Detrimental effects of sanctions on human altruism

Ernst Fehr* & Bettina Rockenbach†

* *University of Zürich, Institute for Empirical Research in Economics, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland*
† *University of Erfurt, Nordhäuser Straße 63, D-99089 Erfurt, Germany*

The existence of cooperation and social order among genetically unrelated individuals is a fundamental problem in the behavioural sciences. The prevailing approaches in biology and economics view cooperation exclusively as self-interested behaviour—unrelated individuals cooperate only if they face economic rewards or sanctions rendering cooperation a self-interested choice. Whether economic incentives are perceived as just or legitimate does not matter in these theories. Fairness-based altruism is, however, a powerful source of human cooperation. Here we show experimentally that the prevailing self-interest approach has serious shortcomings because it overlooks negative effects of sanctions on human altruism. Sanctions revealing selfish or greedy intentions destroy altruistic cooperation almost completely, whereas sanctions perceived as fair leave altruism intact. These findings challenge proximate and ultimate theories of human cooperation that neglect the distinction between fair and unfair sanctions, and they are probably relevant in all domains in which voluntary compliance matters—in relations between spouses, in the education of children, in business relations and organizations as well as in markets.

Human societies are characterized by an unprecedented division of labour supported by a myriad of social and economic exchanges and compliance with social norms[1,2]. Norm compliance and the exchange of favours, goods and services pervade every aspect of human life. They shape the interactions in families, neighbourhoods, schools, firms, markets and politics. The crucial feature of any exchange is that the parties involved have to trust each other. If the exchange partners had doubts about the other's reliability most exchanges would not take place. In fact, much of the economic backwardness in the world can plausibly be explained by a lack of mutual confidence inhibiting cooperation in the production and the exchange of goods and services[3,4]. Yet, what ensures that exchange partners trust each other? In modern societies one solution is to conclude a legally enforceable contract that regulates all aspects of the exchange—in particular, the sanctions imposed on those who breach the contract. If the punishment for breaching the contract is large enough, it is in the self-interest of the involved parties to fulfil their obligations because otherwise they will be heavily punished.

For millennia, however, humans have not been able to rely on contracts because there were no impartial courts enforcing voluntary agreements. Even today it is typically not possible to regulate transactions in every detail, and frequently the courts are unable to verify who violated an agreement. Therefore, even in modern societies the overwhelming majority of all social and economic exchanges are not based on complete and legally enforceable contracts but on implicit agreements and social norms lacking explicitly specified sanctions for non-compliance[5,6]. In such situations there are ample opportunities for cheating the exchange partner to one's own advantage. It is the ubiquity of such cheating opportunities that renders altruistic cooperation important. Altruistic cooperators are willing to cooperate, that is, to abide by the implicit agreement, although cheating would be economically beneficial for them. Much recent research indicates that altruistic cooperation is an important behavioural force[7–11].

However, the prevailing theoretical approaches in biology[12–17] and economics[18,19] neglect costly altruistic behaviour that does not yield future economic benefits for the altruist. Individuals are supposed to do others a favour only if they can themselves reap direct[12,13] or indirect[14–17] future benefits from such others. Likewise, individuals are supposed to cheat on agreements unless they are sufficiently rewarded for compliance or sanctioned for non-compliance. This axiom of self-interested behaviour implies that the only remedy for the compliance problem lies in the provision of sufficient rewards or sanctions. Yet, if these incentives for cooperation undermine altruistic behaviour the remedy may do more harm than good because it aggravates the compliance problem[5,20–23]. Influential social scientists have hypothesized that economic incentives might undermine altruism[20,21] but so far it has been impossible to provide compelling evidence from uncontrolled field studies.

