



OPEN

The role of institutional incentives and the exemplar in promoting cooperation

SUBJECT AREAS:
SOCIAL EVOLUTION
EXPERIMENTAL EVOLUTION

Jia-Jia Wu^{1*}, Cong Li^{1*}, Bo-Yu Zhang^{2*}, Ross Cressman³ & Yi Tao¹

Received
3 April 2014

Accepted
26 August 2014

Published
22 September 2014

Correspondence and requests for materials should be addressed to R.C. (rcressman@wlu.ca) or Y.T. (yitao@ioz.ac.cn)

* These authors contributed equally to this work.

¹Key Lab of Animal Ecology, Institute of Zoology, Chinese Academy of Sciences, Beijing, China, ²School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing, China, ³Department of Mathematics, Wilfrid Laurier University, Waterloo, ON Canada.

People on average do not play their individually rational Nash equilibrium (NE) strategy in game experiments based on the public goods game (PGG) that model social dilemmas. Differences from NE behavior have also been observed in PGG experiments that include incentives to cooperate, especially when these are peer-incentives administered by the players themselves. In our repeated PGG experiment, an institution rewards and punishes individuals based on their contributions. The primary experimental result is that institutions which both reward and punish (IRP) promote cooperation significantly better than either institutions which only punish (IP) or which only reward (IR), and that IP has contribution levels significantly above IR. Although comparing their single-round NE strategies correctly predicts which incentives are best at promoting cooperation, individuals do not play these strategies overall. Our analysis shows that other intrinsic motivations that combine conforming behavior with reactions to being rewarded/punished provide a better explanation of observed outcomes. In our experiments, some individuals who display more cooperation than other individuals can be regarded as the exemplars (or leaders). The role of these exemplars in promoting cooperation provides important insights into understanding cooperation in PGG and the effectiveness of institutional incentives at promoting desirable societal behavior.

The evolution of cooperation is still a puzzle in evolutionary biology and social science. The prisoner's dilemma (PD) game should be considered to be the gold standard for modeling the evolution of cooperation. The public goods game (PGG) is a PD game with multi-players^{1–6}. In a single-round PGG, each player in a group of size n is given a fixed endowment and contributes as much of this as he or she wants to a common pool. The total amount in the pool is multiplied by a factor r with $1 < r < n$ and then redistributed evenly to each player in the group. It is to the group's advantage if all players contribute their total endowment since $r > 1$ but each player, given the contributions of the others, does best by contributing nothing since $r < n$. That is, the only individually Nash equilibrium (NE) of PGG based on monetary considerations is for all players to free ride. However, experiments on PGG do not agree with this NE prediction. Typically, individuals contribute about half their endowment on average to the common pool, and this average tends to decrease as the individuals play the single-round PGG more often among randomly formed groups, starting as high as 70% and decreasing to no lower than 20%^{1,2,7–9}. This outcome is also observed for the repeated PGG game where the same group plays the game over and over^{1,7–9}.

Theoretical research^{10–16} has provided an explanation based on NE behavior (through such effects as reputation, reciprocity and revenge) for the evolution of cooperative behavior in the repeated PGG when individuals are assumed to act solely in their own monetary best interests. Experiments with peer-incentive^{1–6} and with institutional-incentive^{17–21} have also been conducted to test whether reward and/or punishment increases cooperative behavior. Other studies argue that observed behavior in repeated PGG can be described through a combination of these self interests and other intrinsic motivations such as inequity aversion^{22–23} or shame/honour²⁴. However, for the repeated PGG experiments, it is still not clear whether individual behaviors only depend on their own best interests; how individuals avoid (or dilute) the risk under the reward/punishment incentives; how individuals' reactions differ when rewarded compared to punished; and whether individuals being more cooperative increases the cooperation level of other group members.

To reveal mechanisms behind individual behavior in PGG, we designed experiments with institutional incentives. Here, it is not individuals who reward or punish; rather each individual is rewarded (or punished) with known probabilities that increase (or decrease) as the amount he/she contributed goes up²⁵. In our experiments that include an institutional reward (respectively, punishment), exactly one member of the group is selected in this way to be rewarded (respectively, punished). We base our experiments on the repeated multi-player PGG to



Table 1 | Nash equilibrium behaviors in a single round when all group members contribute the same amount. The entries give the NE in Control and in each of the nine treatments

	Const	Up	Down
Control	0		
IR	5.25	5.06	5.79
IP	0 or 20	0 or 20	0 or 20
IRP	20	20	20

reflect that an institution includes several members and uses incentives on an ongoing basis. In fact, the use of institutional incentives is a common feature in nature (e.g. the maintenance of cooperation among subordinates through punishment in hierarchical species²⁶) and many parts of human society such as in businesses and in government institutions^{20–21,27–29}.

