Physica A 605 (2022) 127997

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Costly signals can facilitate cooperation and punishment in the prisoner's dilemma

Kaixuan Guan [a], Yuyou Chen [b,d,c,*], Wanjun Zheng [b], Lulu Zeng [b], Hang Ye [b]

[a] Postdoctoral Research Center, Industrial and Commercial Bank of China, Beijing 100140, China
[b] School of Economics, Center for Economic Behavior and Decision-making (CEBD), Zhejiang University of Finance and Economics, Hangzhou 310018, China
[c] University of Chinese Academy of Science, Beijing 100049, China
[d] Institutes of Science and Development, Chinese Academy of Science, Beijing 100190, China

ARTICLE INFO

Article history:
Received 23 February 2022
Received in revised form 3 June 2022
Available online 6 August 2022

Keywords:
Cooperation
Punishment
Evolutionary game theory
Costly signaling theory
The prisoner's dilemma

ABSTRACT

Costly signal is regarded as one of the mechanisms to explain the emergence of cooperation. Both the cost of cooperation and the cost of punishment can be seen as expensive cost signals. Previous research has used costly signaling theory to explain the establishment of cooperation and punishment in public goods game. However, punishment is less likely to emerge stably in the prisoner's dilemma, and the punisher cannot have additional information to identify and punish the defector. Therefore, it is particularly important to further study whether the expensive cost signal in the prisoner's dilemma can promote the emergence of cooperation and punishment. We distinguish between costly punishing signals and costly cooperative signals in this paper to look at the rule that turns payoff into fitness. The findings reveal that, without punishment, if the costly signal is weak and the cost of cooperation is not too high compared to the benefit of cooperation, cooperation is a better choice than defection. With punishment, if there is a small amount of noise in the costly signaling mechanism and punishment is considered a more expensive signal than cooperation, punishment is a better strategy.

1. Introduction

The evolution of cooperation has been widely studied [1–3]. A number of mechanisms have been proposed to explain the establishment of cooperation. Kin selection is often used to explain cooperation between close relatives [4]. Group selection can explain cooperation in relatively close groups [5]. Direct reciprocity is a continuous game between pairs of individuals that leads to cooperation [6]. Indirect reciprocity promotes cooperation by indirectly influencing the behavior of individuals in the group through the behavior of others in the group [7], as in the reputation effect [8]. In structured populations, certain dynamic rules can promote cooperation through network reciprocity [9–12]. The tagging mechanism can induce individuals to establish cooperation based on tags without reciprocity [13]. It also allows individuals to freely enter and exit the game of public goods and gives individuals relatively stable income when exiting, which can

* Corresponding author at: School of Economics, Center for Economic Behavior and Decision-making (CEBD), Zhejiang University of Finance and Economics, Hangzhou 310018, China.
E-mail addresses: chenyuyou@126.com, chenyuyou@zufe.edu.cn (Y. Chen).

https://doi.org/10.1016/j.physa.2022.127997
0378-4371/

promote cooperation [14]. If cooperation has been established by other mechanisms, punishment can increase the level of cooperation [15]. Rewards have also been shown to positively promote cooperation [16]. Signaling can boost the evolution of cooperation in repeated group interactions [17]. In addition, there are other mechanisms that can improve the level of cooperation, such as social diversity [18,19], commitment [20], resource-based conditional interaction [21], age [22], reputation [23,24], geographical distance [25], knowledge of past information [26], multiple games [27], freedom of choice [28], stochastic interactions [29], and information sharing [30,31].

Punishment is considered a negative reciprocity mechanism that promotes cooperation [32], and usually the punisher improves the level of cooperation by reducing his own payoff to reduce the payoff of the defector in the game. Cooperators have a strong incentive to punish defectors [33]. In the prisoner's dilemma, altruistic punishment is seen as a way to combat defects [34], but it is not considered a mechanism of cooperative evolution [2] because altruistic punishment itself is an evolutionary mechanism to be explained [35,36] and a direct or indirect way of reciprocity [37]. It has been shown that the evolution of altruistic punishment depends on group selection rather than individual selection [38,39]. Several studies have shown that punishment promotes cooperation in n-player games [40–43]. In the two-person prisoner's dilemma, compared to the case without punishment, although the punishment promotes cooperation, the punisher achieves lower returns than the nonpunisher [37], which makes it difficult for the punisher to breed offspring. Another study concludes that if the cost of being punished is large enough, the cooperation and punishment strategies can evolve stably [35]. However, in the absence of other mechanisms, costly punishment is not conducive to individuals adopting punishment strategies, so in direct reciprocal cooperative games, the winner will not choose costly punishment strategies [37], which leads to the second-order social dilemma; that is, the cooperator who does not punish others has a higher fitness than the punisher. If this second-order social dilemma is established, it will make it difficult for altruistic punishment to evolve, but altruistic punishment exists in reality. At this point, it is easy to think of a question: how did the behavior of altruistic punishment evolve? There is debate regarding what exactly punishment does, the mechanism it, and whether a costly signaling mechanism can be introduced to explain the two observed phenomena.

Costly signaling theory was originally derived from the handicap principle proposed by Israeli biologist A Zahavi [44]; that is, when choosing a mate, setting up a handicap to increase the cost to the other party can test the quality of the mate. Grafen then constructed a corresponding evolutionarily stable strategy model based on the handicap principle to explain the evolution of costly signals [45]. Signals are made expensive because high cost guarantees the honesty of signals, since costly signals are expensive to fake [44–46]. Costly signals are also considered to be an indirect reciprocal competitive act of altruism, in which individuals increase their chances of acquiring mates and allies by displaying individual qualities and resources as signals, which are expected to be more highly rewarded in the future [47]. Common behaviors that send costly signals are the following [47]: 1. Honest communication; 2. Public philanthropy; 3. Risk-taking and heroism; 4. Conspicuous consumption; 5. Religious commitment. Costly signals are also used to explain the emergence of cooperation. In the public goods game, individuals show their contributions in the form of costly signals to win potential allies or partners and give priority to cooperation with them [48–50] because long-term commitment-based partnerships are more durable than reciprocity-based relationships [51]. In essence, choosing someone who sends a costly signal as an ally or partner is somewhat altruistic.

We use an agent-based simulation approach to investigate two questions in this study. First, can cooperation or punishment signals among costly signals promote cooperation? Second, if costly signals promote cooperation, what factors influence the efficacy of costly signals, and how? To answer these questions, we transform the costly signal of cooperation and punishment into fitness as the rule to investigate the evolution of cooperation.

2. Model

According to the experiments designed by Dreber et al. [37], we consider three strategies in the prisoner's dilemma game(PDG): the cooperative strategy, the defective strategy and the punishment strategy. The cooperation and defective strategy are exactly the same as the cooperation and defective strategy in the classic prisoner's dilemma, and the punishment is to punish the other strategy unconditionally. The payoff matrix is shown below:

		Player B		
		C	D	P
Player A	C	$b-c$	$-c$	$-c-e$
	D	b	0	$-e$
	P	$b-f$	$-f$	$-f-e$

where b is the benefit from cooperation, c is the cost of cooperation, e is the loss of being punished, and f is the cost of punishment. When the cooperative strategy encounters other strategies, it will pay cooperation cost c. If the cooperative strategy encounters the cooperative strategy, it will also gain benefit b. If other strategies encounter the punishment strategy, all other strategies need to offer loss e. The defector gains benefit b when he encounters the cooperative strategy.The defector neither loses nor gains if he plays with the defector strategy.

