception and implementation of social actions plague every natural situation. The inherent instability of cooperation between non-relatives is often conceptualized with the aid of the Prisoner's Dilemma [1]. Studies of the evolution of cooperation in iterated Prisoner's Dilemmas have suggested various strategies that might maintain reciprocity in noisy environments [1]. These include tit-for-two-tats (TF2T) which 'forgives' a single mistake, generous tit-for-tat (GTFT) that forgives a defection with a small probability and Pavlov which repeats its prior move after receiving a good pay-off, but changes its behaviour after a poor pay-off [2,3]. Another proposed solution is the contrite tit-for-tat (CTFT [4]), where individuals can be in good or bad standing: player A gets into bad standing if it defects when it should have cooperated. However, if A cooperates in the next round, it can restore its standing, and the pair (A, B) can recover from being stuck in mutual defection caused by a simple error in judgement or implementation [5,6]. Fishman [7,8] considered the case of *involuntary defection* that is conceptually similar to an error in implementing a good intent, but in which case the reason for defection is the *inability* to cooperate given current circumstances. He concluded that the possibility of involuntary defection leads to the evolution of an 'emphatic retaliator' but not of 'sympathetic retaliator'. An emphatic retaliator is a rewording of CTFT, whereas the sympathetic retaliator differs from CTFT in which A refrains from punishing B after B has defected, if A assesses that B's defection was involuntary.

Finding good examples of reciprocity in nature has proved challenging [9–12], yet mobbing shows promise in this area. Mobbing is a form of communal defence by prey individuals that can cause a predator to vacate its immediate foraging area [13–17]. Successful mobbing reduces the threat to nearby prey individuals and allows them to resume their daily activities [18]. Because the success of mobbing depends on the group size [19], reciprocity among neighbouring individuals provides the potential for synergistic benefits of cooperation [20]. Recent experimental studies have shown that mobbing behaviour can be explained in terms of reciprocal altruism, and that animals breeding in a neighbourhood follow a tit-for-tat (TFT)-like strategy when driving away predators [21–24]. Occasionally, the absence of individuals from the neighbourhood that results from the need to rest or collect food coincides with the time when their neighbours need their assistance in mobbing [19]. Mistakenly considering such cases as defection would lead to breakdown in cooperation if flycatchers follow a strict TFT.

We carried out a field experiment to test whether breeding pied flycatchers can discriminate between voluntary and involuntary defection of their neighbours. We assigned breeding pied flycatchers to two treatments: the absent-neighbour and the defecting-neighbour treatments. Each treatment consisted of a focal pair and a neighbouring pair. In both treatments, we captured and temporarily removed the neighbouring pair, during which time a predator was presented at the nest of the focal pair (figure 1). In contrast to our previous study [21], there was no sign of the presence of neighbours in the absent-neighbour treatment, while we simulated the presence of neighbours in the defecting-neighbour group by playing back the neighbours' alarm calls from the vicinity of their nests, such that the focal pair (still present in its territory) could hear these calls. The playbacks ensured that the focal pair had no reason to perceive the reason for the absence of their captured neighbours at mobbing as 'forgivable' (i.e. as true absence). One hour later, we conducted the response trial in which the predator was presented at the nest of the previously absent neighbours and we recorded (i) whether focal pairs joined the mobbing and (ii) the intensity of their mobbing.

## 2. Material and methods

The nest-boxes ( $n = 68$ ) were arranged in pairs and placed 50 m apart. We included in the study nest-boxes containing six nestlings, and the pairs of neighbours for which the difference in age of nestlings in the two nests did not exceed 4 days. The birds were colour marked by contact with a piece of non-waterproof ink saturated foam-rubber located at the entrance of their nest-boxes.

In both treatments, we captured birds breeding in the 'neighbour' nest-box at the end of the nestling phase using mist-nets. We kept these birds in a cage under a hide while a stuffed tawny owl (*Strix aluco*) was presented at the focal nest-box 30 min after the capture of their neighbours. It was mounted near the nest-box and placed between the two neighbouring nest-boxes when no flycatchers were detectable in the vicinity. We presented the owl for 15 min after which we removed it. In the defecting-neighbour treatment, we played back neighbour alarm calls that were previously recorded with a Sony PCM-D50 recorder, because pied flycatchers always give alarm calls when they detect a potential predator in the vicinity of their nests. As soon as the captured birds were released, they resumed feeding their nestlings within 10 min. We observed and evaluated the behaviour of flycatchers



**Figure 1.** A male pied flycatcher mobs a stuffed tawny owl. The mobbing flycatcher appears as a shadow. (Online version in colour.)

from a hide where we remained for the hour separating the first and the second trials to ensure that the birds were not involved in anti-predator activities towards us.