Capturing trust and altruistic cooperation

We examine the question of how sanctions intended to prevent cheating affect human altruism in a sequentially played social dilemma experiment with real monetary stakes. In each trial of the experiment two mutually anonymous subjects are involved—one subject in the role of investor, the other one in the role of trustee. First, the investor has the chance of choosing a costly trusting action. Then the trustee is informed about the investor's action and he can honour the investor's trust by taking a costly cooperative action. The payoff rules of the experiment ensure that if the investor chooses a trusting action and the trustee responds cooperatively both players increase their monetary payoff. However, the trustee also has the option of not honouring the investor's trust. In this case the trustee saves the costs of cooperation but the investor is worse off than if he had not been trustful. Because a cooperative action is costly for the trustee the trustee is tempted not to honour the investor's trust. Yet, if the investor expects that the trustee will not honour his trust the investor will not choose a costly trusting action in the first place. Thus, the subjects are caught in a dilemma because trust and cooperation is beneficial for both subjects but the trustee faces the temptation not to cooperate and, therefore, the investor is tempted not to trust.

To examine potential detrimental effects of sanctions on human altruism we implement two experimental conditions—an incentive condition and a trust condition. In the trust condition, cooperation of the trustee cannot be achieved by incentives because there are no rewards or sanctions available to the investor. The trustee's cooperation in this condition provides, therefore, a baseline measure of altruism because—in the absence of economic incentives—purely selfish trustees will never cooperate. In the incentive

condition, the investor has the option of fining a trustee who has not cooperated sufficiently. The incentive condition informs us, therefore, whether there are detrimental effects of sanctions on the trustee's altruism. In particular, if the amount of cooperation is higher in the absence of the sanction that would be evidence of a detrimental effect.

The details of the trust condition are as follows. Both the investor and the trustee receive an endowment of ten money units (MUs). The investor has the option to trust his partner by sending between zero and ten MUs of his endowment to the trustee. The experimenter then triples any amount sent so that if the investor sends, say, 10 MUs the trustee receives 30 MUs. The tripling of the investment mimics a situation where the trustee has superior productive opportunities for the use of economic resources. If the investor transfers money to the trustee he also has to specify a 'desired back-transfer' that can be any amount between zero and the tripled transfer. In the previous example, for instance, where the investor sent 10 and the trustee received 30 MUs, the desired back-transfer can be any amount between zero and 30 MUs. Once the investor has made his investment the trustee is informed about the amount sent and the desired back-transfer. Then the trustee has the option to cooperate by choosing the actual level of the back-transfer. The trustee is free to send any amount between zero and the tripled transfer—in our example, up to 30 MUs—back to the investor. The back-transfer is not tripled by the experimenter—if the trustee sends back, say, 8 MUs the investor receives exactly 8 MUs.

The final payoff of the investor is given by his endowment of 10 MUs minus the transfer to the trustee plus the actual back-transfer of the trustee. The payoff of the trustee is given by the endowment of 10 MUs plus the tripled transfer minus the actual back-transfer. Thus, if there is no trust and cooperation the investor's transfer and the trustee's back-transfer are zero. In this case, both earn just their endowment of 10 MUs. This is the predicted outcome if both subjects are fully selfish and the investor anticipates the trustee's selfishness. In the trust condition, a selfish trustee will never pay back anything and, therefore, a selfish investor, who anticipates the trustee's behaviour, will transfer nothing. Yet, trust and cooperation can render both subjects better off. If the investor trusts fully and sends 10 MUs and the trustee's actual back-transfer is 20 MUs the investor earns 20 MUs, and the trustee earns his endowment of 10 plus the tripled transfer of 30 minus the back-transfer of 20, which in total also gives 20 MUs.