This paper describes the game process with a punishment strategy in five stages.

The first stage is to randomly generate individuals to participate in the prisoner's dilemma game. We choose two individuals at random from a mixed finite population of size M to play the prisoner's dilemma game.

The second stage introduces the calculation method of the individual benefits of participating in the prisoner's dilemma game.

The third stage describes how to calculate evolutionary fitness. Individuals with higher returns usually have higher fitness, as assumed in previous studies; therefore, returns are usually directly converted into fitness, for example, according to the formula $F = e^{wR}$, where F is fitness, R is the return, and w represents choice intensity ($0 < w \leq 1$) [52]. However, from the perspective of costly signaling theory, the behavior that truly increases individual fitness is successful mating or alliance formation. In this case, the payoff alone does not determine fitness, and successful mating or alliance formation can improve fitness. An individual's fitness is determined by his or her benefits to transform the payoff into positive social attention as much as possible. Essentially, potential allies or partners care about how much they can benefit from potential allies or partners, so they need to observe and identify the partners that can benefit them. As a result, if people want to find a partner or allies to support them, they must send signals that indicate not only how much they are getting but also how much they are capable and ready to give to others. The costly signal is defined as the conversion rate between benefits and positive social attention ($0 < \tau < 1$), with $R \cdot \tau$ denoting the social attention holding power (SAHP), which is critical to human evolutionary success [53,54]. We suppose that the following are the rules for transforming payoff into fitness via costly signals: costly signals are a function of altruistic behavior, including cooperative behavior and punishing behavior. The stronger the individuals cooperate or punish, the more their payoff transforms into social attention retention. Therefore, the conversion rate function τ is:

$$\tau = 1 - e^{-(1-\alpha)\theta} \tag{1}$$

where α describes the transformation efficacy of a costly signaling system ($0 \leq \alpha < 1$), and greater α represents more noise. θ represents costly signals. Noise is caused by two distinct factors. At the beginning, individuals may pick their partners or allies for motivations other than altruism. Throughout the ovulation cycle, for example, research has revealed that women dynamically switch between two mate selection mechanisms (handsome or decent husband), in which decent husband uses costly signals to show his commitment to women and their children [55,56]. In addition, the heterogeneity in the weight assignments of different women to the two mechanisms of mate selection may be another source of noise. Furthermore, the ability to witness altruistic behavior is restricted, and any distortion in the transmission of altruistic behavior gossip will erode the costly signaling system's effectiveness. Thus, even if the costly signaling system is in favor of punishing as well as cooperating, defectors still have a chance to win in the population due to the cost of zero [55,56].

When α = 1, altruistic behavior as a costly signal has little effect on enhancing fitness. In this case, τ is $1 - e^{(-1)}$ for defectors, cooperators, or punishers, implying that fitness is determined only by the payoff. Then, this fitness is calculated in the same way as the simulation of the direct conversion from gain to fitness [57]. When $\alpha \to 0$, τ of the defector is 0, and for the cooperator and punisher, $\tau = 1 - e^{-\theta}$. Most research on the evolution of cooperation prior to cost signaling presumes this scenario, in which the punisher is indistinguishable from the cooperator. As a result, this study can be viewed as an attempt to combine the two mechanisms that keep people cooperating (punishment and cooperation costly signal) into a single framework. We use a linear function to represent the effect of altruistic action as a signal with a cost. We define $\theta = k \cdot f + c$, where f is the cost of punishment, c is the cost of cooperation, and k is the punishment signal transformation strength. There is no altruistic signal for defector because the defector pay nothing. This function can represent signals of three behaviors: 1. When $f = 0$ and $c = 0$, it means the strategy of defect, and there is no altruistic signal in this case. 2. When $f = 0$, $c > 0$, it indicates the strategy of cooperation, and the costly signal of cooperation is displayed in this case. 3. When $k > 1$, and $f > 0$, $c > 0$, the cost of punishment is converted into a signal, and the punishment signal is amplified. If $k > 0$, $k < 1$, and $c > 0$, it is a weakened punishment signal, and the cooperation cost signal is mainly a costly signal. At this point, our fitness algorithm in the simulation is as follows:

$$F = e^{wR(1-e^{-(1-\alpha)(kf+c)})} \tag{2}$$

The fourth stage is the replication or policy update stage. We use the imitation mechanism at this stage. The individual will compare its payoff with another randomly selected individual from the population. If its payoff is not higher than the other individual's payoff, then it will imitate the other individual's strategy. This replacement process can be explained as individual replication or imitation of behavior.

The mutation stage is the fifth stage. We suppose that any individual in the population has a slight possibility of switching to one of the other two types at random, independent of the payoff. The parameter μ is called the mutation rate.

The pseudo code of the simulation is as follows:

Algorithm The pseudocode for the simulation

```
Initialize M agents
for step do
    for agent do
        Other agent ← Random(agents)
        Play PDG
        Calculating fitness
    end for
    Agent ← Random(agents)
    Other agent← Random(agents)
    if Agent.payoff ≤ Other agent.payoff then
        Agent imitation Other agent
    end if
    Mutates
end for
```

3. Results

In order to ensure the validity of the results, all the simulation results in this paper are the mean value after running 500 times under the same parameters.

Without punishment ($f = 0, e = 0, k = 0$), the game payoff matrix depicts a classic prisoner's dilemma game. The altruism of the cooperators ($c = 15$) is in the form of a high cost. The results show that the defector dominates the population (the punisher and the defector are the same, Fig. 1a). The costly signaling mechanism is not active at this time. In Fig. 1b, the cost of punishment is 5 ($f = 5$). No punishment signal is transformed($k = 0$). And nobody is punished ($e = 0$). At this time, the game matrix describes the classic prisoner's dilemma with two types of cooperators who pay different cooperation costs. The cost paid by the punisher ($f = 5$, because no individual is punished, the punishment cost is the cooperation cost) is less than the cost paid by the cooperator ($f < c = 15$). The costly signals indicate altruistic cooperation which is not obvious. The cooperators is the lowest in the population, followed by the punishers, and the defectors occupy highest ratio in the population. As shown in Fig. 1c, if somebodies are punished($f = 5, e = 5$), and the signal of punishment is not transformed ($k = 0$), since all the individuals who play against the punisher are be punished, so whether it is a cooperator, a traitor, or a punisher, they get the same payoff owing to be reducing the same payoff. It is still the classic prisoner's dilemma game where there are two types of cooperators who pay different cooperation costs. The cooperators is also the lowest in the population, and the punisher is more than the cooperator, and the defectors is the highest strategy. In Fig. 1d, if the punishment happen ($f = 5$), and the signal of punishment is amplified ($k = 6$), and nobody is punished (e=0), so that the punisher pay for cost. The signal of punishment is greater than the signal of cooperation($k \cdot f > c$). The results show that the punisher the largest in the population, and the cooperator is the lowest strategy. In Fig. 1e, the punishment is 5 ($f = 5$), and the punishment signal is transformed and amplified (k=6), and the individual who plays game with the punisher is punished ($e = 5$). The signal sent by the punisher is greater than that by the cooperator ($k \cdot f > c$). The result is almost the same as in Fig. 1d. Based on this, the following conclusions can be drawn: Firstly, when the cost of cooperation is high, the costly signaling mechanism will not increase cooperation. Secondly, if no one is punished, the punisher has the same role as the cooperator. At this time, there is no costly punishment but costly cooperation. Only costly cooperation signals are transformed, which will not enhance cooperation. Thirdly, even if someone is punished, if the role of punishment and cooperation is the same, the mechanism do not promote the cooperation. Fourthly, If the costly signal of punishment is amplified by a certain number, no matter whether anyone is punished, the number of punishers will be increased, and the number of defects will decrease. Fifthly, no matter whether there is punishment or not, with the help of the costly signaling mechanism, there are only relatively dominant between different strategies.

