

## Research



**Cite this article:** Dong Y, Sasaki T, Zhang B. 2019 The competitive advantage of institutional reward. *Proc. R. Soc. B* **286**: 20190001.  
<http://dx.doi.org/10.1098/rspb.2019.0001>

Received: 1 January 2019

Accepted: 1 March 2019

**Subject Category:**

Evolution

**Subject Areas:**

evolution, behaviour

**Keywords:**

Cooperation, evolutionary game, reward, punishment, efficiency, decision errors

**Author for correspondence:**

Boyu Zhang

e-mail: [zhangby@bnu.edu.cn](mailto:zhangby@bnu.edu.cn)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4428971>.

# The competitive advantage of institutional reward

Yali Dong<sup>1</sup>, Tatsuya Sasaki<sup>2</sup> and Boyu Zhang<sup>3</sup>

<sup>1</sup>School of Systems Science, Beijing Normal University, Beijing 100875, People's Republic of China

<sup>2</sup>F-power Inc., Roppongi 1-8-7-2F, Minato, Tokyo 106-0032, Japan

<sup>3</sup>Laboratory of Mathematics and Complex Systems, Ministry of Education, School of Mathematical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China

TS, 0000-0002-4635-1389; BZ, 0000-0003-0972-3240

Sustaining cooperation among unrelated individuals is a fundamental challenge in biology and the social sciences. In human society, this problem can be solved by establishing incentive institutions that reward cooperators and punish free-riders. Most of the previous studies have focused on which incentives promote cooperation best. However, a higher cooperation level does not always imply higher group fitness, and only incentives that lead to higher fitness can survive in social evolution. In this paper, we compare the efficiencies of three types of institutional incentives, namely, reward, punishment, and a mixture of reward and punishment, by analysing the group fitness at the stable equilibria of evolutionary dynamics. We find that the optimal institutional incentive is sensitive to decision errors. When there is no error, a mixture of reward and punishment can lead to high levels of cooperation and fitness. However, for intermediate and large errors, reward performs best, and one should avoid punishment. The failure of punishment is caused by two reasons. First, punishment cannot maintain a high cooperation level. Second, punishing defectors almost always reduces the group fitness. Our findings highlight the role of reward in human cooperation. In an uncertain world, the institutional reward is not only effective but also efficient.

## 1. Introduction

'How did cooperative behaviour evolve' is a fundamental question in biology and the social sciences [1]. Although kin selection and direct reciprocity provide satisfactory explanations of cooperation in small groups, the sustaining of collective efforts in sizable groups of unrelated individuals remains a problem [2,3]. Examples such as the public goods game (PGG) show that self-interested players should prefer a free-rider strategy, whereas it is to everyone's advantage for all players to contribute [4]. In real-life situations that include this sort of dilemma between individual rationality and collective advantage, incentives are often used to promote cooperation, where contributors may be rewarded and free-riders may be punished. In fact, positive and negative incentives exist not only in human societies but also in animal behaviour [5]. Understanding the consequences of such 'carrots and sticks' in boosting cooperation is a core topic in the evolutionary biology and behavioural sciences (see [6–8] for three recent review papers). Most previous studies have focused on which incentives promote cooperation best. However, in the evolutionary process, groups with higher fitness rather than a higher cooperation level can survive. We thus argue that a better measurement for the success of an incentive is the group average fitness at the evolutionary stable state. Currently, it is unclear which type of incentive is more efficient in preserving group income. Specifically, when individuals are boundedly rational and may make mistakes in decision-making, is it better to use reward or punishment? In this paper, we apply evolutionary game theory to address these questions.

In past decades, several types of incentives have been proposed to promote cooperation in social dilemma games. Most of these investigations addressed so-called peer (or decentralized) incentives, and only a minority considered so-called institutional (or centralized) incentives. In the peer incentive scenario, players are allowed to reward and/or punish others at a cost to themselves (for experimental studies, see, e.g. [9–25]; for theoretical studies, see, e.g. [26–35]). Laboratory experiments have indicated that peer incentives can curb free-riding in human populations. However, this paradigm suffers from several drawbacks. First, the use of incentives is costly, which then raises an issue of second-order free-riding. In fact, the incentive system itself is a common good that can be exploited, and reward and punishment schemes are individually disadvantageous [18]. Second, some players abuse sanctioning opportunities by engaging in antisocial punishment, which harms cooperators [14,15,18,20,22]. In the institutional incentive scenario, it is not individuals who reward or punish. Rather, an institution rewards and punishes individuals based on their contributions (for experimental studies, see, e.g. [36–42]; for theoretical studies, see, e.g. [43–51]). Institutional incentives can overcome the problem of second-order free-riding and avoid antisocial punishment. However, this approach is more wasteful than peer incentives because subjects have to pay a fee to maintain the institution even if no one is being rewarded or punished (which can be viewed as paying for the upkeep of a police force).

The use of institutional incentives is a common feature in many parts of human society such as government institutions and businesses. Theoretical studies based on evolutionary game theory have revealed that the effect of institutional incentives on cooperation can be understood in terms of the incentive size [44,46,47,49]. If the incentive is very small, then both reward and punishment have no effect on promoting cooperation, and selfish players maintain a free-riding strategy. If the incentive is sufficiently large, then both reward and punishment compel all players to cooperate. If the incentive is intermediate, then full contribution can become evolutionarily stable under punishment, and reward can cause only the stable coexistence of free-riders and cooperators. Although punishment promotes cooperation better than a reward for intermediate incentives, this approach appears more socially expensive. In fact, a rewarding institution can increase the payoff of individuals, whereas a punishment institution cannot. Furthermore, when there are free-riders in the group, the decrease in payoffs through punishment may exceed the gains increased from cooperation, which results in an overall reduction of the social welfare [39,42,46].