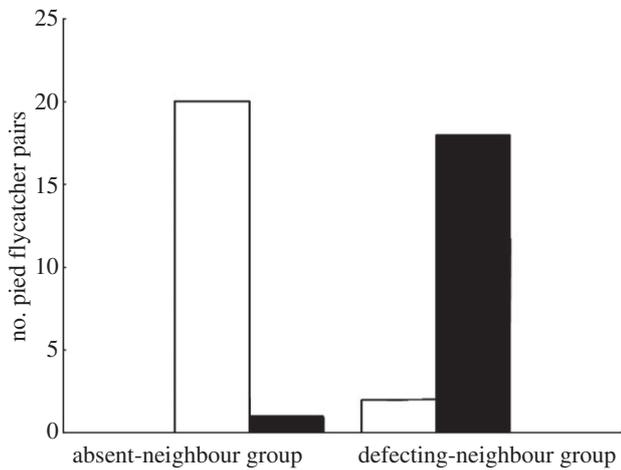
One hour later, we started the response trial and presented the owl for 15 min at the nest of the neighbours. We considered the focal pair to have assisted their neighbours if they approached the owl closer than half of the distance between the two nest-boxes. Our scale of flycatcher mobbing responses consisted of four categories of displays and vocalizations ranging from 'no response' to the predator (0 points) to 'strong response' (3 points) [21].

## 3. Results

Nest owners (the focal pair in the first trial, and their neighbouring pair in the response trial) always mobbed the owl. However, not all of the focal pairs cooperated during the response treatment. Significantly more focal pairs of pied flycatchers arrived to mob the predator near the nest-boxes of their neighbours during the response trial in the absent-neighbour treatment (20 out of 21 focal pairs) than in the defecting-neighbour treatment (two out of 20; two-tailed, sign-test,  $p < 0.001$ ; figure 2). The mobbing intensity of focal pairs at their own nests was not significantly different between treatments (two-tailed, Kruskal–Wallis test,  $\chi^2_3 = 1.59$ ,  $p = 0.66$ ), but in the absent-neighbour treatment, the focal pairs' effort at mobbing was significantly higher when they mobbed at their own nest (first trial) than when they mobbed to assist the neighbours (second trial; two-tailed, Wilcoxon matched-pairs signed-ranks test,  $Z = -2.83$ ,  $n = 21$ ,  $p = 0.005$ ).

## 4. Discussion

Our results show that breeding pied flycatchers did not follow a strict TFT, TF2T or Pavlov strategy. Although birds almost invariably punished a defection, they did not simply



**Figure 2.** Number of pied flycatcher pairs that assisted their neighbours (unfilled bars) and those that remained at their own nests (filled bars) during the second trial in the absent-neighbour group and the defecting-neighbour group.

copy the previous move of their neighbours, which excludes a strict TFT. The birds did not follow a TF2T because neighbours in the defecting-neighbour treatment were not forgiven. Furthermore, the birds did not follow a Pavlov strategy because they did not change their behaviour to defecting in the absent-neighbour group after receiving a poor pay-off. Although the focal pair was not assisted by their neighbours in either treatment, they only retaliated when the neighbours were in the treatment that simulated true defection. We define the *excuse principle* as the decision rule employed by a TFT-like strategy to forgive defection under conditions that make it ‘understandable’. Our results show pied flycatchers use this principle to discriminate between defection and the inability of their neighbours to cooperate. They thus followed the sympathetic retaliator strategy.

The sympathetic retaliator is not the only strategy that uses an excuse principle to forgive defectors. In CTFT, an involuntary or ‘mistaken’ defection by A is followed by a defecting B. In the next move, A excuses B and cooperates, thus reinstating mutual cooperation. Although our experimental design did not follow more than two ‘moves’ of the game, our experiments are sufficient to conclude that CTFT was not used by flycatchers—in our study system, B refrained from defecting if we had manipulated A’s defection to appear involuntary. Thus, the only type of TFT consistent with our findings is the sympathetic retaliator of Fishman [8]. The excuse principle sustains cooperation under some degree of uncertainty: it is not possible to assign the behaviour of the social partner to

the non-cooperative category with absolute accuracy. Our finding of a few cases where the ‘mistaken’ form of behaviour was found fits this idea well.