To investigate the effect of noise in the costly signaling mechanism on cooperative behavior, we explored the changes in cooperation under different noise intense. Without punishment, as the noise increased, the cooperation decreased rapidly until 85% of individuals take the defective strategy (Fig. 2a). With punishment, if the cost of punishment cannot be converted into a costly signal, and the loss due to punishment, the frequency of cooperation will be decreased. But the frequency of punishment will not increase when the noise is 0 (Fig. 2b). Increasing in noise reduces the frequency of cooperation and increases the frequency of defect. If the cost of punishment is transformed into a costly signal, whether or not it is punished, only if the noise is not 1 ($a = 1$), the punishment will prevail (Figs. 2c and 2d). In the costly signal, the influence to cooperation, punishment and defect depend on whether the cost of punishment is converted into a cost signal. Only in the case of extremely high noise, the cost of punishment cannot be converted into costly signal to improve fitness. Therefore, we can draw the following conclusions: Firstly, when the punishment signal is not transformed, the change at low noise will change the frequencies of defect and cooperation in the population. Secondly, if the punishment signal is amplified by a certain multiple, the change at high noise will change the frequencies of defect and punishment in the population.

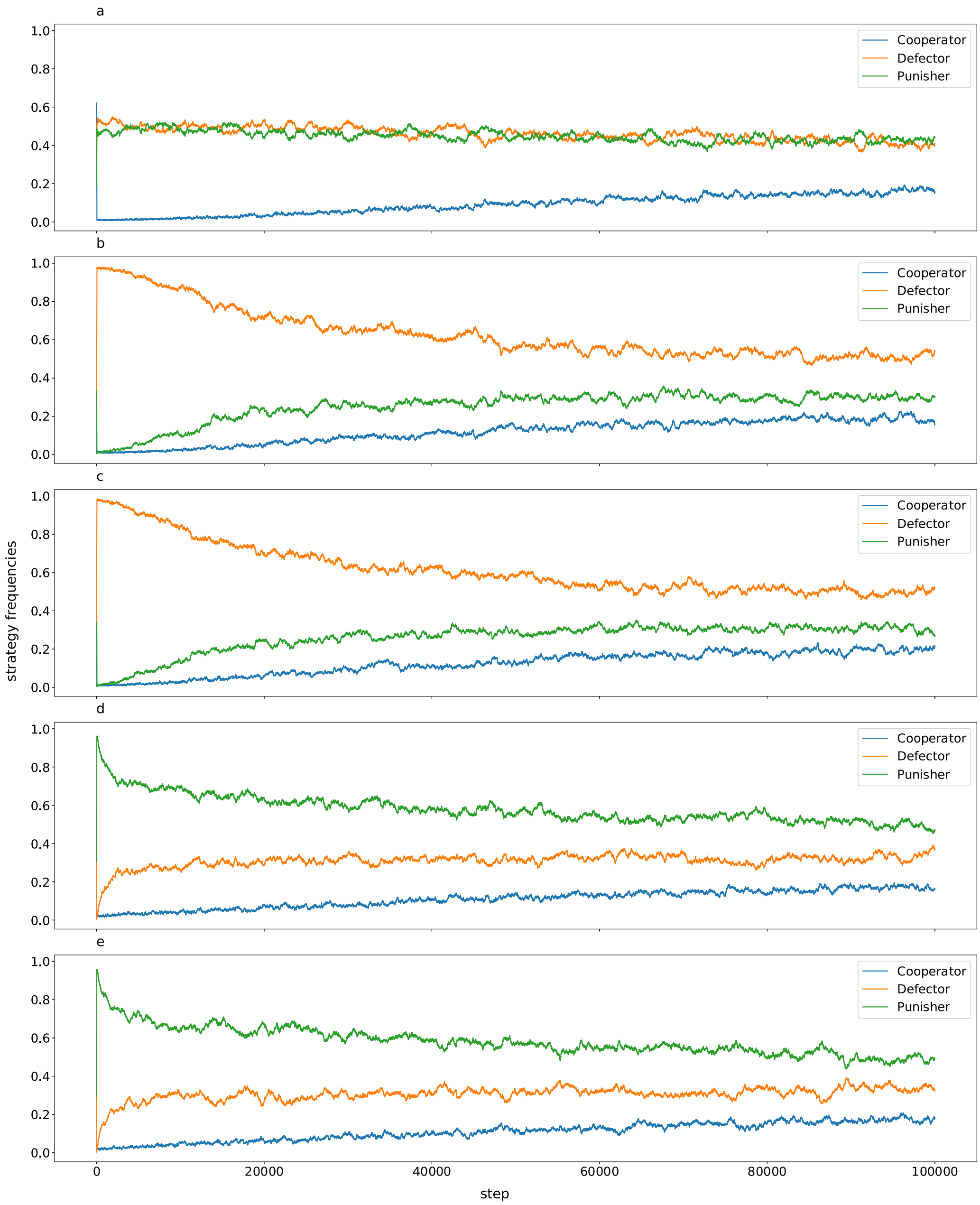

Fig. 1. Cooperative evolution in PDG. The simulated parameter values are M=100, $w = 1$, $\alpha = 0.1$, $b = 16$, $c = 15$, $\mu = 0.001$. (a) $k = 0$, $f = 0$, $e = 0$; (b) $k = 0$, $f = 5$, $e = 0$; (c) $k = 0$, $f = 5$, $e = 5$; (d) $k = 6$, $f = 5$, $e = 0$; (e) $k = 6$, $f = 5$, $e = 5$.

With punishment, a variety of factors will affect the ratio of cooperation, punishment and defect. Surprisingly, increasing the benefits of cooperation can increase the ratio of punishment and decrease the ratio of defect (Fig. 3a). Increasing the cost of cooperation reduces the proportion of cooperation and increases the proportion of punishment and defect (Fig. 3b). Increasing the punishment signal transformation strength will increase the proportion of punishment and decrease the proportion of defect (Fig. 3c). Increasing the cost of punishment also increases the proportion of punishment and decreases the proportion of defect (Fig. 3d). It can be seen from the above: Firstly, the costly signaling mechanism make defect dominate population at high cooperation costs and low cooperation, and make cooperation dominate population at low cooperation costs and high cooperation benefits. Secondly, the costly signaling mechanism makes high cooperation costs and high returns conducive to punishment. Thirdly, the costly signal makes the conversion intensity of the penalty signal and the increase of the punishment cost in the high cooperation cost both beneficial to the punishment, but not to the defect.