In this paper, we consider three types of institutional incentives, namely, reward, punishment, and a mixture of reward and punishment, and investigate which type of incentive is more efficient in preserving group income. Following [39,42,43,46,47], players of a PGG with institutional incentives have to pay a fee to the institution before the joint enterprise takes place, and the institution rewards cooperators and punishes defectors. The incentive institution is established in advance and thus entails running costs even in the case that no one deserves reward or punishment [39,42,43,46]. To better simulate reality, we assume that individuals are boundedly rational, that is, they preferentially imitate the successful strategies and may make mistakes in decision-making [42,43,52,53]. This learning process can be described by the replicator-mutator equation [54]. We compare the efficiencies

of the three types of incentives by analysing the population average fitness at the stable equilibria of the replicator-mutator equation. The main result is that the efficiencies are sensitive to decision-making errors. Without decision errors, a mixture of reward and punishment can maintain a high cooperation level and high group average fitness. However, for intermediate or large decision errors, the use of punishment almost always reduces the group welfare, and reward becomes the most efficient incentive.

## 2. Material and methods

Consider an infinitely large and well-mixed population; from time to time, a sample of size  $n$  is randomly chosen to form a PGG [46,47]. In the PGG, each player can decide whether to contribute a fixed amount  $c > 0$  knowing that this amount will be multiplied by  $r > 1$  and divided equally among all  $n$  players in the group. If  $n_C$  is the number of those players who contribute (i.e. cooperators) and  $n_D$  the number of those who do not (i.e. defectors), then the payoffs of a cooperator and a defector are  $rc(n_C)/n - c$  and  $rc(n_C)/n$ , respectively. In a PGG with institutional incentive, each player has to pay  $C_I$  for the institution, and the total amount of incentive is  $C_I n$ . In the case of institutional reward (IR), the incentive is shared among the  $n_C$  cooperators. Thus, each cooperator obtains a reward of  $(C_I n)/(n_C)$ . In the case of institutional punishment (IP), each defector analogously receives a punishment of  $(C_I n)/(n_D)$ . Finally, we consider an incentive institution that can provide both reward and punishment (IRP). Assume that  $\alpha$  of the total incentive is used for reward and that the remaining  $1 - \alpha$  is used for punishment. Thus, the payoff of each cooperator is increased by  $\alpha(C_I n)/(n_C)$ , and the payoff of each defector is decreased by  $(1 - \alpha)(C_I n)/(n_D)$ . In particular,  $\alpha = 0$  and  $\alpha = 1$  correspond to the cases of IP and IR, respectively. Therefore, IRP with  $\alpha$  provides a more general framework for studying institutional incentives.

We apply evolutionary game theory to study the evolution of cooperation in PGG with institutional incentive [4,54,55]. Let  $x$  be the frequency of cooperators in the population and  $P_C$  and  $P_D$  be the expected payoffs of a cooperator and a defector in a randomly formed PGG, respectively (see electronic supplementary material, §1.1 for payoff calculation). We then define the expected fitness of a cooperator and a defector by  $f_C$  and  $f_D$ , respectively, where  $f_C = \omega P_C + 1 - \omega$  and  $f_D = \omega P_D + 1 - \omega$ . The parameter  $\omega \in [0, 1]$  is interpreted as a selection intensity [54,55]. If  $\omega = 0$ , then the payoff of the PGG does not contribute to fitness. If  $\omega = 1$ , then the fitness is entirely determined by the payoff. Specifically, a positive  $\omega$  does not change the direction of evolution but only affects the speed.

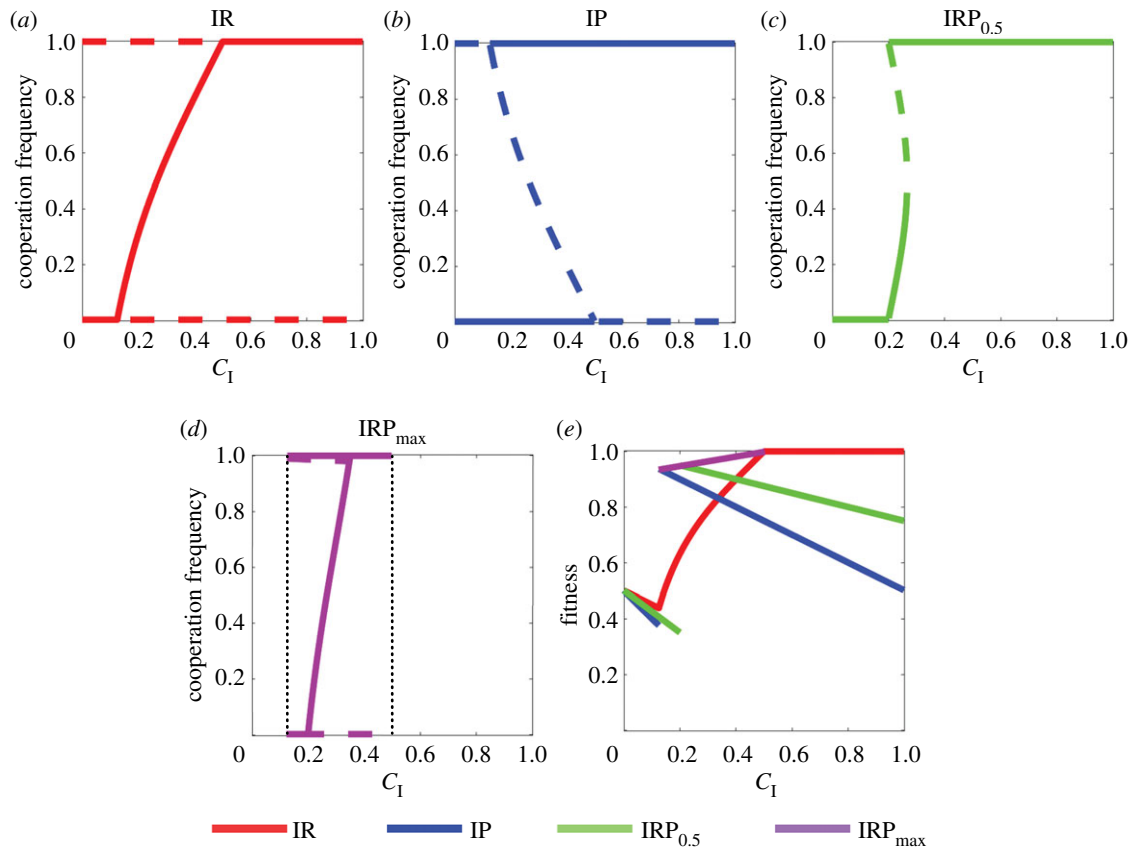
We consider that individuals update their strategies based on the preferential imitation of more successful strategies, where a strategy with a higher fitness is expected to spread. This process can be described by the following replicator equation:

$$\frac{dx}{dt} = x(1-x)(f_C - f_D). \quad (2.1)$$

We now allow decision errors and assume that an individual adopts a random strategy with probability  $\mu$ . Equation (2.1) can then be modified as

$$\frac{dx}{dt} = x(1-x)(f_C - f_D) + \frac{\mu}{2}(1-x) - \frac{\mu}{2}x. \quad (2.2)$$

Equation (2.2) is called the replicator-mutator equation [54]. Biologically, this equation describes a simultaneous action of selection and mutation. Specifically, the two effects are independent, where the first term  $x(1-x)(f_C - f_D)$  is the selection term and the second term  $\mu/2(1-x) - \mu/2x$  is the mutation term [54].



**Figure 1.** PGG without decision errors. Parameters are taken as  $n = 4$ ,  $c = 1$ ,  $r = 2$ ,  $\omega = 0.5$ , and  $\mu = 0$ . (a)–(d) Equilibria of equation (2.1) with different types of incentives, where stable equilibria are denoted by solid curves and unstable equilibria are denoted by dashed curves. (a) IR has a globally stable equilibrium. (b) For IP, the cooperative and the defective equilibria are bistable. (c) IRP may have two interior equilibria in addition to the two boundary equilibria. (d) IRP<sub>max</sub> is defined for  $0.125 < C_1 < 0.5$  (see the dotted lines), and the cooperative equilibrium is stable in this interval. (e) Group average fitness at the stable (cooperative) equilibria for different types of incentives. Overall, the reward can maximize the group average fitness for small and large incentives, and a mixture of reward and punishment performs best for intermediate incentives. (Online version in colour.)

### 3. Results

#### (a) Without decision errors

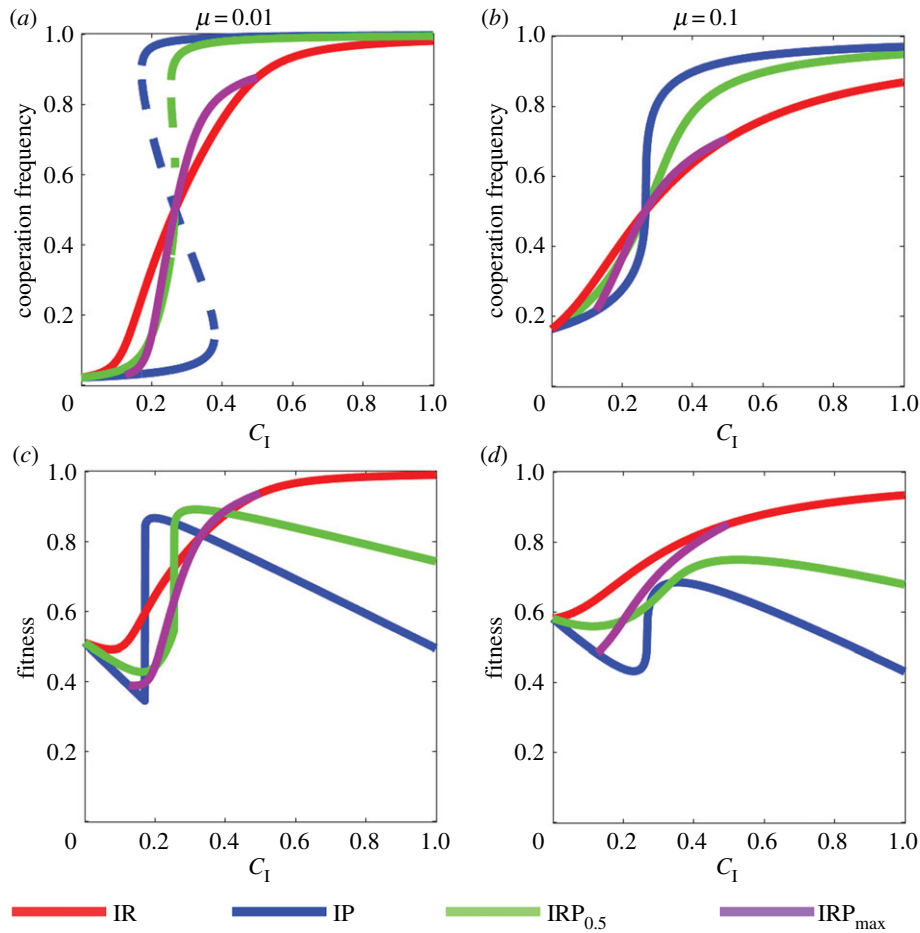
If players do not make a mistake, both the defective state  $x = 0$  and the cooperative state  $x = 1$  are equilibria for equation (2.1). In particular, the defective equilibrium is globally stable when there is no incentive, which corresponds to the state that all individuals in the population are defectors.

