Upstream reciprocity and the evolution of gratitude

Martin A. Nowak1,2,* and Sébastien Roch3

1Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, and 2Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
3Department of Statistics, University of California, Berkeley, CA 94720, USA

If someone is nice to you, you feel good and may be inclined to be nice to somebody else. This everyday experience is borne out by experimental games: the recipients of an act of kindness are more likely to help in turn, even if the person who benefits from their generosity is somebody else. This behaviour, which has been called ‘upstream reciprocity’, appears to be a misdirected act of gratitude: you help somebody because somebody else has helped you. Does this make any sense from an evolutionary or a game theoretic perspective? In this paper, we show that upstream reciprocity alone does not lead to the evolution of cooperation, but it can evolve and increase the level of cooperation if it is linked to either direct or spatial reciprocity. We calculate the random walks of altruistic acts that are induced by upstream reciprocity. Our analysis shows that gratitude and other positive emotions, which increase the willingness to help others, can evolve in the competitive world of natural selection.

Keywords: evolution of cooperation; upstream indirect reciprocity; direct reciprocity; spatial reciprocity; random walks on graphs

1. INTRODUCTION

The evolutionary analysis of altruistic behaviour began with kin selection (Hamilton 1967; Queller 1992; Frank 1998; West et al. 2002; Foster et al. 2006) and group selection (Levin & Kilmer 1974; Wilson 1975; Keller 1999; Michod 1999; Boyd et al. 2003; Traulsen & Nowak 2006). Direct reciprocity is the idea that cooperation emerges in repeated encounters between the same two individuals according to the principle ‘I help you and you help me’ (Trivers 1971; Axelrod & Hamilton 1981; Axelrod 1984; Fudenberg & Maskin 1990; Nowak & Sigmund 1992; Imhof et al. 2005). In contrast, indirect reciprocity embodies the concept ‘I help you and somebody else helps me’ (Alexander 1987; Nowak & Sigmund 1998; Nowak & Sigmund 2005). Indirect reciprocity comes in two flavours. Downstream indirect reciprocity means that a person who has helped in the past has a higher chance of receiving help (Nowak & Sigmund 1998; Wedekind & Milinski 2000; Engelman & Fischbacher 2002; Milinski et al. 2002; Fishman 2003; Brandt & Sigmund 2004; Ohtsuki & Iwasa 2004, 2005; Brandt & Sigmund 2005; Chalub et al. 2005). Upstream indirect reciprocity means that a person who has just received help has an (unreasonable) urge to help someone too. Although observed in experimental settings (Dufwenberg et al. 2001; Güth et al. 2001; Greiner & Levati 2003; Bartlett & DeSteno 2006), upstream reciprocity is harder to understand from an evolutionary perspective (Boyd & Richerson 1989; Pfeiffer et al. 2005). Here, we assume that the recipient of an altruistic act experiences gratitude and is more likely to help either the donor or another person.

Therefore, gratitude is an emotion which can lead to upstream reciprocity. We show that upstream reciprocity alone does not allow evolution of cooperation, but it can evolve if it is linked to a mechanism for the evolution of cooperation. In particular, we study the links to direct reciprocity and spatial reciprocity (Nowak & May 1992; Nowak 2006; Ohtsuki & Nowak 2006a,b; Ohtsuki et al. 2006; Santos et al. 2006).

2. RANDOM WALKS OF UPSTREAM RECIPROCITY

Each altruistic act involves a cost, c, for the donor and a benefit, b, for the recipient. We assume that $b > c$, otherwise cooperation cannot lead to an overall benefit. Consider a large population of players in a game of upstream reciprocity. One player initiates a chain of altruism. The second player passes it on, with a certain probability, to a third player, who in turn might help a fourth player and so on. After a number of steps, the chain dies out. Upstream reciprocity leads to random walks on the set of all players (figure 1).

In the simplest formulation of the game, each player is characterized by two parameters: $q$ denotes the probability to initiate and $p$ the probability to pass on. If everybody in the population uses strategy $S(p, q)$, then the average length of an altruistic random walk is $1/(1-p)$. Each player initiates random walks at the rate $q$. Therefore, the ‘cooperativity’ of a strategy is given by $s = q(1-p)$, which is the expected number of secondary altruistic acts induced by a single player per time-step. In the limit $p \rightarrow 1$, the cooperativity tends to infinity (if $q > 0$), because any one initiation generates a line of altruism which is never to break.

Defectors, $S(0,0)$, never initiate and never pass on. Classical cooperators, $S(0,1)$, are spontaneous altruists: they initiate, but do not pass on. They are cooperative, but not grateful. In contrast, the strategy $S(1,0)$ never initiates,
but always passes on, and the strategy $S(1, 1)$ always initiates and always passes on. These four strategies represent the corner points of our strategy space which is given by the unit square, $[0,1]^2$. Each point in this square represents a pure, probabilistic strategy. Strategies with $p > 0$ are gratitude-related altruists: they have a positive probability to help someone if they have just received help themselves.