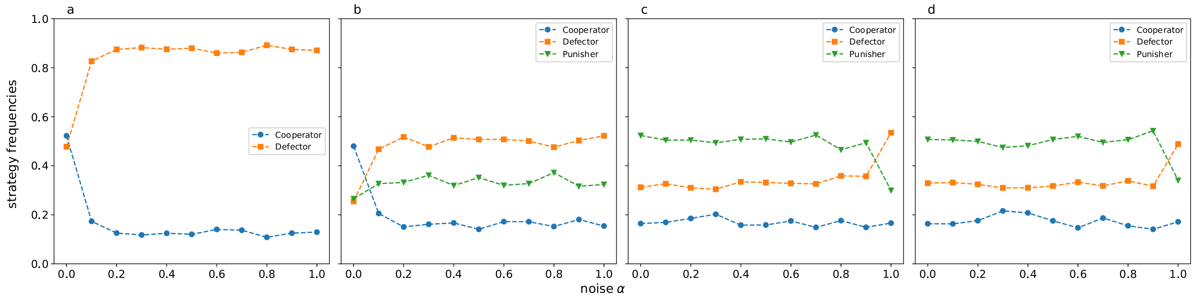

Fig. 2. The effect of noise in the costly signaling mechanism on the frequency of each strategy. The parameter values are M=100, $w = 1$, $b = 16$, $c = 15$, $\mu = 0.001$. (a) $k = 0$, $f = 0$, $e = 0$; (b) $k = 0$, $f = 5$, $e = 5$; (c) $k = 6$, $f = 5$, $e = 0$; (d) $k = 6$, $f = 5$, $e = 5$.

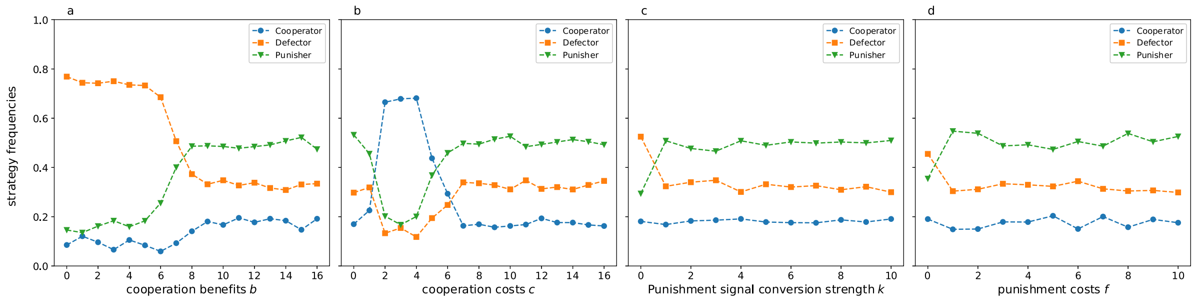

Fig. 3. The effect of punishment as costly signal in different initial combinations. The parameter values are M=100, $w = 1$, $\alpha = 0.1$, $e = 5$, $\mu = 0.001$. (a) $c = 15$, $k = 6$, $f = 5$; (b) $b = 16$, $k = 6$, $f = 5$; (c) $b = 16$, $c = 15$, $f = 5$; (d) $b = 16$, $c = 15$, $k = 6$.

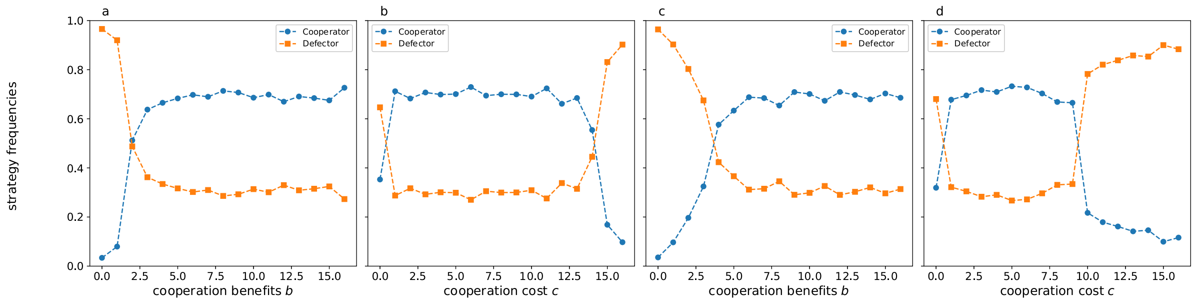

Fig. 4. The effect of benefits and costs of cooperation in the costly signaling mechanism on the frequency of each strategy. The parameter values are M=100, $w = 1$, $mu = 0.001$, $k = 0$, $f = 0$, $e = 0$. (a) $c = 1$, $a = 0.1$; (b) $b = 16$, $a = 0.1$; (c) $c = 1$, $a = 0.5$; (d) $b = 16$, $a = 0.5$.

Without punishment, the establishment of cooperation is affected by the strength of signal noise, the benefits of cooperation, and the cost of cooperation. When the noise intensity is low (a = 0.1), the cooperation is relatively easy to establish (Fig. 4a). If the noise intensity a is equal to 0.5 and the benefit b great than 3, the cooperation establish (Fig. 4c). Increasing the cost of cooperation will increase the difficulty of establishing cooperation (Figs. 4b and 4d). The above results show that: Firstly, without punishment, the costly signals mechanism can affect the ratio of cooperation and defect. Secondly, the effect of costly signals effect will be changed by the signal strength of noise. The stronger the signal, the more significant the effect of costly signals is. Thirdly, increasing the cost of cooperation to improve signal strength cannot make cooperation prevail, but instead make defect prevail. Fourthly, increasing the benefits of cooperation can increase the proportion of cooperation and reduce the proportion of defect.

Changes in low selection intensity have an effect on the ratio of cooperation and punishment. Without punishment, a gradual increase in the intensity of low selection reduces the frequency of cooperation and increases the frequency of defect. However, as the selection intensity further increases, the proportion of cooperation and defection does not change (Figure Fig. 5a). With punishment, if the cost of punishment is not converted into costly signal, the defect still accounts for the largest proportion and the cooperation proportion is the smallest. Both are affected by the change of low selection intensity (Fig. 5b). Regardless of loss due to been punished or not, converting the cost of punishment into costly signal can make the punisher dominant the population and reduce the proportion of defect. This trend is obvious when the selection strength changes at low levels, but not at high level (Figs. 5c and 5d). From the above results, it is obviously that

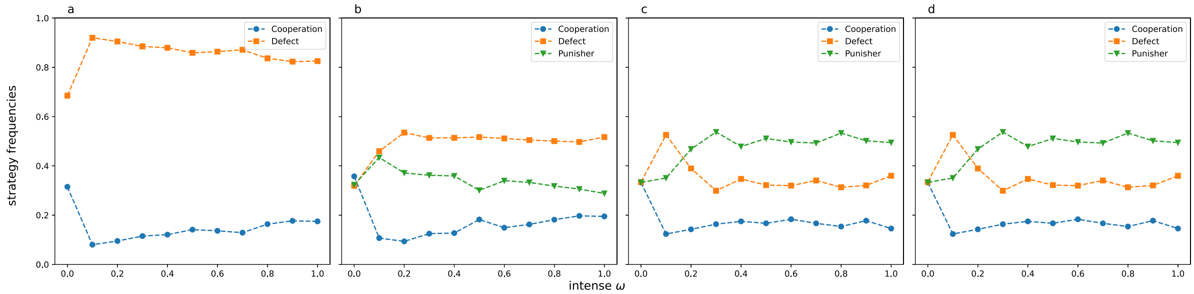

Fig. 5. Efficacy of a costly signaling system with different selection intensity on the frequency of each strategy. The parameter values are M=100, $\alpha = 0.1$, $b = 16$, $c = 15$, $\mu = 0.001$. (a) $k = 0$, $f = 0$, $e = 0$; (b) $k = 0$, $f = 5$, $e = 5$; (c) $k = 6$, $f = 5$, $e = 0$; (d) $k = 6$, $f = 5$, $e = 5$.

