

## Article

# Evolution of “Pay-It-Forward” in the Presence of the Temptation to Free-Ride

Satoshi Uchida <sup>1,2,\*</sup>, Tatsuya Sasaki <sup>3</sup> , Hitoshi Yamamoto <sup>4</sup>  and Isamu Okada <sup>5</sup> <sup>1</sup> Research Center for Ethic-Culture Studies, RINRI Institute, Tokyo 102-8561, Japan<sup>2</sup> High-Tech Research Center, Kokushikan University, Tokyo 154-8515, Japan<sup>3</sup> Department of Community Development, Koriyama Women’s College, Fukushima 963-8503, Japan; t.sasaki@koriyama-kgc.ac.jp<sup>4</sup> Faculty of Business Administration, Rissho University, Tokyo 141-8602, Japan; hitoshi@ris.ac.jp<sup>5</sup> Department of Business Administration, Soka University, Tokyo 192-8577, Japan; okada@soka.ac.jp

\* Correspondence: s-uchida@rinri-jpn.or.jp

**Abstract:** “Paying it forward” is a behavior in which people help someone else because they were helped in the past. Although experimental evidence exists that indicates that real human beings often “pay-it-forward” even in the face of free-rider risks, the theoretical basis for the evolution of this behavior remains unclear. In this paper, we propose a game-theoretical model that explains how pay-it-forward behavior can evolve despite the temptation to free-ride. By assuming that human beings exhibit cognitive distortions, as predicted by prospect theory, and that free-riding is punished with a tiny probability, we demonstrate that pay-it-forward, alongside unconditional altruistic behavior, can evolve and effectively deter free-riding behavior.

**Keywords:** pay-it-forward; cooperation; social dilemma; punishment; evolutionary game theory; prospect theory; expected utility theory



**Citation:** Uchida, S.; Sasaki, T.; Yamamoto, H.; Okada, I. Evolution of “Pay-It-Forward” in the Presence of the Temptation to Free-Ride. *Games* **2024**, *15*, 16. <https://doi.org/10.3390/g15030016>

Academic Editors: Ulrich Berger and Alberto Antonioni

Received: 17 February 2024

Revised: 16 April 2024

Accepted: 19 April 2024

Published: 25 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

People often “pay-it-forward” by passing on benefits they have received from someone in the past to others in the future. Experimental and observational studies have provided evidence that people indeed engage in decisions reflective of pay-it-forward behavior [1,2]. For instance, in [1], game experiments conducted in a laboratory setting revealed that the altruistic behavior of one game player spread to other players within the social network. This propagation of altruistic behaviors from one player to another encourages cooperation among players.

The presence of pay-it-forward behaviors in the real world presents a significant challenge to game theory, especially within the context of evolutionary game theory. Essentially, when a player chooses to pass on benefits received from someone in the past to others, they do not gain any immediate personal profits. Consequently, individuals who opt not to pay-it-forward (i.e., those who engage in “free-riding”) often find themselves at a greater advantage compared to those who participate in pay-it-forward activities (“pay-it-forwarders”). Without introducing an additional mechanism, natural selection or social learning processes are expected to favor non-participants [3]. Thus, identifying a mechanism that discourages free-riding behavior has become a crucial theoretical challenge.

Unconditional altruistic behavior, in which individuals help others without expecting anything in return, often leads to what is known as the social dilemma, or free-rider problem, as documented in the literature [4–10]. Pay-it-forward practices, however, represent a form of conditional altruistic behavior where individuals help others only if they themselves have received help in the past. This fact highlights that conditional altruistic behavior shares the same underlying structure as the social dilemma, characterized by the following: (1) individuals in a society have the choice to either engage in pay-it-forward

practices or to free-ride, (2) a society comprised mainly of pay-it-forwarders is generally more beneficial than one filled with free-riders, and (3), on an individual level, free-riders tend to fare better than pay-it-forwarders. This is because pay-it-forwarders distribute the benefits they have received to others, whereas free-riders keep any advantages for themselves. In situations presenting a social dilemma, opting for free-riding becomes more appealing for individuals irrespective of the choices made by others.