Fishman [8] concludes, however, that sympathetic retaliators should be rare in nature for two reasons. First, he argues that this particular form of retaliation is cognitively more demanding than its alternatives. Second, he suggests that the stability of this strategy is undermined by a class of individuals that are never able to cooperate. Success of the sympathetic retaliator in this case is therefore equal to that of an always-altruistic individual in a population of defectors, and thus will be selected against. Our empirical example shows that neither caveat is necessarily a problem in nature. The sympathetic retaliator B needs relatively reliable information about A’s circumstances, but it does not need to keep a memory of its own past behaviour, which simplifies the cognitive task. Pied flycatchers appear to possess the cognitive ability to assess whether a neighbour should be excused for its behaviour. For obvious reasons, flycatchers have to be near their nest for much of the time, and if neighbour presence is easy to assess, the cognitive demands are not necessarily great. The same fact that parent birds must spend much of their time near the nest also makes it difficult to argue that there could be a subclass of individuals that are permanently unable to cooperate (or try to cheat by claiming that they are unable). Moreover, if some individuals are permanently unable to help, and this consistent inability is visible to others, then these others will be able to exclude unhelpful subclasses from representing potential partners in reciprocity games. The fact that individual recognition has been demonstrated in this [25] and other bird species [26] makes this a biologically plausible scenario and suggests that maintenance of reciprocity is achieved through partner choice.

Without an excuse principle, TFT cannot persist, as its appetite for revenge proves self-destructive [27,28]: it requires defection of a social partner even if the partner is unable to cooperate [5,6]. The excuse principle makes mandatory defection unnecessary, dramatically increasing the stability of TFT-like behavioural strategies [29,30].

All animal manipulations comply with the current laws in Latvia.

**Funding statement.** The study was supported by the Academy of Finland (to I.K. and M.J.R.) and by the Ministry of Education and Science of the Republic of Latvia and the Science Council of Latvia (to I.K. and T.K.). Social Fund within the project ‘Support for the implementation of doctoral studies at Daugavpils University’ No. 2009/0140/1DP/1.1.2.1.2/ 09/IPIA/VIAA/015 supported J.V.

## References

- Nowak MA. 2006 Five rules for the evolution of cooperation. *Science* **314**, 1560–1563. (doi:10.1126/science.1133755)
- Brandt H, Sigmund K. 2006 The good, the bad and the discriminator: errors in direct and indirect reciprocity. *J. Theor. Biol.* **239**, 183–189. (doi:10.1016/j.jtbi.2005.08.045)
- Wu JZ, Axelrod R. 1995 How to cope with noise in the iterated Prisoners-Dilemma. *J. Confl. Resol.* **39**, 83–189. (doi:10.1177/0022002795039001008)
- Boyd R. 1989 Mistakes allow evolutionary stability in the repeated Prisoner’s Dilemma game. *J. Theor. Biol.* **136**, 47–56. (doi:10.1016/S0022-5193(89) 80188-2)
- Lotem A, Fishman MA, Stone L. 1999 Evolution of cooperation between individuals. *Nature* **400**, 226–227. (doi:10.1038/22247)
- Sherratt TN, Roberts G. 2001 The importance of phenotypic defectors in stabilizing reciprocal altruism. *Behav. Ecol.* **12**, 313–317. (doi:10.1093/beheco/12.3.313)
- Fishman MA. 2003 Indirect reciprocity among imperfect individuals. *J. Theor. Biol.* **225**, 285–292. (doi:10.1016/S0022-5193(03)00246-7)
- Fishman MA. 2006 Involuntary defection and the evolutionary origins of empathy. *J. Theor. Biol.* **242**, 873–879. (doi:10.1016/j.jtbi.2006.05.004)
- Dugatkin LA. 1988 Do guppies play tit for tat during predator inspection visits? *Behav. Ecol. Sociobiol.* **23**, 395–399. (doi:10.1007/BF00303714)