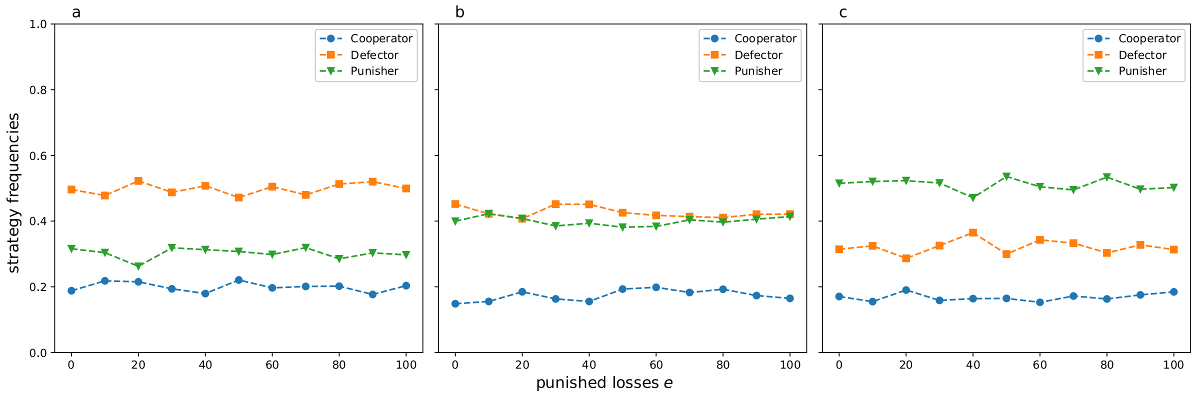

Fig. 6. The effect of the loss due to punishment on the frequency of each strategy. The parameter values are M=100, $\alpha = 0.1$, $b = 16$, $c = 15$, $w = 1$, $\mu = 0.001$. (a) $k = 0$, $f = 5$; (b) $k = 6$, $f = 0$; (c) $k = 6$, $f = 5$.

the influence of selection intensity on each strategy exists only when the intensity changes at low levels, and this effect disappears as the selection intensity increases.

Next, we further simulate the effect of the size of the loss due to punishment on the frequency of each strategy. We can see that the proportion of cooperation, punishment, and defection in the population does not change significantly due to the increasing of the loss been punished (Fig. 6). What determines the proportion of each strategy is the cost of punishment and the punishment signal transformation strength. If the cost of punishment is not transformed into a costly signal, defectors still dominate the population (Fig. 6a). When there is no punishment, the proportion of cooperation is not increased, and the proportion of defect is high (Fig. 6b, the punishment is considered as defect at this time). Only if the cost of punishment is converted into a costly signal, the punisher occupies the population (Fig. 6c).

In addition, we investigate the effect of mutation rate on the evolution of cooperation. It difficult to capture the differences between strategies with high mutation rate (Figs. 7a and 7b). Even if a certain strategy is advantageous in the game, it can be difficult to maintain because of the high mutation rate. If the mutation rate is low, it can amplify the disadvantage of certain strategies at the beginning of the game. It can also reduce the evolutionary rate (Fig. 7d). The most significant decrease in the evolutionary rate of the cooperative strategy is observed.

Finally, to further examine the role of noise and signal conversion degree in establishing cooperation, we obtain the relationship between the 100,000th cooperator frequency and noise α and cooperation cost c (Fig. 8a), as well as the punishment frequency and noise α and the punishment signal transformation strength k relationship (Figs. 8b and 8c). Without punishment, if the cost of cooperation increases, the cooperation frequency decreases. However, if increase the strength of noise, the cooperation will decrease rapidly (Fig. 8a). All the results are consistent with the above analysis. Consistently, the establishment of cooperation is positively related to the strength of the signal noise and negatively related to the cost of cooperation. With punishment, the noise α and the punishment signal hardly affect the frequency of cooperators (Fig. 8b), and the frequency of cooperation has been maintained at a low level. However, these two significantly affect the frequency of punishment and defectors (Fig. 8c). As the punishment signal transformation strength k increases, the punishment frequency increases. But as the noise increases, the punishment frequency decreases. This means that if the punishment signal transformation strength is increased or noise α decrease, the frequency of cooperation will not change, but the frequency of defect will reduce, the frequency of cooperation will increase.

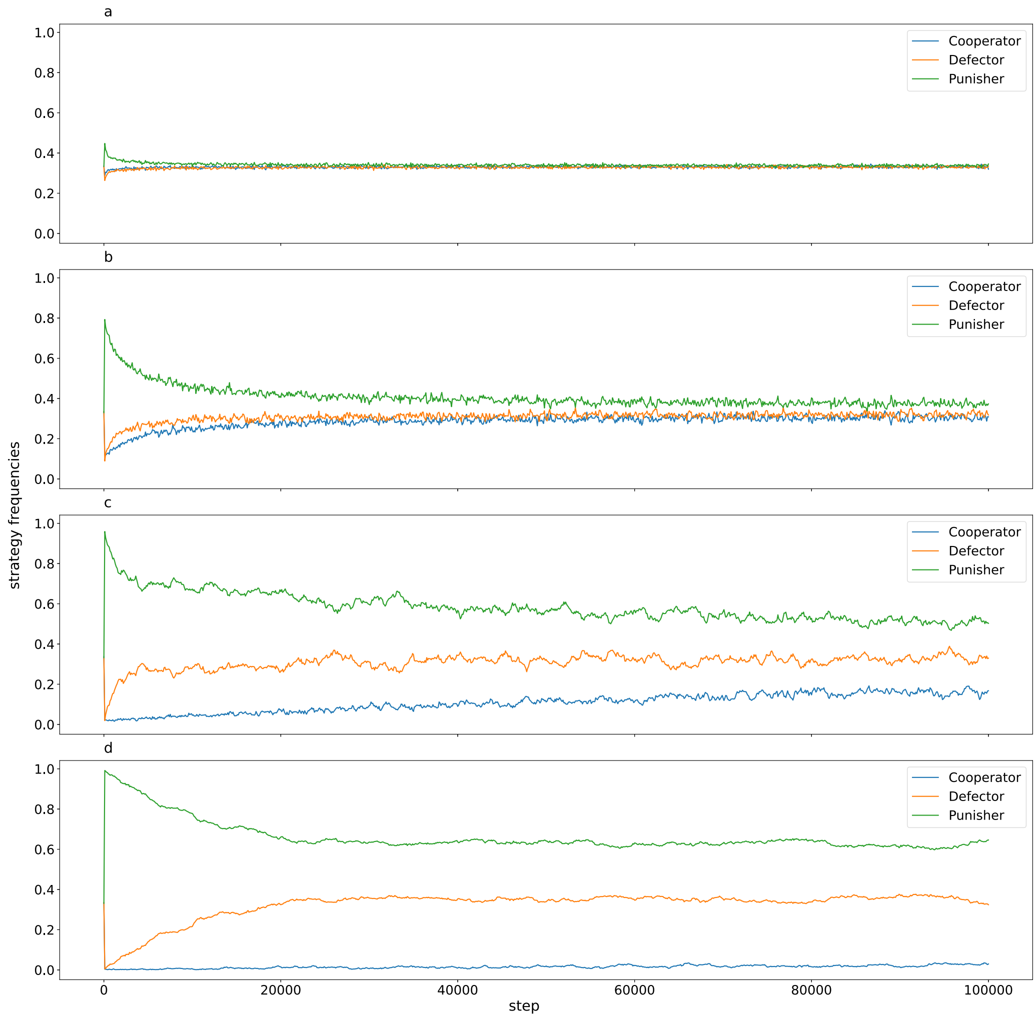

Fig. 7. Cooperative evolution in PDG. The simulation parameters are M=100, $w = 1$, $b = 16$, $c = 15$, $k = 6$, $f = 5$, $e = 5$. (a) $\mu = 0.1$; (b) $\mu = 0.01$; (c) $\mu = 0.001$; (d) $\mu = 0.0001$.