In this paper, we show that pay-it-forward can be sustainably maintained if individuals adhere to what is known as prospect theory (PT) and if punishment for not helping others is applied, albeit infrequently. This stands in contrast to scenarios where individuals are so rational that they follow the (linear) expected utility theory (EUT), under which pay-it-forward practices do not persist. In our previous research, we established that prospect theory facilitates cooperation through peer punishment [11], whereas EUT does not lead to the same outcome when free-riders are only occasionally punished. We hypothesize that a similar rationale applies to the evolution of cooperation through pay-it-forward mechanisms. Investigating this hypothesis is the main focus of the current paper.

Prospect theory is widely recognized as one of the most successful frameworks for describing the cognitive biases inherent in human decision-making [12–14]. Within this theory, distortions in the perception of probabilities and the evaluation of outcomes are represented through a weighting probability function and a value function, respectively. Both functions are characterized by their non-linearity and asymmetry, capturing the essence of human irrationality. This theoretical approach highlights how individuals often deviate from rational decision-making models, particularly in the context of risk and uncertainty, by assigning disproportionate weights to certain outcomes and probabilities.

Most studies on the evolution of cooperation in evolutionary game theory have traditionally been conducted under the assumption that individuals adhere to EUT [5]. EUT posits that individuals are perfectly rational, making their calculations regarding potential benefits, or so-called expected payoffs, with precision. Contrary to this assumption, experimental research on both behavioral economics and experimental economics has consistently demonstrated that real humans do not conform to EUT [15,16]. In reality, humans are incapable of estimating probabilities and values without bias. The capacity to accurately calculate probabilities and values, devoid of any distortions, underpins the concept of linearly calculated expected payoffs [17–22]. However, the payoffs, as perceived by real humans, are subjective, and this subjectivity is a critical factor in the actual decision-making process. This discrepancy between the theoretical assumptions of EUT and the practical observations under PT highlights the complexities and nuances in understanding human behavior in evolutionary game-theoretic contexts.

In the subsequent sections of this paper, we demonstrate that, when individuals exhibit the level of irrationality encapsulated by PT, pay-it-forward behaviors emerge, and cooperation persists even within parameter regions where conditional altruistic actions cannot be sustained under the rationality assumptions of EUT. Specifically, the risk associated with free-riding behaviors is mitigated by the cognitive biases and decision-making patterns described by PT. Furthermore, the strategy of pay-it-forward behaviors has the potential to overturn unconditional altruistic behavior, highlighting a critical distinction between paying-it-forward and unconditional cooperation, despite their superficial similarities.

Following a brief review of the relevant literature in the next section, we will outline our model before presenting our findings. The paper concludes with a section dedicated to our conclusions, summarizing the implications of our research on the understanding of cooperative behaviors and the influence of cognitive biases on the interactions among individuals.

## 2. Brief Literature Review

There are several recent papers that have combined evolutionary game theory with prospect theory, aiming to derive practical implications for addressing social issues. For instance, the study referenced as [23] considers the irrationality of players, as described by prospect theory, and constructs an evolutionary game-theoretical model to explore the

decision-making process of enterprises when choosing between innovation and imitation. A paper by Sun, Gao, and Li published in 2022 [24] analyzes interactions between enterprises and governments, particularly in the context of establishing carbon emission regulations. Meanwhile, the paper referenced as [25] focuses on the decision-making processes of polluting enterprises and their interactions with government bodies. A more recent study [26] investigates the recycling of power batteries within the electric vehicle industry, proposing a game model that involves three distinct types of players (governments, manufacturers, and recycling companies). Each of these papers develops an evolutionary game model that incorporates aspects of prospect theory.

The evolution of cooperation has been extensively studied within the field of evolutionary game theory, as highlighted in [27]. A recent theoretical study [28] attempts to elucidate the evolution of cooperative (and coordinated) behavior by incorporating the effects of prospect theory into a model. This is achieved through the lens of a coordination game known as the stag-hunt game.

In this paper, we aim to show a path to the evolution of cooperation through pay-it-forward behaviors using a prisoner's-dilemma-type game called the "donation game". The mechanism to sustain cooperation via pay-it-forward is a type of so-called indirect reciprocity mechanism. Indirect reciprocity is one of the five typical mechanisms to sustain cooperation [29]. In indirect reciprocity, beneficial acts are returned not by the recipient, as in the direct reciprocity mechanism, but by third parties. In the literature [6], indirect reciprocity is classified into the following two categories: downstream reciprocity and upstream reciprocity. Pay-it-forward is categorized as upstream reciprocity, which represents the "I-help-you-because-someone-helped-me" attitude. Downstream reciprocity represents the "I-help-you-because-you-helped-someone" attitude, which is the mainstream in the study of indirect reciprocity.