In this study, we emphasize the prominent roles that imitative and reactive behaviors play in explaining observed outcomes in our repeated PGG experiments. In particular, we study the effects of conformists who base their behavior on contributions rather than income by changing their contribution in the next round in the direction of the average group contribution in the current round³⁰. On the other hand, we also consider the difference between the reward and punishment incentives in promoting cooperation based on individual reactive behavior to these incentives. Finally, some individuals in our experiments can be regarded as the exemplars (leaders) in cooperation since they display more cooperation than other individuals, and the role of these exemplars in promoting cooperation (i.e. the effect these exemplars have on the average group contribution level) is considered.

Methods

A total of 792 university students participated in our PGG experiments at the School of Mathematical Sciences Computer Lab, Beijing Normal University. Subjects interacted anonymously via computer screens for 50 rounds of the repeated game among the same four players (Supplementary Information (SI), Methods). The control experiment (Control) is a standard four-player repeated PGG^{1,6}. In each round, every subject receives an endowment of 20 monetary units and decides how much to contribute to the public pool, keeping the rest for himself. The contributions are increased by 60% and split evenly among the four group members.

To reveal how individuals respond to institutional punishment/reward in PGG, we designed three different incentive schemes (SI, Methods), called Const, Up and Down, respectively. The amount of punishment/reward is fixed at 20 monetary units in Const and increases linearly from 16 to 25.6 (or decreases from 25.6 to 16) monetary units per round as a function of the group's total contribution in Up (or Down). For each of Const, Up and Down, there are three treatments. In these three treatments, each round of PGG is followed by a second stage, which corresponds respectively to an institutional punishment (IP), an institutional reward (IR), or both institutional reward and punishment (IRP). As described in the following paragraph (see also SI, Methods), after each round of the IP (respectively, IR) treatment, one of

the four subjects is selected to be punished (respectively rewarded). The probability a particular subject is selected decreases (respectively, increases) as the contribution the subject makes to the common pool increases¹⁵. Examples of these probabilities relevant for their experiment as well as the amount of the punishment or reward are given to the participants as part of their instructions (SI, Methods). From these examples they gain experience with the monetary consequences of different single-round contribution levels. In the IRP treatment, one of the four subjects is selected to be punished with the IP probabilities above and one subject is independently selected to be rewarded with the IR probabilities. The Control and the nine treatments had approximately 20 groups apiece resulting in almost 800 participants.

In a single round of our repeated PGG experiment with institutional incentives, an individual's expected payoff is $\pi^C \equiv 20 - c + 1.6\bar{c}$ in Control when he/she contributes c and the average contribution of a group member is \bar{c} . For the treatments, an individual's expected payoff are $\pi^{IR} = \pi^C + P^{IR}A$ in IR; $\pi^{IP} = \pi^C - P^{IP}A$ in IP; and $\pi^{IRP} = \pi^C + (P^{IR} - P^{IP})A$ in IRP where A denotes the amount of incentive (which is $A = 20$ in Const, $A = 16 + 0.48\bar{c}$ in Up and $A = 16 + 0.48(20 - \bar{c})$ in Down) and P^{IR} (respectively, P^{IP}) is the probability the individual is rewarded in IR (respectively, punished in IP). Specifically, $P^{IR} = (c+1)/4(\bar{c}+1)$ and $P^{IP} = (21-c)/4(21-\bar{c})$ (SI, Analysis 2.2). In particular, an individual who contributes the group average receives the incentive 25% of the time. Furthermore, an individual who cooperates by contributing 20 units when the rest of the group free-rides has a high chance of being rewarded (87.5%) whereas the others are rewarded only about 4% of the time. This latter small probability reflects the possibility that the institution makes a mistake as to who the cooperator is. In particular, no participant can guarantee he/she will receive the reward in IR (or prevent punishment in IP). Due to these institutional "mistakes", it is reasonable that the same participant can be rewarded and punished in the same round of IRP (i.e. the reward and punishment mechanisms are independent). In fact, many peer-incentive experiments^{3–4,6} allow the same individual to be both rewarded and punished. We also mention here that a previous study used a different exogenous mechanism to examine efficient probabilistic punishment²⁵.