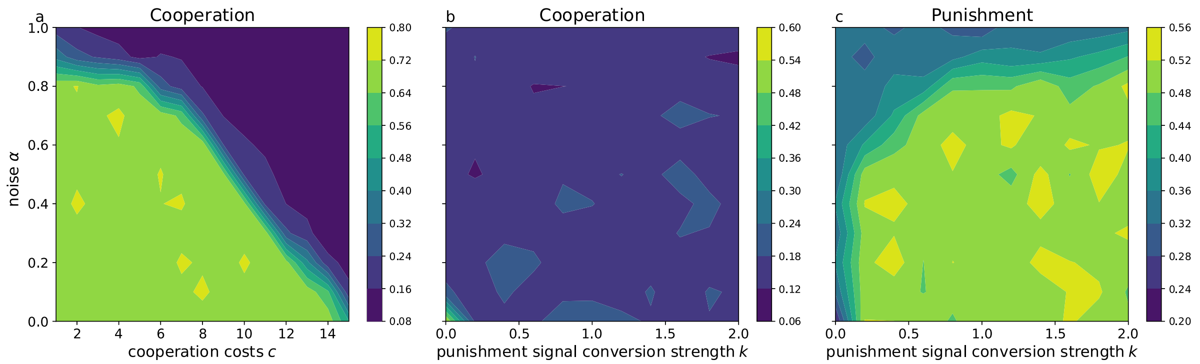

Fig. 8. Contour plots of cooperation and punishment frequencies. The simulation parameters are M=100, $w = 1$, $\mu = 0.001$, $b = 16$. (a) $k = 0$, $f = 0$, $e = 0$; (b) $c = 15$, $f = 5$, $e = 5$; (c) $c = 15$, $f = 5$, $e = 5$.

4. Discussion

The simulations in this paper are based on 4 assumptions: The first assumption is that game participants have no memory of costly signals, which means that individuals will completely ignore or forget the behavioral decisions made by the punisher in the previous round. The second assumption is that once a decision is made to adopt a punishing strategy, no matter what strategy the opponent adopts, the punishment strategy is adopted. That is, there is no more information about the strategy adopted by the opponent. The third assumption is that the costly signal is transformed into a monotonically increasing relationship between fitness and the cost. The fourth assumption is that individuals in the game are able to distinguish costly signals between punishment and cooperation. The first and second are harsh assumptions, which are not conducive to the establishment of cooperation and the emergence of punishment. If the individual's cooperative behavior and punishment behavior are regarded as special signals, even under such harsh assumptions, punishment may still appear, then cooperation and punishment will also occur under the more relaxed assumption. The third and fourth assumptions ensure that costly signaling mechanisms work.

In some studies, punishers end up dominating the entire population, and unpunished cooperator would all but die out if not for mutation [57,58]. Our simulation results show that no matter what parameter combination is used, no strategy dominates the entire population, and no strategy perishes. Different parameters only affect the relative proportion of strategies, which is in line with the real world.

Without punishment, the costly signaling mechanism can make cooperation dominant, provided that the noise of the costly signal is not too much and the cost of cooperation cannot be too high. The magnitude of noise affects the efficacy of costly signals, and previous research on costly signals in public goods has similar conclusions [59]. Interestingly, the higher the cost of cooperation, the higher the ratio of the transmitted costly signal to fitness should be. However, the simulation results show that the higher the cost of cooperation, the more dominant defection. Some studies have shown that when the cost of arranging a commitment deal lies within certain limits, substantial levels of cooperation can be achieved [60]. It means that the costly signaling mechanism has a conditional limit. If the cost of cooperation is too high, the mechanism will fail. That is to say, individuals need to weigh the appropriate cost of cooperation in order to obtain an evolutionary advantage. If the cooperation cost is too high, the return from the game will be too low to increase the fitness. If the cooperation cost is too low, the return from the game is not enough to improve fitness. This is similar to the behavior shown by most individuals in reality. It is not a complete defector, but they adopt different level of cooperation and pays different costs to obtain the maximum benefit. It suggests that the costly signal is not the better the higher the cost in terms of depositing cooperation. A study based on the Philip Sidney game shows that honest signals must be costly if there is a conflict of interest between signaler and receiver, but cost-free signals can be honest if there is no such conflict [61]. It suggests that altruism as a sign of honesty is not necessarily high cost.

With punishment, the key to increasing the ratio of punishers by the costly signaling mechanism is not how much loss other strategies suffered, but the strength of the punishment signal. The strength of the signal can come from the amplification of the punishment cost, or from the high punishment costs. A study based on the Philip Sidney game shows that the effect of honest signaling varies greatly depending on the type of individual being punished [62]. However, our study does not consider the type of punished individuals. The costly signaling mechanism can explain the long-term existence of punishment strategies in the population. Not to reduce the benefits of other strategies, but to improves their own fitness. Previous studies have emphasized that punishing defectors to reduce the benefits of defectors to improve the level of cooperation [35,63], however, the purpose of punishment under the costly signaling mechanism is not to reduce others' payoff, but to show their altruism.

If our assumption 3 holds, the signal of altruism transformed by costly signaling mechanism in the form of a multiple of scaling up or down the punishment cost. When the cost of punishment can be amplified as a signal of altruism, the punishment as a strategy is not difficult to survive in the population. On the contrary, punishment is far more advantageous than cooperation, and in some cases it is also advantageous over defect. Although the punishment payoff is higher than the cooperation, because the punishment is a more expensive signal, and has a more competitive advantage. "In a cooperative game of direct reciprocity, the winner will not choose a costly punishment strategy"[37], which is not contradictory to our simulation result, because it is an indirect reciprocity.

There are some controversies over whether the introduction of cost punishment in prisoner's dilemma games increases cooperation. An experimental study from China show that the introduction of punishment did not significantly increase or decrease cooperation, which may be caused by different attitudes towards cooperation and punishment in Chinese culture [64]. Another study shows that the introduction of a costly punishment can significantly improve the level of cooperation, but does not increase the average return of the population, and the winner in the population usually does not adopt the punishment strategy [37]. Our simulation results show that costly punishment do not affect the level of cooperation without costly signals. Even with the introduction of cost signals, it only increases the frequency of punishment and does not affect the frequency of cooperation. It seems that our results confirm the research of Wu et al. [64].

Experiments have shown that third-party punishers receive less rewards (compensations) than third-party helpers [65]. There are two things to note about this result: First, third-party punishers are not as compensated as third-party helpers, and our simulation does not involve the role of third parties. Second, some researchers believe that punishers may benefit from being treated as long-term partners rather than being rewarded in the real world, thus requiring more realistic experimental conditions [65]. For example, while punishers are less often chosen as temporary partners in trust games [66], they are more frequently used as providers of resources [67]. Considering the above experimental evidence that punishers have an adaptive advantage in that they may be selected as long-term partners rather than as recipients of rewards, we argue that fitness algorithms that directly transform payoffs into fitness may not be appropriate to explain why altruism punishment can evolve. Thus, we emphasize in our simulations from a costly signal perspective that it is not the rewards themselves that determine fitness, but the extent to which those rewards can be translated into social attractiveness that brings allies and partners.