The incentive condition is exactly identical to the trust condition except for one feature: the investor can impose a fine of four MUs on the trustee if less than the desired amount has been sent back. Yet, the investor can also refrain from imposing the fine. The trustee need not be fined even if he pays back less than the desired amount. At the time when the investor chooses the transfer and the desired back-transfer he also has to specify whether to impose a fine for the case that the trustee's actual back-transfer is lower than the desired one. Thus, when the trustee makes his decision in the incentive condition he knows whether the investor has chosen the fining option. To avoid evocative language the experimental instructions do not include value-laden terms like fine or punishment. Instead, the fine is described, more neutrally, as a deduction from the trustee's payoff. The fine affects only the trustee's payoff by reducing his earnings by 4 MUs in the case of an insufficient back-transfer. The fine does not affect the investor's payoff directly. However, the investor can use the fine to affect the trustee's back-transfer. In the incentive condition, a selfish trustee who does not face a fine will never pay back anything. Only if he faces a fine will he be ready to pay back up to 4 MUs. For instance, if the desired back-transfer is 3 MUs the trustee is better off by actually paying back 3 MUs instead of paying the fine of 4 MUs. Therefore, if the world were populated only by selfish subjects we should observe that all investors impose a fine and that only if a fine is imposed will the trustees pay back a positive amount. In such a world the available sanction would be used to enforce the cooperation of selfish trustees, and altruistic cooperation would be absent.

All the interactions in the experiment took place anonymously and each pair of subjects played the experiment only once. This means that repeated interactions[12,13] or reputation formation[14–17] cannot account for positive back-transfers. In the absence of a fine, positive back-transfers can, therefore, be characterized as altruistic cooperation. Likewise, if the investor imposes a fine, back-transfers exceeding the fine of 4 MUs represent altruistic cooperation because selfish trustees will never pay back more than 4 MUs.

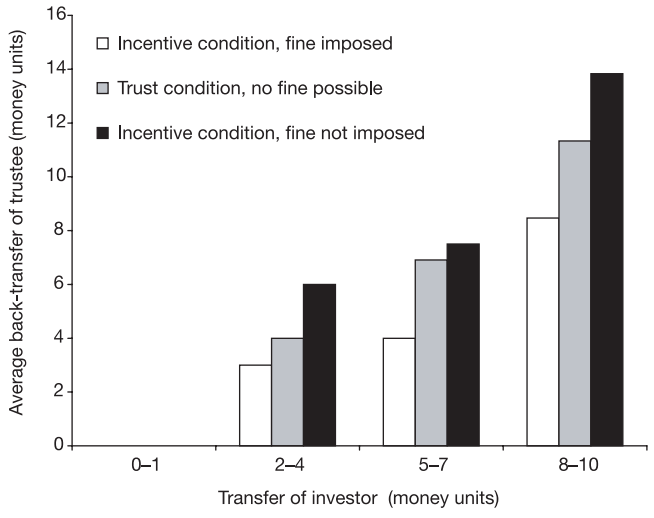

Figure 1 The average back-transfer of the trustees plotted as a function of the investors' transfers. In the trust condition and the incentive condition the back-transfers of the trustees increase with the investors' transfers—irrespective of whether the fine is imposed or not. If the investors impose a fine in the incentive condition, trustees reduce their back-transfers, indicating a detrimental effect of the incentive on altruistic cooperation. Trustees' back-transfers are highest if the investor deliberately refrains from imposing the fine.

Sanctions and altruistic cooperation

In our first experiment, 24 pairs of subjects participated in the trust condition and 45 pairs in the incentive condition. To examine the motives behind the investors' choices in more depth we later conducted an additional experiment with 50 pairs in the incentive condition. Figure 1 shows the behaviour of trustees in the two conditions of our first experiment. In contrast to the self-interest hypothesis the trustees paid back substantial amounts of money in all conditions. In addition, back-transfers were increasing in the investors' transfer, indicating the reciprocal nature of human altruism. This kind of altruism should not, however, be confused with the notion of "reciprocal altruism" as defined by Trivers[12]. In our experiment there are no repeated interactions so that trustees cannot hope to receive future rewards from their current altruistic acts. 19 of the 24 trustees (79%) in the trust condition paid back more than zero. 29 of the 45 trustees (64%) in the incentive condition paid back more than 4 MUs. This confirms that altruism is an important force in our setting. More importantly, however, there were striking differences in trustees' altruism across conditions (Fig. 1). For any interval of the investors' transfers the trustees' back-transfers were highest when the investor voluntarily refrained from the fine in the incentive condition and lowest when the investor imposed the fine. The back-transfers in the trust condition were always at an intermediate level. When the investor voluntarily refrained from the fine no trustee chose a back-transfer of zero and 47% (7 out of 15 subjects) chose a back-transfer of 15 MUs or more, yielding an average back-transfer of 12.5 MUs (Table 1). In contrast, if a fine was imposed, 33% (10 out of 30 subjects) paid back nothing and only 13% (4 out of 30 subjects) paid