Following the methods of a recent theoretical study¹⁵, the predicted contribution of a rational player in a single round (i.e. his/her NE behavior) for all three incentive schemes can be easily calculated (SI, Analysis 2.3). As noted above, these single round PGG results are also relevant for our repeated game since outcomes are similar whether the game is played among randomly chosen groups or the same group plays over and over^{1,2,7–9}. If an individual contributes c and the average contribution of his/her other group members is \bar{c} , then the contribution level \bar{c} is called a NE if the actual expected payoff of this individual is highest when $c = \bar{c}$. It is well-known that the NE for Control is to free-ride^{1–6}. From SI (Analysis 2.3), there is a unique NE in IR at about 5, IP has two NE 0 and 20, and IRP has a unique NE at 20. Specifically, for each of Const, Up and Down, the NE behaviors in Control and in treatments IR, IP and IRP are shown in Table 1. Thus, single-round NE behavior predicts our rewards provide a powerful incentive to get cooperation started but are less well suited to maintain cooperation and the converse holds for our punishments^{12,15,16,20,31} (SI, Analysis 2.3 and Figure S1). Furthermore, we expect that the NE outcomes of Table 1 will appear in the corresponding theoretical studies since all these NE are stable under evolutionary dynamics that model how myopic individuals change their contribution in the direction of increasing payoff⁵.

Results

The primary experimental result (Figure 1) is that for all three incentive schemes, IRP is significantly better than either IP or IR in promoting cooperative behavior. Furthermore, IP has contribution levels significantly above Control while the levels in IR are above Control on average but never significantly so. The statistic analysis of Figure 1 is given in SI (SI, Tables S2–4 in Analysis 2.1).

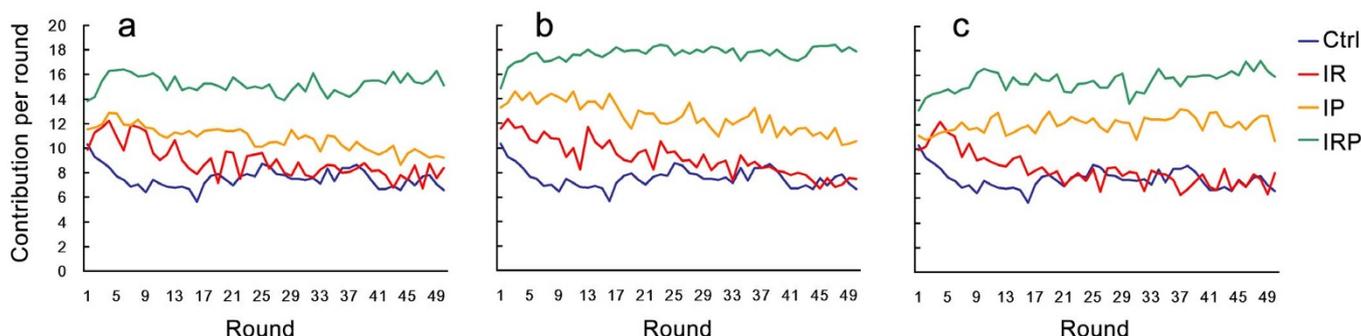


Figure 1 | Average contribution per round for institutional incentives compared to Control (Ctrl). Panels (a), (b) and (c) show the results in the three schemes Const, Up and Down, respectively. The data is analyzed at the level of the group (i.e. average group contribution per capita) to avoid interdependence of outcomes for members of a given group. The average contribution per session (i.e. over all 50 rounds) are: 7.55 ± 4.65 in Control; IRP 15.23 ± 2.08, IP 10.79 ± 4.65 and IR 8.95 ± 2.08 in Const; IRP 17.69 ± 1.64, IP 12.55 ± 4.25 and IR 9.18 ± 2.85 in Up; IRP 15.52 ± 2.20, IP 11.80 ± 4.35 and IR 8.32 ± 2.34 in Down.