We first analyse the stability of equation (2.1) for different types of incentives (see electronic supplementary material, §1.2 for equilibria calculation and stability analysis). The cases of IR and IP were previously investigated in [47]. For both cases, if  $C_1$  is greater than an upper bound  $C_1^+ = c - (rc)/n$ , then the existence of an interior equilibrium is impossible, and the cooperative equilibrium  $x = 1$  is globally stable. By contrast, if  $C_1$  is smaller than a lower bound  $C_1^- = c/n - (rc)/(n^2)$ , then the defective equilibrium  $x = 0$  is globally stable. Therefore, we focus mainly on  $C_1^- < C_1 < C_1^+$  in later discussions, i.e. the incentive size is intermediate in the sense that it has some effects on promoting cooperation but is not sufficient to guarantee the global stability of the cooperative equilibrium. For IR with intermediate incentive, both the cooperative and the defective equilibria are unstable, and equation (2.1) has a globally stable interior equilibrium  $x^*_R$  (see figure 1a). For IP with intermediate incentive, the cooperative and the defective equilibria are bistable, and equation (2.1) has an unstable interior equilibrium  $x^*_P$  (see figure 1b). Finally, for IRP with intermediate incentive, equation (2.1) may have (at most) two

interior equilibria in addition to the two boundary equilibria (figure 1c). Furthermore, the cooperative equilibrium is stable if and only if  $C_1 > ((n - r)c)/(n(\alpha + (1 - \alpha)n))$ .

We next look at the efficiencies of the three types of incentives by comparing their average fitness at the stable (cooperative) equilibria. First, although IP can maintain full contribution, whether this strategy is more efficient than IR depends on the incentive size  $C_1$ . In fact, in IP, the average fitness is maximized when  $C_1$  barely exceeds the lower bound  $C_1^-$ , while in IR, the maximum fitness is achieved when  $C_1$  reaches the upper bound  $C_1^+$  [46]. Second, IRP is more efficient than IP whenever the cooperative equilibrium is stable. This result implies that a mixture of reward and punishment works better than punishment only. Finally, whether IRP has a higher fitness than IR depends on both  $C_1$  and  $\alpha$  (see figure 1d). For intermediate incentive, the fitness of IRP is maximized when  $\alpha_{\max} = (n - (n - r)c/nC_1)/(n - 1)$ , i.e. the maximum  $\alpha$  such that the cooperative equilibrium is stable. With this  $\alpha_{\max}$ , IRP must be more efficient than IR (see figure 1e).

We now show the optimal incentive that maximizes the group average fitness at the stable (cooperative) equilibrium. If  $C_1 \leq C_1^-$ , then the defective state is the only stable equilibrium. In this case, the optimal incentive is IR because the use of punishment decreases the social welfare. On the other hand, if  $C_1 \geq C_1^+$ , then the cooperative state is the only stable equilibrium. The optimal incentive in this case is also IR because the punishment institution is costly even if no one is punished. Finally, if  $C_1^- < C_1 < C_1^+$ , then the optimal



**Figure 2.** PGG with decision errors. Parameters are taken as  $n = 4$ ,  $c = 1$ ,  $r = 2$ ,  $\omega = 0.5$ . (a),(b) Equilibria of equation (2.2) with different types of incentives and decision errors, where stable equilibria are denoted by solid curves and unstable equilibria are denoted by dashed curves.  $IRP_{max}$  is defined for  $0.125 < C_1 < 0.5$ . (a) For small errors, IR and  $IRP_{max}$  have a unique stable equilibrium, and IP and  $IRP_{0.5}$  may have three equilibria. (b) For intermediate errors, all the incentives have a unique stable equilibrium. (c),(d) Group average fitness at the stable (cooperative) equilibria for different types of incentives and decision errors. The optimal incentive is sensitive to decision errors. (c) For small errors, the incentives that lead to the highest group average fitness are, from  $C_1 = 0$  to 1, IR, IP,  $IRP_{0.5}$ ,  $IRP_{max}$  and IR. (d) For intermediate errors, IR performs best for all  $C_1$ . (Online version in colour.)

incentive is IRP with  $\alpha_{max}$  (denoted by  $IRP_{max}$ ). Overall, reward alone is the most efficient incentive for both small and large  $C_1$ , and a mixture of reward and punishment is the most efficient incentive for intermediate  $C_1$  (see figure 1e).