Table 1 **Average behaviour and payoffs of investors and trustees**

	Trust condition	Incentive condition, fine chosen	Incentive condition, no fine chosen
Investment	6.5	6.8	8.7
Desired back-transfer as a percentage of tripled investment	59.9	67.4	63.7
Actual back-transfer	7.8	6.0	12.5
Actual back-transfer as a percentage of tripled investment	40.6	30.3	47.6
Actual back-transfer as a percentage of desired back-transfer	74.4	54.5	74.1
Investor's payoff	11.3	9.2	13.8
Trustee's payoff	21.8	22.4	23.5
Number of observations	24 pairs	30 pairs	15 pairs

back 15 MUs or more, causing an average back-transfer of only 6 MUs. This difference in back-transfers is significant (Mann–Whitney test, $z = -2.988$, $P = 0.003$, two-tailed), providing a first indication that the imposition of the fine had detrimental effects on altruistic behaviour.

One reason for the higher back-transfers in the absence of fines could be that those investors who did not impose a fine transferred higher amounts to the trustees (Table 1). Investors who imposed a fine transferred 6.8 MUs on average whereas those who refrained from fining transferred 8.7 MUs (Mann–Whitney test, $z = -1.961$, $P = 0.050$, two-tailed). Trustees tended to pay back more money when they had received higher transfers (Fig. 1), so the higher back-transfers in the absence of the fine could be due to the higher transfers of the investors. To control for this effect we computed the actual back-transfer as a percentage of the tripled investment (Table 1). We find that if the fine was imposed the trustees paid back 30.3% of the tripled investment, and if investors voluntarily refrained from fining, significantly more, 47.6%, was paid back (Mann–Whitney test, $z = -1.930$, $P = 0.054$, two-tailed). A similar picture emerges if we examine how much the trustees paid back relative to the desired back-transfer. If the fine was imposed the trustees paid back only 54.5% of the desired level whereas in the absence of the fine 74.1% of the desired level was paid back (Table 1).

The negative impact of sanctions on altruistic responses is also supported by a regression analysis, which controls for the investor's transfers and the desired back-transfers. Keeping these other factors constant, using the fine in the incentive treatment reduced the back-transfers significantly by 4.56 MUs ($t = 2.380$, $P = 0.020$, two-tailed). As a consequence, investors who did not use the fine earned significantly more money (Mann–Whitney test, $z = -2.283$, $P = 0.022$): 13.8 MUs in the absence of the fine and only 9.2 MUs when the fine was imposed (Table 1). The imposition of the fine reduced the back-transfers by 4.56 MUs, so it is tempting to assume that trustees who decided to pay back less than the desired amount—and hence had to pay the fine—reduced the back-transfer by the amount of the fine. This was, however, not the case. Those trustees who were actually fined paid back 2.94 MUs on average, whereas those who transferred back the desired amount, so that they did not have to pay the fine, paid back 9.5 MUs on the average. The median back-transfer of those who actually had to pay the fine was zero MUs.

Why do so many investors in the incentive treatment—two-thirds (30 out of 45)—threaten to sanction the trustee if the threat is associated with much lower back-transfers? We hypothesized that there are two potential reasons for this behaviour. First, preferences for strong reciprocity[24,25], which have been shown to prevail in many situations[8,9,11], could be the proximate cause. A strong reciprocator is willing to sacrifice resources for rewarding behaviour that is perceived as kind or fair and for punishing behaviour perceived as hostile or unfair, even if reciprocation is costly and provides no present or future material benefits whatsoever[24]. The variance in trustees' back-transfers is considerable, so the investor always faces the risk of being cheated. Therefore, strong reciprocators will impose a fine in case of malfeasance even if it reduces their overall earnings. Second, investors may simply not have anticipated the detrimental effects of the incentive on the back-transfers.