Table 2 | Definitions for the three behavioral types; conforming, cooperating and defecting. For instance, if an individual contributes less than the group average in round $k-1$ (first row of table), he is conforming between this round and the next if he increases his contribution (last column of table) and defecting otherwise (other two columns)

		Round k		
		Decrease contribution	Same contribution	Increase contribution
Round $k-1$	Contribution lower than group average	defecting	defecting	conforming
	Contribution same as group average	defecting	conforming	cooperating
	Contribution higher than group average	conforming	cooperating	cooperating

From Table 1, the evolutionary outcome in the incentive schemes will also be more cooperative than Control (this is true in IP as well since some groups will free-ride while others will become fully cooperative and contribute 20). Although the single-round NE predictions do not exactly match the experimental results (Figure 1), this evolutionary behavior is consistent with the experimental results where the trend in IRP (respectively, Control and IR) is to increase (respectively, decrease) their contributions toward the single NE as they play more rounds (Figure 1). However, although motivations based on single-round payoffs correctly predict the evolutionary direction, individuals overall do not play these NE strategies and so we now consider behaviors based on other intrinsic motivations.

Group selection cannot be used to explain our results. For example, IRP is a self-financing incentive scheme³² (i.e. the reward and punishment amounts cancel themselves out) and so groups with the same contribution have the same total payoff in one round of Control as they do in IRP. Thus, group imperatives cannot explain why these two protocols always display the largest difference in contributions. Unlike peer-incentive PGG experiments^{1–6} where group payoffs are typically examined in conjunction with group contribution to judge the effectiveness of different protocols in promoting cooperation (e.g. group selection in favor of higher contribution becomes a bigger factor when such contributions produce larger effects on group payoffs), group selection arguments are inconsistent with our results. Instead, we turn to other individual motivations as contributing causes for the observed differences in contribution levels among our experimental protocols.

Conforming behavior has been used to explain observed outcomes in PGG experiments and other social dilemmas^{30,33}. In the theoretical analysis of a sequential PGG³⁰, a conformist (called an imitator there) contributes the average contribution of all individuals who have already played. In the analysis of a repeated PGG experiment³⁴, a conformist moves his contribution in the next round in the direction of the group average for the current round. In particular, conformists base their behavior on contributions rather than income. This latter approach is used in Table 2 to define conforming behavior as well as defecting and cooperating behavioral types that partition the participants in each round of the experiment into three behavioral types.

From Figure 2a, the proportion of participants who display conforming behavior is essentially constant at about 52.5% in all our experimental protocols (SI, Figure S4 and Table S5 in Analysis 2.4), supporting the conclusion that some, if not all, decisions are indeed based on contributions. Furthermore, the analysis of Figure 2b (SI, Analysis 2.5) shows participants are not randomly choosing their contributions without regard to payoffs or the contributions of other group members. It is also clear from Figure 2a that cooperating behavior is more pronounced in treatments that include punishment. From SI (Analysis 2.6), there is a higher between-round transition rate to cooperating behavior in these treatments as well. In fact, the limiting distribution of the Markov process generated by the transitions among behavioral types between rounds can be used to predict the outcomes in all our experimental protocols (SI, Analysis 2.6).

For our incentive experiments, a more challenging question is how participants' reactions differ when rewarded compared to punished.

These reactions, which are now a combination of decisions based on payoffs and on contributions, can best be explained by comparing them in the IR and IP protocols. Figure 2c (respectively, Figure 2d) shows the proportions of individuals in all three IR (respectively, IP) protocols who increase (+), decrease (–) and do not change (0) their contribution in the round that follows receiving a reward (respectively, punishment), given their current contribution is between 1 and 19. Players who currently contribute 0 (respectively, 20) are not included since they cannot decrease (respectively, increase) their contribution. Similarly, the reactions of participants who are not rewarded in IR (not punished in IP) are plotted in Figure S6a (Figure S6b) (see SI, Analysis 2.7). From Figure 2c–d (SI, Tables S8a–c in Analysis 2.7), there are significantly more individuals who increase their contribution after being punished in IP than after being rewarded in IR. In fact, this result can be attributed to the difference in reaction to reward or punishment among those who are currently contributing between 11 and 19 (i.e. among high contributors) since there is no significant difference among those who contributed less than 10. Moreover, there are the following differences in behavior between individuals who receive an incentive and those who do not. From Figure 2c–d and Figure S6a–b, among high contributors, individuals who are punished in IP (respectively, rewarded in IR) are significantly more likely to increase (respectively, decrease) their contribution than those who are not punished in IP (respectively, those not rewarded in IR). These reactions to being the recipient of an incentive are consistent with single-round NE behavior. Specifically, Table 1 shows that high contributors in IR should decrease their contribution in order to move in the direction of the unique NE whereas high contributors in IP move toward their nearest NE by increasing their contribution. It is also interesting to note that Table 1 implies that punishment is more efficient in promoting cooperation among high contributors than low contributors since low contributors in IP decrease their contribution to move toward their nearest NE (see also the Discussion). This argument is in line with experimental results questioning the effectiveness of punishment^{21,35}.