Our simulation results are very similar to some experimental results. Experimental studies have shown that women are attracted to men who display heroism and altruism. And women prefer to long-term partners than short-term partners [68].

5. Conclusion

A long-standing problem in the biological and social sciences is understanding the conditions required for cooperation to emerge and maintain in evolving populations. Costly signaling mechanisms have been proposed as possible mechanisms explaining the emergence and maintenance of cooperation. In this study, punishment is introduced into the classic prisoner's dilemma game. Cooperation and punishment are considered as different level of costly signals. The costly signal is convert into fitness to improve the situation of cooperation and punishment. The simulation results of this study show that the costly signaling mechanism can explain the emergence and maintenance of cooperation and punishment under certain conditions. Firstly, without punishment, if there is a small noise in the costly signaling mechanism and the difference between cost and benefits of cooperation is very small, cooperation is a better choice than defection. Secondly, with punishment, if there is a small noise in the costly signaling mechanism and punishment is consider as more expensive signal than cooperation, punishment would be a better strategy.

CRediT authorship contribution statement

Kaixuan Guan: Methodology, Software, Writing – original draft. **Yuyou Chen:** Conceptualization, Writing – original draft, Writing – review & editing. **Wanjun Zheng:** Visualization. **Lulu Zeng:** Data curation. **Hang Ye:** Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank anonymous reviewers for their valuable comments.

References

[1] R. Axelrod, W.D. Hamilton, The evolution of cooperation, Science 211 (4489) (1981) 1390–1396, http://dx.doi.org/10.1126/science.7466396.
[2] M.A. Nowak, Five rules for the evolution of cooperation, Science 314 (5805) (2006) 1560–1563, http://dx.doi.org/10.1126/science.1133755.
[3] J.L. Sachs, U.G. Mueller, T.P. Wilcox, J.J. Bull, The evolution of cooperation, Q. Rev. Biol. 79 (2) (2004) 135–160, http://dx.doi.org/10.1086/383541.
[4] K.R. Foster, T. Wenseleers, F.L. Ratnieks, Kin selection is the key to altruism, Trends Ecol. Evol. 21 (2) (2006) 57–60.
[5] D.S. Wilson, A theory of group selection, Proc. Natl. Acad. Sci. U S A 72 (1) (1975) 143–146, http://dx.doi.org/10.1073/pnas.72.1.143.
[6] D.G. Rand, H. Ohtsuki, M.A. Nowak, Direct reciprocity with costly punishment: generous tit-for-tat prevails, J. Theoret. Biol. 256 (1) (2009) 45–57, http://dx.doi.org/10.1016/j.jtbi.2008.09.015.
[7] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity, Nature 437 (7063) (2005) 1291–1298, http://dx.doi.org/10.1038/nature04131.
[8] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity by image scoring, Nature 393 (6685) (1998) 573–577, http://dx.doi.org/10.1038/31225.
[9] H. Ohtsuki, C. Hauert, E. Lieberman, M.A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks, Nature 441 (7092) (2006) 502–505, http://dx.doi.org/10.1038/nature04605.
[10] G. Szabó, G. Fath, Evolutionary games on graphs, Phys. Rep. 446 (4–6) (2007) 97–216.
[11] Z. Wang, A. Szolnoki, M. Perc, Interdependent network reciprocity in evolutionary games, Sci. Rep. 3 (1) (2013) 1183, http://dx.doi.org/10.1038/srep01183.
[12] X. Chen, F. Fu, L. Wang, Interaction stochasticity supports cooperation in spatial prisoner's dilemma, Phys. Rev. E 78 (5) (2008) 051120.
[13] R.L. Riolo, M.D. Cohen, R. Axelrod, Evolution of cooperation without reciprocity, Nature 414 (6862) (2001) 441–443.
[14] C. Hauert, S. De Monte, J. Hofbauer, K. Sigmund, Volunteering as red queen mechanism for cooperation in public goods games, Science 296 (5570) (2002) 1129–1132.
[15] R. Boyd, H. Gintis, S. Bowles, Coordinated punishment of defectors sustains cooperation and can proliferate when rare, Science 328 (5978) (2010) 617–620.
[16] A. Szolnoki, M. Perc, Antisocial pool rewarding does not deter public cooperation, Proc. R. Soc. B Biol. Sci. 282 (1816) (2015) 20151975.
[17] L.A. Martinez-Vaquero, F.C. Santos, V. Trianni, Signalling boosts the evolution of cooperation in repeated group interactions, J. R. Soc. Interface 17 (172) (2020) 20200635.
[18] F.C. Santos, M.D. Santos, J.M. Pacheco, Social diversity promotes the emergence of cooperation in public goods games, Nature 454 (7201) (2008) 213–216, http://dx.doi.org/10.1038/nature06940.
[19] M. Perc, A. Szolnoki, Social diversity and promotion of cooperation in the spatial prisoner's dilemma game, Phys. Rev. E Stat. Nonlin Soft Matter Phys. 77 (1 Pt 1) (2008) 011904, http://dx.doi.org/10.1103/PhysRevE.77.011904.
[20] L.M. Pereira, T. Lenaerts, et al., Evolution of commitment and level of participation in public goods games, Auton. Agents Multi-Agent Syst. 31 (3) (2017) 561–583.
[21] X. Li, W. Han, W. Yang, J. Wang, C. Xia, H.-j. Li, Y. Shi, Impact of resource-based conditional interaction on cooperation in spatial social dilemmas, Physica A 594 (2022) 127055.