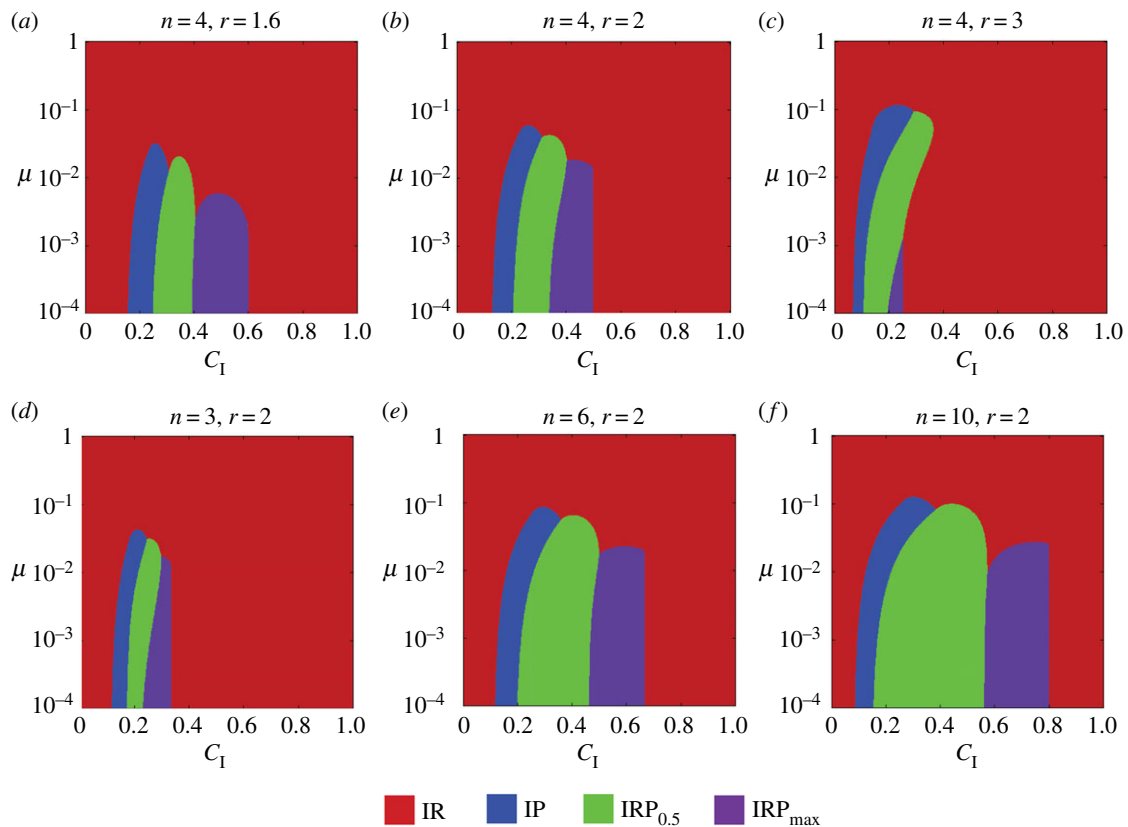
### (b) With decision errors

Decision errors can dramatically change the evolutionary process. Foremost, when subjects can make mistakes, the cooperative state and the defective state are no longer in equilibria. As a result, cooperators and defectors stably coexist, and no incentive can maintain full cooperation. Second, equation (2.2) has a unique and globally stable interior equilibrium for larger  $\mu$  and smaller  $\omega$ . In particular, a crucial  $C_1^*$  exists such that IR leads to the highest cooperation level for  $C_1 < C_1^*$  and that IP is the most effective incentive in promoting cooperation for  $C_1 > C_1^*$  (see figure 2b; electronic supplementary material, § 1.3).

We further study the properties of the equilibria of equation (2.2) through numerical simulations (see electronic supplementary material, § 3.1 for simulation codes). Figures 2a,b show equilibria of equation (2.2) for different types of incentives under small and intermediate decision errors. On the one hand, both IR and IP are robust to small  $\mu$ , i.e. IR has a globally stable equilibrium for all  $C_1$  and IP has two stable equilibria and one unstable equilibrium for intermediate  $C_1$ . However, IRP is sensitive to  $\mu$ . As shown

in figure 2b,  $IRP_{max}$  is more like IR with its unique equilibrium, and IRP with  $\alpha = 0.5$  (denoted by  $IRP_{0.5}$ ) is more like IP, which can have three equilibria. On the other hand, all the incentives have a unique globally stable equilibrium for intermediate  $\mu$ . Specifically,  $C_1$  plays a predominant role in determining the equilibrium value, where the cooperation rate is continuously increasing in  $C_1$ . Finally, IRP is never superior in promoting cooperation; rather, it is dominated by IR for smaller  $C_1$  and by IP for larger  $C_1$ .

We next compare the group average fitness for different types of incentives. For small decision errors  $\mu = 0.01$ , the optimal incentive is IR for small and large  $C_1$  and is IP or IRP for intermediate  $C_1$  (see figure 2c). Notably, as  $\mu$  increases to 0.1, IR becomes the most efficient incentive for all  $C_1$  (see figure 2d). Figure 3 systematically illustrates the effect of decision errors on incentive efficiencies, indicating that the above observations are robust against a variety of  $n$  and  $r$ . Overall, for small decision errors, the optimal incentives are, from  $C_1 = 0$  to 1, IR, IP,  $IRP_{0.5}$ ,  $IRP_{max}$ , and IR. If we exclude  $IRP_{max}$  then this order is consistent with the case of  $\mu = 0$ . This result reveals that the efficiencies of IR, IP, and IRP with constant  $\alpha$  are stable against small mutations. By contrast,  $IRP_{max}$  can no longer maintain high fitness even for small decision errors. Finally, for intermediate and larger decision errors, IR leads to the highest group average fitness for all  $C_1$ ,  $n$  and  $r$ .



**Figure 3.** Optimal incentives for different  $C_1$ ,  $\mu$ ,  $n$ , and  $r$ .  $c = 1$  and  $\omega = 0.5$  in all six subfigures. IR, IP,  $IRP_{0.5}$ , and  $IRP_{max}$  lead to the highest group average fitness in the red, blue, green, and purple regions, respectively. For small errors, the incentives that lead to the highest group average fitness are, from  $C_1 = 0$  to 1, IR, IP,  $IRP_{0.5}$ ,  $IRP_{max}$ , and IR. For intermediate and large errors, IR is the optimal incentive for all  $C_1$ ,  $n$  and  $r$ .

### (c) Robustness analysis

In electronic supplementary material, §2, we evaluate the robustness of the results with respect to the following model variants. (i) We consider an ‘others only’ variant of the PGG (also called a ‘mutual aid’ game) in which a cooperator does not benefit from its own contribution [4,43,46,56–58]. (ii) The incentive institution can be modelled in different ways. On the one hand, one may assume that the total amount of reward is multiplied by a factor  $r_1$  before it is distributed among cooperators, and the total amount of punishment is multiplied by a factor  $r_2$  before it is distributed among defectors [33,35]. On the other hand, the incentive could be probabilistic where only one cooperator (or defector) is exemplarily rewarded (punished) [36,40,41,44]. (iii) The institution makes a mistake in incentive distribution [47]. For instance, the institution may punish a cooperator or reward a defector due to a lack of information or observation errors. We find that none of the variants (i)–(iii) qualitatively affect our results regarding the equilibria of the evolutionary dynamics and the efficiencies of the different types of incentives. For intermediate or large decision errors, the optimal incentive is IR for all  $C_1$ .