To show how individual behavior based on contribution affects the average contribution, we define a participant as a conforming-individual (cooperating-individual, defecting-individual) if the frequency that he/she displays conforming behavior (cooperating behavior, defecting behavior) over all 49 rounds (i.e. from round 2 to round 50) is larger than 50%. Obviously, this definition concerns the holistic characteristic of each individual's behavior. According to this definition, the proportions of conforming-individuals, cooperating-individuals and defecting-individuals in all our experimental protocols are shown in Figure 3a along with the proportion who fit none of these three categories. It is easy to see that the proportion of conforming-individuals is about 60% in all treatments. For Control and treatments IR, IP and IRP, the change in the proportion of cooperating-individuals (defecting-individuals) has the pattern with Control < IR < IP < IRP (Control \approx IR > IP \approx IRP) (Figure 3a and SI, Table S9). This result strongly suggests it is only the change in proportion of cooperating-individuals (or defecting-individuals) that responds sensitively to the incentives, since the proportion of conforming-individuals seems to be independent of the incentives. This result exactly matches the observations in Figure 2a as well.

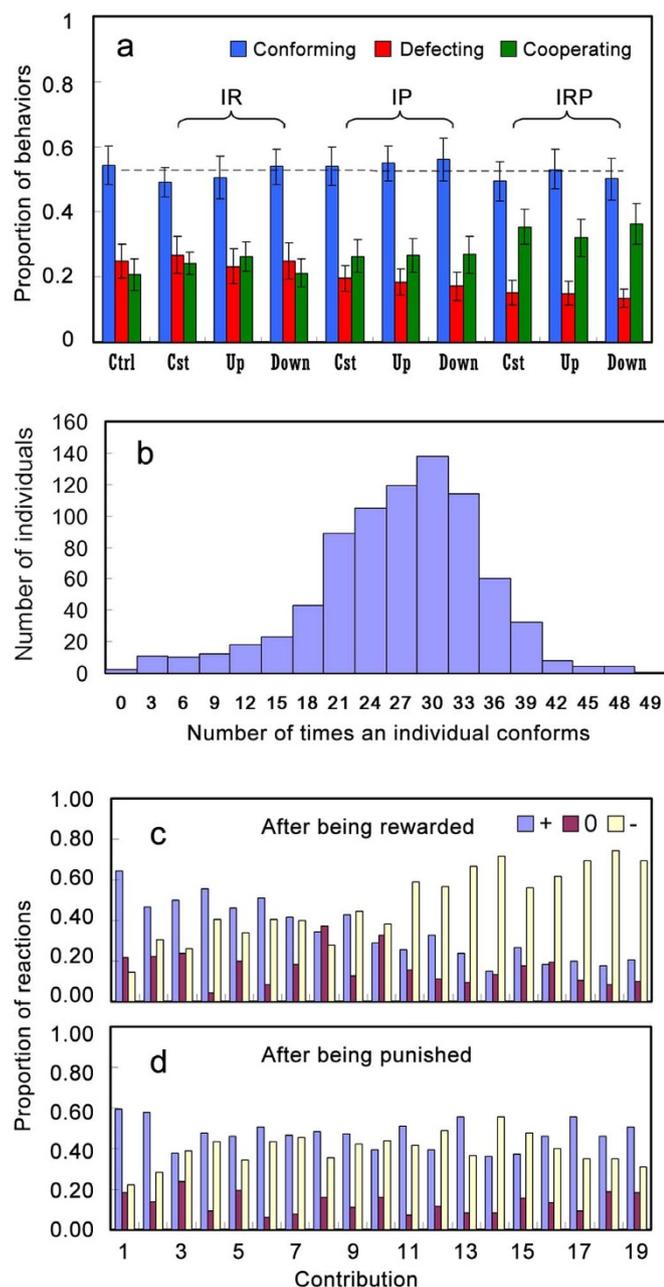


Figure 2 | Individual behavior based on contribution and individual reactive behavior to reward/punishment incentives. (a) For Control and treatments IR, IP and IRP in each of Const, Up and Down, the average proportions of participants who display conforming, defecting and cooperating behaviors over all 49 rounds, respectively, are plotted. (b) The distribution of the number of times (of a possible 49) that an individual displays conforming behavior from one round to the next. (c) The reaction to being rewarded in IR in the current round (where the symbols “+”, “0” and “-” represent “increase contribution”, “no change in contribution” and “decrease contribution”, respectively, in the following round). The proportions of these three reactions over all 49 rounds are shown for a given contribution in the current round. (d) The reaction to being punished in IP in the current round.