[22] A. Szolnoki, M. Perc, G. Szabo, H.U. Stark, Impact of aging on the evolution of cooperation in the spatial prisoner's dilemma game, Phys. Rev. E Stat. Nonlin Soft Matter Phys. 80 (2 Pt 1) (2009) 021901, http://dx.doi.org/10.1103/PhysRevE.80.021901.
[23] H.-X. Yang, J. Yang, Reputation-based investment strategy promotes cooperation in public goods games, Physica A 523 (2019) 886–893.
[24] Q. Jian, X. Li, J. Wang, C. Xia, Impact of reputation assortment on tag-mediated altruistic behaviors in the spatial lattice, Appl. Math. Comput. 396 (2021) 125928.
[25] H.-X. Yang, L. Sun, Heterogeneous donation game in geographical small-world networks, Physica A 540 (2020) 123255.
[26] Z. Danku, M. Perc, A. Szolnoki, Knowing the past improves cooperation in the future, Sci. Rep. 9 (1) (2019) 262, http://dx.doi.org/10.1038/s41598-018-36486-x.
[27] Z. Li, D. Jia, H. Guo, Y. Geng, C. Shen, Z. Wang, X. Li, The effect of multigame on cooperation in spatial network, Appl. Math. Comput. 351 (2019) 162–167.
[28] L. Shi, I. Romić, Y. Ma, Z. Wang, B. Podobnik, H.E. Stanley, P. Holme, M. Jusup, Freedom of choice adds value to public goods, Proc. Natl. Acad. Sci. 117 (30) (2020) 17516–17521.
[29] X. Li, G. Hao, Z. Zhang, C. Xia, Evolution of cooperation in heterogeneously stochastic interactions, Chaos Solitons Fractals 150 (2021) 111186.
[30] A. Szolnoki, M. Perc, Information sharing promotes prosocial behaviour, New J. Phys. 15 (5) (2013) 053010.
[31] C. Xia, X. Li, Z. Wang, M. Perc, Doubly effects of information sharing on interdependent network reciprocity, New J. Phys. 20 (7) (2018) 075005.
[32] T.H. Clutton-Brock, G.A. Parker, Punishment in animal societies, Nature 373 (6511) (1995) 209–216, http://dx.doi.org/10.1038/373209a0.
[33] E. Fehr, S. Gächter, Cooperation and punishment in public goods experiments, Amer. Econ. Rev. 90 (4) (2000) 980–994.
[34] T. Yamagishi, The provision of a sanctioning system as a public good, J. Personal. Soc. Psychol. 51 (1) (1986) 110.
[35] R. Boyd, P.J. Richerson, Punishment allows the evolution of cooperation (or anything else) in sizable groups, Ethol. Sociobiol. 13 (3) (1992) 171–195.
[36] J.H. Fowler, Altruistic punishment and the origin of cooperation, Proc. Natl. Acad. Sci. U S A 102 (19) (2005) 7047–7049, http://dx.doi.org/10.1073/pnas.0500938102.
[37] A. Dreber, D.G. Rand, D. Fudenberg, M.A. Nowak, Winners don't punish, Nature 452 (7185) (2008) 348–351, http://dx.doi.org/10.1038/nature06723.
[38] S. Bowles, H. Gintis, The Evolution of Cooperation in Heterogeneous Populations, WP Santa Fe Institute, 2003, pp. 03–05.
[39] R. Boyd, H. Gintis, S. Bowles, P.J. Richerson, The evolution of altruistic punishment, Proc. Natl. Acad. Sci. U S A 100 (6) (2003) 3531–3535, http://dx.doi.org/10.1073/pnas.0630443100.
[40] S. Gachter, B. Herrmann, Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment, Philos. Trans. R. Soc. Lond. B Biol. Sci. 364 (1518) (2009) 791–806, http://dx.doi.org/10.1098/rstb.2008.0275.
[41] A. Chaudhuri, Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature, Exp. Econ. 14 (1) (2011) 47–83.
[42] D. Balliet, L.B. Mulder, P.A. Van Lange, Reward, punishment, and cooperation: a meta-analysis, Psychol. Bull. 137 (4) (2011) 594–615, http://dx.doi.org/10.1037/a0023489.
[43] M. Milinski, B. Rockenbach, On the interaction of the stick and the carrot in social dilemmas, J. Theoret. Biol. 299 (2012) 139–143, http://dx.doi.org/10.1016/j.jtbi.2011.03.014.
[44] A. Zahavi, Mate selection-a selection for a handicap, J. Theoret. Biol. 53 (1) (1975) 205–214, http://dx.doi.org/10.1016/0022-5193(75)90111-3.
[45] A. Grafen, Biological signals as handicaps, J. Theoret. Biol. 144 (4) (1990) 517–546.
[46] A. Zahavi, A. Zahavi, The Handicap Principle: A Missing Piece of Darwin's Puzzle, Oxford University Press, 1999.
[47] F.T. McAndrew, Costly signaling theory, in: Encyclopedia of Evolutionary Psychological Science, 2021, pp. 1525–1532.
[48] R. BliegeBird, E. Smith, Signaling theory, strategic interaction, and symbolic capital, Curr. Anthropol. 46 (2) (2005) 221–248.
[49] E.A. Smith, R.L.B. Bird, Turtle hunting and tombstone opening: Public generosity as costly signaling, Evol. Hum. Behav. 21 (4) (2000) 245–261.
[50] E.A. Smith, R.B. Bird, D.W. Bird, The benefits of costly signaling: Meriam turtle hunters, Behav. Ecol. 14 (1) (2003) 116–126.
[51] I. Back, A. Flache, The viability of cooperation based on interpersonal commitment, J. Artif. Soc. Soc. Simul. 9 (1) (2006).
[52] A. Traulsen, N. Shoresh, M.A. Nowak, Analytical results for individual and group selection of any intensity, Bull. Math. Biol. 70 (5) (2008) 1410–1424.
[53] P. Gilbert, J. Price, S. Allan, Social comparison, social attractiveness and evolution: How might they be related? New Ideas Psychol. 13 (2) (1995) 149–165.
[54] P. Gilbert, The evolution of social attractiveness and its role in shame, humiliation, guilt and therapy, Br. J. Med. Psychol. 70 (2) (1997) 113–147, http://dx.doi.org/10.1111/j.2044-8341.1997.tb01893.x, Pt 2.
[55] C.M. Larson, M.G. Haselton, K.A. Gildersleeve, E.G. Pillsworth, Changes in women's feelings about their romantic relationships across the ovulatory cycle, Horm. Behav. 63 (1) (2013) 128–135.
[56] K. Gildersleeve, M.G. Haselton, M.R. Fales, Do women's mate preferences change across the ovulatory cycle? A meta-analytic review, Psychol. Bull. 140 (5) (2014) 1205.
[57] C. Hauert, A. Traulsen, H. Brandt, M.A. Nowak, K. Sigmund, Via freedom to coercion: the emergence of costly punishment, Science 316 (5833) (2007) 1905–1907, http://dx.doi.org/10.1126/science.1141588.
[58] H. Ye, S. Chen, J. Luo, F. Tan, Y. Jia, Y. Chen, Increasing returns to scale: The solution to the second-order social dilemma, Sci. Rep. 6 (1) (2016) 31927, http://dx.doi.org/10.1038/srep31927.
[59] H. Niu, Y. Chen, H. Ye, H. Zhang, Y. Li, S. Chen, Distinguishing punishing costly signals from nonpunishing costly signals can facilitate the emergence of altruistic punishment, Appl. Math. Comput. 371 (2020) 124945.
[60] T.A. Han, L.M. Pereira, F.C. Santos, T. Lenaerts, Good agreements make good friends, Sci. Rep. 3 (1) (2013) 1–7.
[61] J.M. Smith, Honest signalling: the philip sidney game, Anim. Behav. (1991).
[62] D. Catteeuw, T.A. Han, B. Manderick, Evolution of honest signaling by social punishment, in: Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, 2014, pp. 153–160.
[63] T.A. Han, L.M. Pereira, T. Lenaerts, Avoiding or restricting defectors in public goods games? J. R. Soc. Interface 12 (103) (2015) 20141203.
[64] J.-J. Wu, B.-Y. Zhang, Z.-X. Zhou, Q.-Q. He, X.-D. Zheng, R. Cressman, Y. Tao, Costly punishment does not always increase cooperation, Proc. Natl. Acad. Sci. 106 (41) (2009) 17448–17451.
[65] N.J. Raihani, R. Bshary, Third-party punishers are rewarded, but third-party helpers even more so, Evolution 69 (4) (2015) 993–1003.
[66] J.J. Jordan, M. Hoffman, P. Bloom, D.G. Rand, Third-party punishment as a costly signal of trustworthiness, Nature 530 (7591) (2016) 473–476.
[67] Y. Horita, Punishers may be chosen as providers but not as recipients, Lett. Evol. Behav. Sci. 1 (1) (2010) 6–9.
[68] L. Margana, M.S. Bhogal, J.E. Bartlett, D. Farrelly, The roles of altruism, heroism, and physical attractiveness in female mate choice, Pers. Individ. Differ. 137 (2019) 126–130.