Let us now consider a population of players using different strategies of upstream reciprocity. We can show that natural selection always reduces the cooperativity, $s$. A strategy with lower cooperativity always outcompetes a strategy with higher cooperativity. Defectors dominate all strategies with positive cooperativity, $s > 0$. We conclude that upstream reciprocity alone does not facilitate the evolution of cooperation.

3. UPSTREAM RECIPROCITY CAN EVOLVE AS A BY-PRODUCT OF DIRECT RECIPROCITY

The situation changes dramatically if we allow some level of direct reciprocity. Denote by $r$ the probability that help is immediately returned to the donor. Thus, a recipient reciprocates to the donor with probability $r$. From the simplified payoff function, we can see that two strategies are best replies to themselves. Finally, if both $r_1 > r_2$ and $s_1 < s_2$, then there is stable equilibrium between the two strategies. If $r_1 > c/b > r_2$ and $s_1 > s_2$, then both strategies are best replies to themselves. Finally, if both $r_1$ and $r_2$ are less than $c/b$, then selection favours the strategy with lower cooperativity.

From the simplified payoff function, we can see immediately that a strategy, $S(p, q, r)$, is evolutionarily stable against invasion by another strategy with lower cooperativity if $r > c/b$. The probability to reflect the random walk to the donor, $r$, has to exceed the cost-to-benefit ratio, $c/b$, of the altruistic act. In this case, direct reciprocity allows the evolution of cooperation, and upstream reciprocity can hitch-hike on direct reciprocity. The interpretation of this finding is: if there is direct reciprocity in a population, then upstream reciprocity will evolve too (figure 2).

If $r > c/b$, then selection tends to increase all of the three parameters $p$, $q$, and $r$, and evolution will lead to more upstream reciprocity, more initiation of random walks and more direct reciprocity. Adaptive dynamics (Nowak & Sigmund 1990; Dieckmann et al. 1995; Metz et al. 1996; Hofbauer & Sigmund 2003; Doebeli et al. 2004) are described in Appendix A. If the parameter $r$ evolves very close to 1, then upstream reciprocity becomes less relevant. For many social interactions, however, it is natural to assume an upper bound $r_{max}$: a direct
Figure 2. Upstream reciprocity evolves if there is direct reciprocity. (a) Suppose there is a certain probability, $r$, that the recipient of an altruistic act will immediately reciprocate to the donor. With probability $(1-r)p$ the recipient will help someone else. The probabilities $p$ and $r$ denote the levels of upstream and direct reciprocity, respectively. If $r$ exceeds the cost-to-benefit ratio of the altruistic act, $r > c/b$, then direct reciprocity allows the evolution of cooperation. But in this case, upstream reciprocity will evolve too, because selection will tend to increase the parameter $p$. (b) In order to calculate the payoff for one strategy (light blue) versus another (dark blue), we consider a large mixed population and average over all possible random walks initiated by each player (see the electronic supplementary material). Three random walks are shown (red, green and yellow arrows).

Reciprocity is not possible or meaningful if there is no opportunity or need to help the donor immediately. In this case, there is a continuous selection pressure leading to more and more upstream reciprocity ($p \rightarrow 1$).

The parameter $q$ characterizes spontaneous altruism, while $p$ denotes the gratitude-related altruism of upstream reciprocity. The total amount of altruism in a population increases linearly with $q$, but hyperbolically with $p$. Therefore, gratitude-related altruism can lead to much higher levels of cooperativity. Moreover, if $r > c/b$, then spontaneous cooperators (with $p = 0$) are outcompeted by gratitude-related cooperators ($p > 0$).

4. THE SYNERGISTIC INTERACTION BETWEEN UPSTREAM AND SPATIAL RECIPROCITY

As a second model, we investigate the interaction between upstream and spatial reciprocity. Imagine a cluster of cooperators surrounded by defectors. Altruistic acts are initiated by cooperators and lead to random walks of cooperation that die within the cluster or after hitting a defector. We need to calculate the expected payoff of all individuals by summing over all possible random walks. For the evolutionary dynamics, let us consider ‘imitation updating’ (Ohtsuki et al. 2006): at each time-step, an individual is chosen at random to update its strategy. It will compare its own payoff to the payoff of its immediate neighbours and then choose one of the strategies proportional to fitness. We can perform an exact mathematical analysis if we study the simplest possible geometry: all individuals are aligned in a one-dimensional array (Ellison 1993; Nakamaru et al. 1997; Lieberman et al. 2005; Ohtsuki & Nowak 2006a). The details of the calculation are shown in the electronic supplementary material. Here, we only state the results.

As before, we consider strategies $S(p, q)$, where $p$ denotes the probability to pass on and $q$ the probability to initiate. It turns out that spontaneous altruists, $S(0, 1)$, can outcompete defectors, $S(0, 0)$, if $b/c > 4$. In contrast, gratitude-related altruists, $S(1, 1)$, can outcompete defectors if $b/c > 10/7$. Therefore, upstream reciprocity working together with spatial reciprocity makes evolution of cooperation easier (figure 3).