Finally, Figure 3b shows that the average group contribution is proportional to the proportion of cooperating-individuals. This means that the cooperating-individuals can be regarded as “exemplars” (or “leaders”) whose behaviors are tracked by conformists who follow the resultant change in group average contribution. In particular, since there is on average one cooperating-individual in each IRP

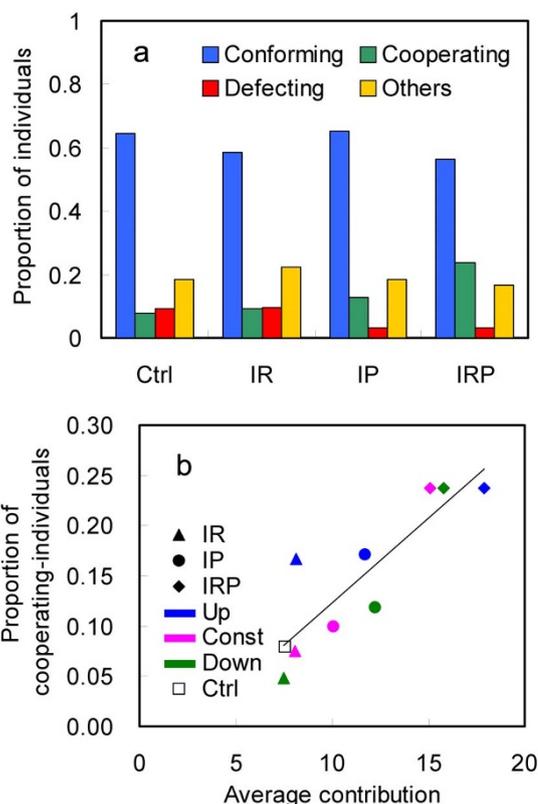


Figure 3 | Conforming-, cooperating- and defecting-individuals and their effects on the average contribution. (a) The blue, green and red bars denote the proportions of conforming-, cooperating- and defecting-individuals, respectively. The yellow bar (Others) denotes the proportion of individuals who display each of conforming, cooperating and defecting behaviors in less than 50% of the rounds. (b) The average group contribution versus the proportion of cooperating-individuals. This shows clearly that the increase in the proportion of cooperating-individuals will lead to the increase of the average group contribution.

group (i.e. the proportion of cooperating-individuals in all three IRP protocols is about 25% compared to less than 18% in all other treatments in Figure 3b), this exemplar leads his/her group to higher group average contribution. This conclusion is also consistent with the observation of a previous study³⁶, in which cooperating-individuals are particularly good at leading by example.

Discussion

Although we mostly attributed the reaction of a participant who receives an incentive to its effect on payoffs, these reactions may also be a reflection of society’s different attitudes to rewards and punishments. Participants who are punished in IP even though they are contributing at a high level want to avoid the shame of being punished again and so contribute more²⁴. Shame is probably not an important factor for low contributors. Punishment will not be effective at increasing their contribution if they decide it is unlikely they will be punished two rounds in a row when there are other low contributors in their group. On the other hand, both low and high contributors who are rewarded in IR are satisfied with their reward and the status quo^{37–38} and so see no reason to increase their chance of being rewarded.