10. Raihani NJ, Bshary R. 2011 Resolving the iterated Prisoner's Dilemma: theory and reality. *J. Evol. Biol.* **24**, 1628–1639. (doi:10.1111/j.1420-9101.2011.02307.x)
11. Milinski M. 1987 Tit for tat in sticklebacks and the evolution of cooperation. *Nature* **325**, 433–435. (doi:10.1038/325433a0)
12. Clutton-Brock T. 2009 Cooperation between non-kin in animal societies. *Nature* **462**, 51–57. (doi:10.1038/nature08366)
13. Cheney DL, Moscovice LR, Heesen M, Mundry R, Seyfarth RM. 2010 Contingent cooperation between wild female baboons. *Proc. Natl Acad. Sci. USA* **107**, 9562–9566. (doi:10.1073/pnas.1001862107)
14. Davies NB, Welbergen JA. 2009 Social transmission of a host defense against cuckoo parasitism. *Science* **324**, 1318–1320. (doi:10.1126/science.1172227)
15. Griesser M. 2009 Mobbing calls signal predator category in a kin group-living bird species. *Proc. R. Soc. B* **276**, 2887–2892. (doi:10.1098/rspb.2009.0551)
16. Templeton CN, Greene E, Davis K. 2005 Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science* **308**, 1934–1937. (doi:10.1126/science.1108841)
17. Welbergen J, Davies N. 2009 Strategic variation in mobbing as a front line of defense against brood parasitism. *Curr. Biol.* **19**, 235–240. (doi:10.1016/j.cub.2008.12.041)
18. Flasskamp A. 1994 The adaptive significance of avian mobbing. V. An experimental test of the 'Move On' hypothesis. *Ethology* **96**, 322–333. (doi:10.1111/j.1439-0310.1994.tb01020.x)
19. Krams I, Berzins A, Krama T. 2009 Group effect in nest defence behaviour of breeding pied flycatchers, *Ficedula hypoleuca*. *Anim. Behav.* **77**, 513–517. (doi:10.1016/j.anbehav.2008.11.007)
20. Kruuk H. 1964 Predators and anti-predator behaviour of the blackheaded gull (*Larus ridibundus* L.). *Behaviour* **11**, 1–129.
21. Krams I, Krama T, Igaune K, Mand R. 2008 Experimental evidence of reciprocal altruism in the pied flycatcher. *Behav. Ecol. Sociobiol.* **62**, 599–605. (doi:10.1007/s00265-007-0484-1)
22. Olendorf R, Getty T, Scribner K. 2004 Cooperative nest defence in red-winged blackbirds: reciprocal altruism, kinship or by-product mutualism? *Proc. R. Soc. Lond. B* **271**, 177–182. (doi:10.1098/rspb.2003.2586)
23. Krams I, Bērziņš A, Krama T, Wheatcroft D, Igaune K, Rantala MJ. 2010 The increased risk of predation enhances cooperation. *Proc. R. Soc. B* **277**, 513–518. (doi:10.1098/rspb.2009.1614)
24. Krama T, Vrublevska J, Freeberg TM, Kullberg C, Rantala MJ, Krams I. 2012 You mob my owl, I'll mob yours: birds play tit-for-tat game. *Sci. Rep.* **2**, 800. (doi:10.1038/srep00800)
25. Lampe HM, Slagsvold T. 1998 Female pied flycatchers respond differently to songs of mates, neighbours and strangers. *Behaviour* **135**, 269–285. (doi:10.1163/156853998793066249)
26. Falls JB. 1982 Individual recognition by sounds in birds. In *Acoustic communication in birds* (eds DE Kroodsmas, EH Miller), pp. 237–278. New York, NY: Academic Press.
27. Nowak MA. 2008 Generosity: a winner's advice. *Nature* **456**, 579. (doi:10.1038/456579a)
28. Rand DG, Ohtsuki H, Nowak MA. 2009 Direct reciprocity with costly punishment: generous tit-for-tat prevails. *J. Theor. Biol.* **256**, 45–57. (doi:10.1016/j.jtbi.2008.09.015)
29. Axelrod R, Hamilton WD. 1981 The evolution of cooperation. *Science* **211**, 1390–1396. (doi:10.1126/science.7466396)
30. Trivers RL. 1971 Evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)