## 4. Discussion

Reward and punishment are often used to promote cooperation in social dilemmas. Most of the previous studies have focused on which type of incentive best promotes cooperation (see, e.g. [6–8,33]). However, a high cooperation level does not always lead to high fitness. For instance, punishment can effectively promote cooperation, but experiments

rarely find a significant increase in net payoff in that context [11,18,23]. In this paper, we argue that a better measurement for the success of an incentive is the group average fitness at the evolutionary stable state. The reason is twofold: on the one hand, the group average fitness plays a predominant role in the evolutionary process. In fact, only groups with higher fitness can win the competition for survival, and groups with lower fitness tend to go extinct. On the other hand, if players update strategies based on payoff-driven decision rules such as ‘best-response’ or ‘imitate-the-better’ (these two decisions rules have been widely used in the previous studies, e.g. [4,43–49]), then the players should also prefer the incentive institution that can lead to a higher income rather than a higher cooperation level.

In this paper, we compare the efficiencies of three types of institutional incentive, namely, reward, punishment, and a mixture of reward and punishment. We emphasize that players are boundedly rational and may make mistakes in decision-making. Recent studies in neuroscience revealed that errors are inevitable in human decision-making, even reward and punishment cannot eliminate unintended errors [59]. Therefore, it is important to investigate how decision errors affect the efficiencies of different types of incentives. By analysing the group average fitness at the stable equilibria of the replicator-mutator equation, we find that the optimal incentive is sensitive to decision errors. Without decision errors, the reward can maximize the group average fitness for small and large incentives, and a mixture of reward and punishment performs best for intermediate incentives. However, if players make mistakes in decision-making, then the use of punishment almost always reduces the group welfare, and the reward is revealed as the most efficient incentive for intermediate and large decision

errors. Numerical simulations show that this result is robust in the standard PGG and alternative model variants.

The failure of IP and IRP is caused by two reasons. On the one hand, these strategies can no longer maintain a high cooperation level when error exists. As shown in figures 1 and 2, the stable equilibria in IP and IRP with  $\mu = 0.1$  are different from that of  $\mu = 0$  and  $\mu = 0.01$ . For smaller  $C_I$ , the cooperation levels in IP and IRP are always lower than that of IR. On the other hand, punishing defectors significantly reduces the group fitness. This problem is not severe when players do not make mistakes. Because in this case, full cooperation can be stable in IP and IRP, and no one is punished. However, for intermediate or large decision errors, the proportion of defectors is not rare, and the loss of punishment is typically inevitable.

We note that the effects of institutional incentives and peer incentives in promoting group fitness are substantially different. On the one hand, intermediate errors can increase cooperation and improve the group average fitness in the PGG with peer punishment [52,60]. However, errors are harmful to institutional punishment. On the other hand, (budget-balanced) peer reward can never lead to a high cooperation level [4,26,30], but the institutional reward can promote cooperation regardless of the presence of errors. This result implies that institutional rewards have advantages over other types of incentives, as they can promote cooperation without hurting the group welfare and are robust against decision errors.

In our model, every player has to pay for the incentive institution before contributing to the PGG. In fact, this payment can be seen as an entry fee for the PGG. If the payment is voluntary, then the issue of second-order free-riding is raised because the incentive institution itself is a common good that can be exploited. Previous studies show that IP with voluntary payment is functional if punishment is also imposed on second-order free-riders (i.e. subjects who cooperate but do not pay for the institution) [39,42,43]. By contrast, if IR is budget-balanced, then rational players will not pay for the institution [35]. Since the entry fee in our model is compulsory, a subsequent question is whether real players will choose such an incentive institution. A recent experimental study indicated that approximately half of the subjects would like to pay 20% of their wealth for a reward mechanism before playing the PGG

[61]. This finding suggests that the incentive institution we proposed could be sustainable.

In addition, the incentive institution considered in this paper is not adaptive, i.e.  $\alpha$  does not change over the cooperation level. Notably, a ‘first carrot, then stick’ (FCTS) control of rewards and punishments promotes cooperation best [47]. According to FCTS, the incentive institution should provide a reward if the cooperation level is lower than 0.5 and provide punishment if the cooperation level is higher than 0.5. Our study confirms that this adaptive institution also works when players make mistakes, because reward promotes cooperation best in a selfish population and punishment performs best in an altruistic population (see figure 2*a,b*). However, FCTS does not always maximize the group average fitness. For intermediate and large  $\mu$ , IR is the most efficient incentive even though IP promotes cooperation better than IR in a cooperative population.

In summary, our research deepens the understanding into the influences of reward and punishment in curbing selfishness and improving social welfare. Previous research indicated that institutional punishment is a cheaper and more reliable way of inducing cooperation than is institutional reward [44,46]. However, when individuals are boundedly rational and may make mistakes, we show that institutional reward is not only effective but also efficient. This finding demonstrates the competitive advantage of reward institution in the evolutionary process and highlights the role of positive incentives in promoting cooperation. Our results also suggest that in mechanism design, uncertainty should be taken into account because the optimal incentive could be sensitive to decision errors.

**Data accessibility.** The simulation code supporting this study has been uploaded as part of the electronic supplementary material in pdf formats.

**Authors' contributions.** Y.D., T.S., and B.Z. all contributed to the development of the original idea and the conception of the mathematical model, B.Z. wrote the first draft of the manuscript and all authors contributed substantially to revisions.