Whatever the ultimate reasons behind these reactions to rewards and punishments, it is also important to understand why individuals would consider conforming behavior in our experiments. Conforming is a good choice for individuals who want to avoid risk in a complex game since it can guarantee that their income level will be



at the average (SI, Figure S7 in Analysis 2.8). That is, even though the conformist cannot win the game, he/she does not lose either and so it is ‘not a bad choice’ to follow group behavior, especially if his/her NE behavior is difficult to analyze³⁹.

Most previous theoretical and empirical studies on the repeated PGG assumed that an individual’s behavioral decision mainly depends on benefits (or incomes)^{10–16}. However, in our experiments, observed outcomes can be better explained by dividing individuals into conforming, cooperating and defecting behavioral types³⁴ based on contributions. The prevalence of conforming in all experimental protocols agrees with the fact that this behavior is common in nature and human society^{39–40}. Conformists in our experiments are able to avoid (or dilute) risk by guaranteeing their income will be close to average. We also found that participants who are the recipient of an institutional incentive react differently than those who are not and that reactions to receiving a reward are different to being punished. Our analysis shows that these different reactions predict the effect of incentives in our experiments.

Finally, based on the holistic characterization of an individual’s behavior over all 50 rounds of the repeated PGG, we found that “cooperating-individuals” can be regarded as the exemplars (or leaders) and that their behaviors are tracked by conformists who follow the resultant change in group average contribution. This opinion is strongly supported by the experimental evidence. In particular, the proportion of “conforming-individuals” is almost the same for all incentive treatments and the change in the proportion of “cooperating-individuals” (respectively, “defecting-individuals”) corresponding to Control and treatments IR, IP and IRP has the pattern Control < IR < IP < IRP (respectively, Control ≈ IR > IP ≈ IRP). That is, this experimental evidence can be used to explain why IRP is significantly better than either IP or IR in promoting cooperation and why IP is better than IR (see the primary results in Figure 1). The role of these exemplars in promoting cooperation provides an important insight into the mechanisms behind cooperation in our repeated PGG experiments that can be used to predict the effectiveness of institutional incentives at promoting desirable societal behavior.

- Fehr, E. & Gächter, S. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000).
- Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
- Walker, J. M. & Halloran, M. Rewards and sanctions and the provision of public goods in one-shot settings. *Exp. Econ.* **7**, 235–247 (2004).
- Sefton, M., Schupp, R. & Walker, M. The effect of rewards and sanctions in provision of public goods. *Econ. Inq.* **45**, 671–690 (2007).
- Andreoni, J., Harbaugh, W. T. & Vesterlund, L. The carrot or the stick: rewards, punishments and cooperation. *Am. Econ. Rev.* **93**, 893–902 (2003).
- Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D. & Nowak, M. A. Positive interactions promote public cooperation. *Science* **325**, 1272–1275 (2009).
- Ledyard, O. Public goods: some experimental results. *Handbook of Experimental Economics* Kagel, J. & Roth, A. (ed.) (Princeton Univ. Press, Princeton, 1995).
- Fischbacher, U. & Gächter, S. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *Am. Econ. Rev.* **100**, 541–556 (2010).
- Chaudhuri, A. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Exp. Econ.* **14**, 47–83 (2011).
- Sigmund, K., Hauert, C. & Nowak, M. A. Reward and punishment. *Proc. Natl. Acad. Sci. USA* **98**, 10757–10762 (2001).
- Hauert, C., Haiden, N. & Sigmund, K. The dynamics of public goods. *Discrete Cont. Dyn.-B* **4**, 575–587 (2004).
- Hilbe, C. & Sigmund, K. Incentives and opportunism: from the carrot to the stick. *Proc. R. Soc. B* **277**, 2427–2433 (2010).
- Sigmund, K. *The Calculus of Selfishness* Princeton Univ. Press, Princeton, (2010).
- Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466**, 861–863 (2010).
- Cressman, R., Song, J. W., Zhang, B. Y. & Tao, Y. Cooperation and evolutionary dynamics in the public goods game with institutional incentives. *J. Theor. Biol.* **299**, 144–151 (2012).