To discriminate between these two hypotheses we conducted a second experiment with 50 pairs of subjects in the incentive treatment. In this experiment, we informed the investors about the behaviour of the trustees in the first incentive treatment. We gave them two scatter diagrams that showed the trustees' back-transfers for every level of the investors' transfers. One diagram depicts the trustees' behaviour when they faced the fine, the other one shows trustees' behaviour in the absence of the fine. If the investors' preferences for fining were caused by false expectations about the effects of the fine they should change their decisions when informed about the true effects of the fine. This did, however, not occur. Again, roughly two-thirds of the investors (34 out of 50) preferred to fine the trustees in case of non-compliance. This suggests that the desire to punish cheaters is the reason for the frequent use of the fine.

Why do sanctions reduce altruistic cooperation?

All the other qualitative findings observed in the first incentive treatment were also replicated in our second experiment. In particular, we again observed a sizeable increase in the back-transfers if the investors voluntarily waived the possibility of imposing the fine. How can we explain this behaviour by trustees? The notion of strong reciprocity can provide an answer to this question, too. First, refraining from the threat of fining, although the threat is available, could itself be perceived as a fair act, which

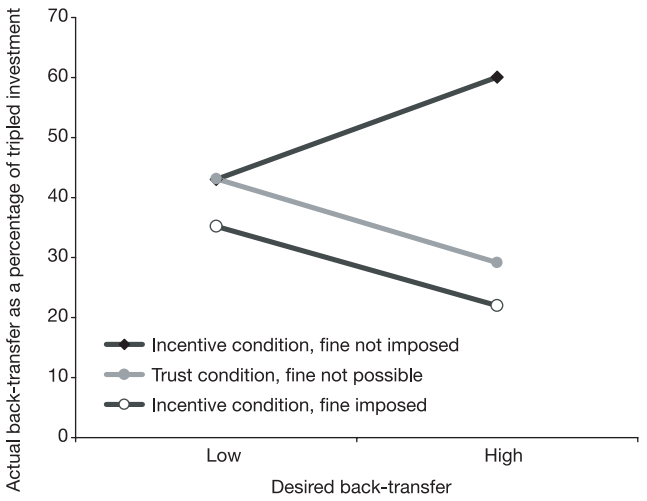

Figure 2 Actual back-transfer as a percentage of tripled investment plotted as a function of the investors' desired back-transfers. The desired back-transfer is categorized as low if, in the case that the trustee pays this back-transfer, the investor would earn the same or less than the trustee. It is categorized as high if the investor would earn more than the trustee. At low desired back-transfers the sanctioning threat reduces actual back transfers in the incentive treatment but the effect is not statistically significant. At high desired back-transfers the sanctioning threat has a large negative impact on the actual back-transfers in the incentive condition, indicating that if sanctions are used to achieve a distributional advantage they strongly undermine altruistic cooperation.

induces the trustees to increase their cooperation. Second, attempts to use the sanction to enforce an unfair distribution of income may be perceived as hostile acts, inducing the trustees to reduce cooperation.

A back-transfer of two-thirds of the tripled transfer equalizes the trustee's and the investor's payoff. Therefore, we classified desired back-transfers of two-thirds or less of the tripled transfer as low, and desired back-transfers above two-thirds of the tripled transfer were classified as high. We find that for low desired back-transfers the imposition of the fine decreases the actual back-transfers in the incentive condition from 43% to 35.2% of the tripled transfer (Fig. 2). This is evidence that the fine itself diminishes altruistic responses but the effect is not significant (Mann–Whitney test, $z = -0.520$, $P = 0.603$, two-tailed). A significantly large effect occurs, however, if the fine is combined with high desired back-transfers. In this case, the imposition of the fine leads to back-transfers of only 22% of tripled investments whereas voluntarily refraining from fining elicits back-transfers of 60.1% (Fig. 2, Mann–Whitney test, $z = -2.262$, $P = 0.024$, two-tailed). In fact, if the fine is imposed and the desired back-transfers are high, 46% of the trustees give back nothing and the average back-transfers are only 3.82 MUs, which suggests that altruistic responses have largely vanished.