More precisely, the game is determined by the following two functions:

$h(p) = \frac{8 + 2p + 8\sqrt{1 - p^2}}{3 + 4p + \sqrt{1 - p^2}}$

and

$g(p) = \frac{p}{1 + 2p} \frac{3 + 3p + \sqrt{1 - p^2}}{1 + p - \sqrt{1 - p^2}}$

If $b/c > h(p)$, then upstream cooperators, $S(p, 1)$, win against both the defectors, $S(0, 0)$, and the classical cooperators, $S(0, 1)$. If $b/c > g(p)$, then natural selection favours small increases in $p$ and $q$, thereby enhancing the cooperativity, $s = \min(1 - p)$. Note that $h(0) = g(0) = 4$, while $h(1) = 10/7$ and $g(1) = 1$. Moreover, we have $h(p) \geq g(p)$ for all probabilities $p$. All these results are derived in the electronic supplementary material.

5. CONCLUSION

In summary, we conclude that upstream reciprocity alone does not select for cooperation, but can promote cooperation if it is linked to a mechanism for the evolution of cooperation. Here, we have studied the links to direct and spatial (or network) reciprocity. If direct reciprocity allows cooperation ($r > c/b$), then upstream reciprocity will evolve too. In the model with spatial reciprocity, we find that evolution of cooperation is easier with upstream reciprocity ($b/c > 10/7$) than without upstream reciprocity ($b/c > 4$). In both models, spontaneous cooperators are outcompeted by gratitude-related upstream cooperators. Moreover, upstream reciprocity greatly enhances the level of altruism in a population.
Upstream reciprocity and gratitude

Our models are psychologically plausible. Experiments with human subjects demonstrate that gratitude, which is the positive emotion one feels after having received something of value, fosters prosocial behaviour (Carlson et al. 1988; Emmons & McCullough 2004; Bartlett & DeSteno 2006): the recipient of a favour is more likely to help both the donor (direct reciprocity) and a stranger (upstream reciprocity). Thus, gratitude may be the key to understanding upstream reciprocity. Our analysis demonstrates that gratitude and other positive emotions, which enhance the willingness to help (given by the parameter p in our model), can evolve by natural selection. It is also conceivable that upstream reciprocity is a by-product of emotions that have primarily evolved to facilitate direct reciprocity.

This paper is a first step towards studying the enormous consequences of upstream reciprocity for human and animal behaviour. We expect that upstream reciprocity will also work in synergy with kin selection, group selection and downstream indirect reciprocity, which is based on reputation. It will be of particular interest to study upstream reciprocity as a random walk on general social networks. Another promising extension is the idea that a person, who has just received help, may not only help one other person (as considered here) but also several other people. This can lead to an ‘epidemiology of altruism’ resulting in an explosive increase of altruistic acts. For a change, this is a pandemic which would be welcomed by all of us.

Support from the John Templeton Foundation, the NSF/NIH joint program in Mathematical Biology (NIH grant 1R01GM078986-01), NSERC (Canada) and FQRNT (Quebec, Canada) is gratefully acknowledged. The Program for Evolutionary Dynamics at Harvard University is sponsored by Jeffrey Epstein.

APPENDIX A
(a) Adaptive dynamics
The idea of adaptive dynamics is the following (Nowak & Sigmund 1990). Consider an infinitely large homogeneous population using strategy \( S_1 = S(p_1, q_1, r_1) \). A mutant strategy \( S_2 = S(p_2, q_2, r_2) \) is generated, which is very close to \( S_1 \) in the space of strategies. In our context, this means that \( p_2 \) is very close to \( p_1 \), \( q_2 \) is very close to \( q_1 \), and \( r_2 \) is very close to \( r_1 \). If the payoff for the mutant, \( A(S_2, S_1) \), is less than the payoff for the resident, \( A(S_1, S_1) \), then the mutant cannot invade and becomes extinct again. If however \( A(S_2, S_1) > A(S_1, S_1) \), then the mutant can invade. If additionally \( A(S_1, S_2) < A(S_2, S_2) \), then the resident becomes extinct. The population has moved from \( S_1 \) to \( S_2 \). Now another mutant \( S_3 \) is generated and so on. In the limit of the mutant strategies being infinitesimally close to the resident, the evolutionary trajectories of adaptive dynamics can be described by ordinary differential equations. Denote by \( S_i \) the invading strategy and by \( S_j \) the resident. From \( A(S_0, S_i) = S_i, S_j \) we obtain

\[
\frac{dp}{dt} = \frac{\partial A(S_i, S_j)}{\partial p_i} \bigg|_{S_i=S_j} = \frac{K}{1-p},
\]

\[
\frac{dq}{dt} = \frac{\partial A(S_i, S_j)}{\partial q_i} \bigg|_{S_i=S_j} = \frac{K}{q},
\]

\[
\frac{dr}{dt} = \frac{\partial A(S_i, S_j)}{\partial r_i} \bigg|_{S_i=S_j} = \frac{K}{1-r}.
\]

where

\[ K = \frac{(br - c)q}{(1-r)^2(1-p)}. \]  

References


NOTICE OF CORRECTION

The corresponding author’s affiliation is now correct. 21 December 2006