**Competing interests.** We declare that we have no competing interests.

**Funding.** Y.D. and B.Z. acknowledge support by the National Science Foundation of China (grant no. 71771026). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References

- Pennisi E. 2011 How did cooperative behavior evolve? *Science* **309**, 93–93. (doi:10.1126/science.309.5731.93)
- Hamilton WD. 1964 The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–52. (doi:10.1016/0022-5193(64)90038-4)
- Trivers RL. 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)
- Sigmund K. 2010 *The calculus of selfishness*. Princeton, NJ: Princeton University Press.
- Clutton-Brock TH, Parker GA. 1995 Punishment in animal societies. *Nature* **373**, 209–216. (doi:10.1038/373209a0)
- Balliet D, Mulder LB, Van Lange PAM. 2011 Reward, punishment, and cooperation: a meta-analysis. *Psychol. Bull.* **137**, 594–615. (doi:10.1037/a0023489)
- Guala F. 2012 Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *Behav. Brain Sci.* **35**, 1–15. (doi:10.1017/S0140525X11000069)
- Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, Szolnoki A. 2017 Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51. (doi:10.1016/j.physrep.2017.05.004)
- Fehr E, Gächter S. 2000 Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994. (doi:10.1257/aer.90.4.980)
- Andreoni J, Harbaugh WT, Vesterlund L. 2003 The carrot or the stick: rewards, punishments and cooperation. *Am. Econ. Rev.* **93**, 893–902. (doi:10.1257/00022803322157142)
- Fehr E, Fischbacher U. 2003 The nature of human altruism. *Nature* **425**, 785–791. (doi:10.1038/nature02043)
- Fehr E, Rockenbach B. 2003 Detrimental effects of sanctions on human altruism. *Nature* **422**, 137–140. (doi:10.1038/nature01474)
- Walker JM, Halloran M. 2004 Rewards and sanctions and the provision of public goods in one-shot settings. *Exp. Econ.* **7**, 235–247. (doi:10.1023/B:EXEC.0000040559.08652.51)
- Cinyabuguma M, Page T, Putterman L. 2006 Can second-order punishment deter perverse punishment. *Exp. Econ.* **9**, 265–279. (doi:10.1007/s10683-006-9127-z)
- Denant-Boemont L, Masclet D, Noussair C. 2007 Punishment, counterpunishment and sanction