Previous research on 'public good' experiments has shown that altruistic punishment is a highly effective means of enforcing cooperation[26]. Why do the sanctions in the public good context enhance cooperation but reduce cooperation in our context? Our results suggest that the moral legitimacy of the sanction is a crucial factor. In the public good context the punishment of 'free-riders' is an altruistic act that is considered as morally legitimate. Therefore, the disciplining effect of the sanctions is not undermined by a decline in altruistic cooperation. In the present context, punishment serves the punisher's self-interest and, if it is used to enforce an unfair payoff distribution in favour of the punisher, it decreases altruistic responses. This interpretation is also supported by the following fact. If the fine is imposed in the incentive condition the actual back-transfers as a percentage of tripled investments are significantly lower when the desired back-transfers are high than when the desired back-transfers are low (Mann–Whitney test, $z = -1.85$, $P = 0.064$, two-tailed).

In human societies, social order and cooperation rely on both the use of rewards and sanctions, which ensures the compliance of self-interested actors, and on the presence of people willing to perform altruistic acts. However, as our experiments indicate, sanctions intended to deter non-cooperation may backfire because they undermine altruistic cooperation. Whereas altruistically motivated sanctions for the benefit of the group enhance cooperative behaviour, sanctions that are imposed to enforce an unfair distribution of resources have the opposite effect. This is in contrast to prevailing approaches in economics, biology and behavioural ecology, which predict cooperation-enhancing effects of sanctions, regardless of the moral legitimacy or purpose of each sanction.

Our results, however, do not imply that economic incentives generally have negative motivational effects. There is evidence indicating that rewarding incentives and moralistic sanctions in repeated interactions have large positive effects on cooperation[27]. This suggests that economic incentives cause mainly negative effects on altruistic cooperation if they come in the form of sanctions and if they are associated with greedy or selfish intentions. We believe that these insights are important for everyone who needs the voluntary compliance of other people. □

Methods

The experiments were conducted with 238 students at the University of Bonn. They took place in a large lunch room for students where we recruited subjects on the spot to participate in the experiment. At noon the lunch room is visited by hundreds of students from various disciplines. Each pair of subjects played the game just once and no subject could participate in the experiment for a second time. At registration, subjects received an instruction sheet (available from the authors upon request) that explained the rules of the game, the payoff structure, the random and anonymous matching with another student subject recruited at the same occasion, the conversion rate of experimental MUs into German Marks and the payment mode. Subjects were privately paid immediately after the experimental session. After having read the instructions, each subject was randomly assigned to one of two roles (investor or trustee) and anonymously matched with a subject in the other role. Investors received a blank decision sheet in which they had to fill in the transfer and the desired payback amounts. In the incentive condition, they additionally had to mark one of two options 'no deduction' and 'a deduction of 4 MUs' for the case where the trustee would pay back less than the desired amount. While making their decisions subjects were seated in a polling booth that we had brought to the lunch room. After an investor had made his three decisions a monitor checked whether all the necessary information was on the decision sheet. Then the monitor handed the decision sheet to the trustee with whom the investor was matched. The trustees had just one decision to make, namely, to specify the back-transfer to the investor. After completion the decision sheet was handed to the monitor who controlled for completeness and calculated the payoffs. Immediately after their game was over subjects were paid. Special care was taken to ensure the anonymity of the subjects and the privacy of the payment procedure. Each subject received a flat fee of DM3 (€1.53) for participation and each MU earned was rewarded with DM0.5 (€0.26). On average an investor earned DM8.45 (€4.32) and a trustee earned DM14.23 (€7.28). On average half an hour elapsed between registration and payment.