- Sasaki, T., Brannstrom, A., Dieckmann, U. & Sigmund, K. The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proc. Natl. Sci. USA* **109**, 1165–1169 (2012).
- Yamagishi, T. The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.* **51**, 110–116 (1986).
- Sutter, M., Haigner, S. & Kocher, M. G. Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Rev. Econ. Studies* **77**, 1540–1566 (2010).
- Putterman, L., Tyran, J. L. & Kamel, K. Public goods and voting on formal sanction schemes. *J. Public Econ.* **95**, 1213–1222 (2011).
- Cressman, R., Wu, J. J., Li, C. & Tao, Y. Game experiments on cooperation through reward and punishment. *Biological Theory* **8**, 158–166 (2013).
- Zhang, B. Y., Li, C., De Silva, H., Bednarik, P. & Sigmund, K. The evolution of sanctioning institutions: an experimental approach to the social contract. *Exp. Econ.* **17**, 285–303 (2014).
- Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868 (1999).
- Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785–791 (2003).
- Jacquet, J., Hauert, C., Traulsen, A. & Milinski, M. Shame and honour drive cooperation. *Biol. Lett.* **7**, 899–901 (2011).
- Qin, X. & Wang, S. Using an exogenous mechanism to examine efficient probabilistic punishment. *J. Econ. Psychol.* **39**, 1–10 (2013).
- Clutton-Brock, T. H. & Parker, G. A. Punishment in animal societies. *Nature* **373**, 209–216 (1995).
- Morgan, J. Financing public goods by means of lotteries. *Rev. Econ. Studies* **67**, 761–784 (2000).
- Kosfeld, M., Okada, A. & Riedl, A. Institution formation in public goods games. *Am. Econ. Rev.* **99**, 1335–1355 (2009).
- Andreoni, J. & Gee, L. Gun for hire: Delegated enforcement and peer punishment in public goods provision. *J. Public Econ.* **96**, 1036–1046 (2012).
- Cartwright, E. & Patel, A. Imitation and the incentive to contribute early in a sequential public good game. *J. Public Econ. Theory* **12**, 691–708 (2010).
- Hauert, C. Replicator dynamics of reward & reputation in public goods games. *J. Theor. Biol.* **267**, 22–28 (2010).
- Falkinger, J., Fehr, E., Gächter, S. & Winter-Ebmer, R. A. Simple mechanism for the efficient provision of public goods: experimental evidence. *Am. Econ. Rev.* **90**, 247–264 (2000).
- Kurzban, R. & Houser, D. Experiments investigating cooperative type in humans: a complement to evolutionary theory and simulations. *Proc. Natl. Sci. USA* **102**, 1803–1807 (2005).
- Keser, C. & van Winden, F. Conditional cooperation and voluntary contributions to public goods. *Scand. J. Econ.* **102**, 23–39 (2002).
- Wu, J. J., Zhang, B. Y., Zhou, Z. X., He, Q. Q., Zheng, X. D., Cressman, R. & Tao, Y. Costly punishment does not always increase cooperation. *Proc. Natl. Sci. USA* **106**, 17448–17451 (2009).
- Gächter, S., Nosenzo, D., Renner, E. & Sefton, M. Who makes a good leader? Cooperativeness, optimism, and leading-by-example. *Econ. Inquiry* **50**, 953–967 (2012).
- Kohn, A. Why incentive plans cannot work. *Harv. Bus. Rev.*, **71**, 2–7 (1993).
- Eisenberg, R. & Cameron, J. Detrimental effects of reward. *Am. Psychol.* **11**, 1153–1166 (1996).
- Bentley, A., Earls, M. & O’Brien, M. J. *I’ll Have What She’s Having – Mapping Social Behavior* (MIT Press, Cambridge, 2011).
- Wilson, E. O. *Social Biology – The New Synthesis* (Harvard Univ. Press, Cambridge, 2000).

Acknowledgments

This study was supported by the National Science Foundation of China (Grant No. 31270439 and Grant No. 11301032), “the Fundamental Research Funds for the Central Universities” of China and the Natural Sciences and Engineering Research Council of Canada (Individual Discovery Grant).

Author contributions

The authors have made the following declarations about their contributions: Y.T., J.W., C.L. and B.Z. designed the study; J.W., C.L. and B.Z. performed the research; J.W., C.L., B.Z., Y.T. and R.C. provided new methods or materials; and Y.T. and R.C. wrote the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wu, J.-J., Li, C., Zhang, B.-Y., Cressman, R. & Tao, Y. The role of institutional incentives and the exemplar in promoting cooperation. *Sci. Rep.* **4**, 6421; DOI:10.1038/srep06421 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative

Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>