- enforcement in a social dilemma experiment. *Econ. Theory* **33**, 145–167. (doi:10.1007/s00199-007-0212-0)
16. Gurek O, Irlenbusch B, Rockenbach B. 2006 The competitive advantage of sanctioning institutions. *Science* **312**, 108–111. (doi:10.1126/science.1123633)
  17. Sefton M, Schupp R, Walker M. 2007 The effect of rewards and sanctions in provision of public goods. *Econ. Inq.* **45**, 671–690. (doi:10.1111/ecin.2007.45.issue-4)
  18. Dreber A, Rand DG, Fudenberg D, Nowak MA. 2008 Winners don't punish. *Nature* **452**, 348–351. (doi:10.1038/nature06723)
  19. Egas M, Riedl A. 2008 The economics of altruistic punishment and the maintenance of cooperation. *Proc. R. Soc. B* **275**, 871–878. (doi:10.1098/rspb.2007.1558)
  20. Herrmann B, Thoni C, Gächter S. 2008 Antisocial punishment across societies. *Science* **319**, 1362–1367. (doi:10.1126/science.1153808)
  21. Gächter S, Renner E, Sefton M. 2008 The long-run benefits of punishment. *Science* **322**, 1510–1510. (doi:10.1126/science.1164744)
  22. Nikiforakis N. 2008 Punishment and counter-punishment in public good games: can we really govern ourselves? *J. Public Econ.* **92**, 91–112. (doi:10.1016/j.jpubeco.2007.04.008)
  23. Rand DG, Dreber A, Ellingsen T, Fudenberg D, Nowak MA. 2009 Positive interactions promote public cooperation. *Science* **325**, 1272–1275. (doi:10.1126/science.1177418)
  24. Wu JJ, Zhang B, Zhou ZX, He QQ, Zheng XD, Cressman R, Tao Y. 2009 Costly punishment does not always increase cooperation. *Proc. Natl Acad. Sci. USA* **106**, 17 448–17 451. (doi:10.1073/pnas.0905918106)
  25. Rand DG, Nowak MA. 2011 The evolution of antisocial punishment in optional public goods games. *Nat. Commun.* **2**, 434. (doi:10.1038/ncomms1442)
  26. Sigmund K, Hauert C, Nowak MA. 2001 Reward and punishment. *Proc. Natl Acad. Sci. USA* **98**, 10 757–10 762. (doi:10.1073/pnas.161155698)
  27. Boyd R, Gintis H, Bowles S, Richerson PJ. 2003 The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* **100**, 3531–3535. (doi:10.1073/pnas.0630443100)
  28. Fowler JH. 2005 Altruistic punishment and the origin of cooperation. *Proc. Natl Acad. Sci. USA* **102**, 7047–7049. (doi:10.1073/pnas.0500938102)
  29. Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K. 2007 Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907. (doi:10.1126/science.1141588)
  30. Hilbe C, Sigmund K. 2010 Incentives and opportunism: from the carrot to the stick. *Proc. R. Soc. B* **277**, 2427–2433. (doi:10.1098/rspb.2010.0065)
  31. Hilbe C, Traulsen A. 2012 Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Sci. Rep.* **2**, 458. (doi:10.1038/srep00458)
  32. dos Santos M. 2015 The evolution of anti-social rewarding and its countermeasures in public goods games. *Proc. R. Soc. B* **282**, 20141994. (doi:10.1098/rspb.2014.1994)
  33. Gao L, Wang Z, Pansini R, Li YT, Wang RW. 2015 Collective punishment is more effective than collective reward for promoting cooperation. *Sci. Rep.* **5**, 17 752. (doi:10.1038/srep17752)
  34. Okada I, Yamamoto H, Toriumi F, Sasaki T. 2015 The effect of incentives and meta-incentives on the evolution of cooperation. *PLoS Comput. Biol.* **11**, e1004232. (doi:10.1371/journal.pcbi.1004232)
  35. Szolnoki A, Perc M. 2015 Antisocial pool rewarding does not deter public cooperation. *Proc. R. Soc. B* **282**, 20151975. (doi:10.1098/rspb.2015.1975)
  36. Yamagishi T. 1986 The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.* **51**, 110–116. (doi:10.1037/0022-3514.51.1.110)
  37. Sutter M, Haigner S, Kocher MG. 2010 Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Rev. Econ. Stud.* **77**, 1540–1566. (doi:10.1111/roes.2010.77.issue-4)
  38. Putterman L, Tyran JL, Kamei K. 2011 Public goods and voting on formal sanction schemes. *J. Public Econ.* **95**, 1213–1222. (doi:10.1016/j.jpubeco.2011.05.001)
  39. Traulsen A, Röhl T, Milinski M. 2012 An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proc. R. Soc. B* **279**, 3716–3721. (doi:10.1098/rspb.2012.0937)
  40. Cressman R, Wu JJ, Li C, Tao Y. 2013 Game experiments on cooperation through reward and punishment. *Biol. Theory* **8**, 158–166. (doi:10.1007/s13752-013-0106-2)
  41. Wu JJ, Li C, Zhang B, Cressman R, Tao Y. 2014 The role of institutional incentives and the exemplar in promoting cooperation. *Sci. Rep.* **4**, 6421. (doi:10.1038/srep06421)
  42. Zhang B, Li C, De Silva H, Bednarik P, Sigmund K. 2014 The evolution of sanctioning institutions: an experimental approach to the social contract. *Exp. Econ.* **17**, 285–303. (doi:10.1007/s10683-013-9375-7)
  43. Sigmund K, De-Silva H, Traulsen A, Hauert C. 2010 Social learning promotes institutions for governing the commons. *Nature* **466**, 861–863. (doi:10.1038/nature09203)
  44. Cressman R, Song JW, Zhang B, Tao Y. 2012 Cooperation and evolutionary dynamics in the public goods game with institutional incentives. *J. Theor. Biol.* **299**, 144–151. (doi:10.1016/j.jtbi.2011.07.030)
  45. Perc M. 2012 Sustainable institutionalized punishment requires elimination of second-order free-riders. *Sci. Rep.* **2**, 344. (doi:10.1038/srep00344)
  46. Sasaki T, Brännström A, Dieckmann U, Sigmund K. 2012 The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proc. Natl Acad. Sci. USA* **109**, 1165–1169. (doi:10.1073/pnas.1115219109)
  47. Chen X, Sasaki T, Brännström Å, Dieckmann U. 2015 First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* **12**, 20140935. (doi:10.1098/rsif.2014.0935)
  48. Schoenmakers S, Hilbe C, Blasius B, Traulsen A. 2014 Sanctions as honest signals - the evolution of pool punishment by public sanctioning institutions. *J. Theor. Biol.* **356**, 36–46. (doi:10.1016/j.jtbi.2014.04.019)
  49. Sasaki T, Uchida S, Chen X. 2015 Voluntary rewards mediate the evolution of pool punishment for maintaining public goods in large populations. *Sci. Rep.* **5**, 8917. (doi:10.1038/srep08917)
  50. Oya G, Ohstuki H. 2017 Stable polymorphism of cooperators and punishers in a public goods game. *J. Theor. Biol.* **419**, 243–253. (doi:10.1016/j.jtbi.2016.11.012)
  51. Szolnoki A, Perc M. 2017 Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment. *Phys. Rev. X* **7**, 041027. (doi:10.1103/PhysRevX.7.041027)
  52. Traulsen A, Hauert C, De Silva H, Nowak MA, Sigmund K. 2009 Exploration dynamics in evolutionary games. *Proc. Natl Acad. Sci. USA* **106**, 709–712. (doi:10.1073/pnas.0808450106)
  53. Traulsen A, Semmann D, Sommerfeld RD, Krambeck HJ, Milinski M. 2010 Human strategy updating in evolutionary games. *Proc. Natl Acad. Sci. USA* **107**, 2962–2966. (doi:10.1073/pnas.0912515107)
  54. Hofbauer J, Sigmund K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.
  55. Nowak MA. 2006 *Evolutionary dynamics: exploring the equations of life*. Cambridge, UK: Harvard University Press.
  56. Wilson DS. 1975 A theory of group selection. *Proc. Natl Acad. Sci. USA* **72**, 143–146. (doi:10.1073/pnas.72.1.143)
  57. Sugden R. 1986 *The economics of rights, cooperation and welfare*. Oxford, UK: Blackwell.
  58. Fletcher JA, Zwick M. 2004 Strong altruism can evolve in randomly formed groups. *J. Theor. Biol.* **228**, 303–313. (doi:10.1016/j.jtbi.2004.01.004)
  59. Glimcher PW. 2011 *Foundations of neuroeconomic analysis*. New York, NY: Oxford University Press.
  60. Zhang B, Hofbauer J. 2016 Quantal response methods for equilibrium selection in  $2 \times 2$  coordination games. *Games Econ. Behav.* **97**, 19–31. (doi:10.1016/j.geb.2016.03.002)
  61. Yang C, Zhang B, Charness G, Li C, Lien JW. 2018 Endogenous rewards promote cooperation. *Proc. Natl Acad. Sci. USA* **115**, 9968–9973. (doi:10.1073/pnas.1808241115)