Received 29 November 2002; accepted 3 February 2003; doi:10.1038/nature01474.

1. Smith, A. *The Wealth of Nations* (The Modern Library, New York, 1776).
2. Durkheim, E. *De La Division Du Travail Social* (Presses Universitaires de France, Paris, 1902).
3. Arrow, K. Gifts and exchanges. *Phil. Publ. Affairs* **1,** 343–362 (1972).
4. Knack, S. & Keefer, P. Does social capital have an economic payoff? *Q. J. Econ.* **112,** 1251–1288 (1997).
5. Bowles, S. *Economic Institutions and Behavior* (Princeton Univ. Press, Princeton, in the press).
6. Milgrom, P. & Roberts, J. *Economics, Organization & Management* (Prentice Hall, Englewood Cliffs, 1992).
7. Dawes, R. M. Social Dilemmas. *Annu. Rev. Psych.* **31,** 169–193 (1980).
8. Fehr, E., Kirchsteiger, G. & Riedl, A. Does fairness prevent market clearing? *Q. J. Econ.* **108,** 437–460 (1993).
9. Berg, J., Dickhaut, J. & McCabe, K. Trust, reciprocity and social history. *Games Econ. Behav.* **10,** 122–142 (1995).
10. Andreoni, J. Cooperation in public goods experiments: kindness or confusion. *Am. Econ. Rev.* **85,** 891–904 (1995).
11. Ostrom, E. A behavioral approach to the rational choice theory of collective action. *Am. Pol. Sci. Rev.* **92,** 1–22 (1998).
12. Trivers, R. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46,** 35–57 (1971).
13. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211,** 1390–1396 (1981).
14. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **393,** 573–577 (1998).
15. Milinski, M., Semmann, D. & Krambeck, H. J. Reputation helps solve the 'tragedy of the commons'. *Nature* **415,** 424–426 (2002).
16. Leimar, O. & Hammerstein, P. Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. Lond. B* **268,** 745–753 (2001).
17. Zahavi, A. The cost of honesty (further remarks on the handicap principle). *J. Theor. Biol.* **67,** 603–605 (1977).
18. Mas-Colell, A., Whinston, M. D. & Green, J. R. *Microeconomic Theory* (Oxford Univ. Press, Oxford, 1995).
19. Laffont, J. J. & Martimort, D. *The Theory of Incentives* (Princeton Univ. Press, Princeton, 2002).
20. Titmuss, R. M. *The Gift Relationship* (Allen and Unwin, London, 1970).
21. Frey, B. S. *Not Just For the Money. An Economic Theory of Personal Motivation* (Edward Elgar, Cheltenham, 1997).
22. Bewley, T. F. *Why Wages Don't Fall During A Recession* (Harvard Univ. Press, Cambridge, Massachusetts, 1999).
23. Bohnet, I., Frey, B. S. & Huck, S. More order with less law: on contract enforcement, trust and crowding. *Am. Pol. Sci. Rev.* **95,** 131–144 (2001).
24. Rabin, M. Incorporating fairness into game theory and economics. *Am. Econ. Rev.* **83,** 1281–1302 (1993).
25. Gintis, H. Strong reciprocity and human sociality. *J. Theor. Biol.* **206,** 169–179 (2000).
26. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415,** 137–140 (2002).
27. Fehr, E. & Falk, A. Psychological foundations of incentives. *Eur. Econ. Rev.* **46,** 687–724 (2002).

Acknowledgements This paper is part of the European Network for the Development of Experimental Economics (ENDEAR). E.F. gratefully acknowledges support by the Ludwig Boltzmann Institut for the Analysis of Economic Growth, by the Swiss National Science Foundation and by the MacArthur Foundation Network on Economic Environments and the Evolution of Individual Preferences and Social Norms. B.R. gratefully acknowledges support by the Deutsche Forschungsgemeinschaft and the Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen for providing the subject payments.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.F. (e-mail: efehr@iew.unizh.ch).