While the literature on the evolution of cooperation via downstream reciprocity is ample [30–38], evolutionary theory has yet to explain the effectiveness of upstream reciprocity in regard to the evolution of cooperation. A well-cited paper on upstream reciprocity in the context of evolutionary game theory was published in 2007 [3], but it concludes that upstream reciprocity can evolve as a "by-product of direct reciprocity", i.e., upstream reciprocity works if it is combined with another effective mechanism, direct reciprocity. In [39], the evolution of upstream reciprocity was explained in a complex network, and in [40], it was shown that cooperation can evolve by upstream reciprocity only in small interaction groups. According to a thorough review on the theoretical research of indirect reciprocity published recently [41], although there is some experimentation on upstream reciprocity, explaining it via theoretical viewpoints is still challenging.

### 3. Model

#### 3.1. Game

To maintain simplicity and facilitate analysis of the model, in line with several studies in the field of the evolution of cooperation [11,42–44], we consider a large, well-mixed population of individuals (or players). Two players are chosen at random from the population from time to time. Let us call the chosen players A and B. Players A and B engage in the following "donation game" (or "giving game") [6]: each player decides whether to help the opponent by paying a cost  $c$ . If Player A (B) chooses to help B (A), the opponent receives a benefit  $b > c$ ; otherwise A (or B) does not pay anything and B (A) obtains nothing. Every player in the population is selected as a game player (infinitely) many times and experiences decision-making processes. From here on, we denote the action "cooperate with its opponent (or just cooperate)" by "C" and "defect its opponent (or just defect)" by "D".

Table 1 shows the payoffs Player A obtains when playing the donation game with Player B. It consists of four possible outcomes, as both players have two options each. If both players choose C, Player A gains a benefit of  $b - c$ . If Player A chooses C while Player B opts for D, Player A incurs a loss of  $c$ . Conversely, if Player A selects D and Player B

chooses C, then Player A receives a benefit of  $b$ . If both players select D, Player A receives no benefit (and neither does Player B).

**Table 1.** Payoffs player A obtains in donation game.

Player B's Option Player A's Option	Cooperate (C)	Defect (D)
Cooperate (C)	$b - c$	$-c$
Defect (D)	$b$	$0$

### 3.2. Game, Strategies and Payoff Matrix

We consider the following three strategies: AC, AD, and S [42]. Strategy AC prescribes to help others with probability  $1 - \epsilon$ . Due to the shortage of resources, AC-players do not give help with probability  $\epsilon$ , even if they would like to. We call  $\epsilon$  an implementation error and set its value to 0.05. Players using strategy AD are free-riders and never help others. We call those players AD-players. Strategy S players correspond to pay-it-forward, which prescribes to choosing C (with probability  $1 - \epsilon$ ) if they were helped in the last game. If they were not helped in the last game, S-players choose D in the current game. We also refer to the strategies by numbers: AC = strategy 1, AD = strategy 2, and S = strategy 3.

### 3.3. Payoff Matrix

We examine a type of punishment known as pool punishment [45–51]. In this framework, a punishment institution exists in the population and executes punishment to those who chose D in the donation game. Players in the population pay a cost  $d$ , which will be used to maintain the punishment institution.

In this paper, we assume that the institution executes punishment with only a small probability to account for situations where the institution functions insufficiently. From the viewpoints of players, after the two selected players have played the donation game, they have a small chance of being punished by the external institution if they chose D in the donation game. The probability of being punished is  $\delta$ , which is set to be 0.2 in the following numerical simulations.

We assume that AD-players, who never give help to others, do not pay  $d$  (See Section 3.5 for more details). In this sense, AD-players are not only first-order free-riders but also second-order free-riders. Here, first-order free-riders mean that they do not engage in cooperative activities, while second-order free-riders are defined as those who do not contribute to punishment activities. AD-players are assumed to be both first-order and second-order free-riders at the same time. This is a harsher assumption for the evolution of cooperation than the assumption that AD-players also pay the cost to maintain the punishment institution, as other types of players are at more of a disadvantage if AD-players do not incur the punishment cost.

The strength of punishment is denoted by  $s$ . Thus, if a player chooses D in the donation game, and if punishment is executed by the institution (which occurs with probability  $\delta$ ), the player suffers a fine,  $s$ . If punishment is not executed, nothing is imposed on the player. As a result, we obtain Table 2 as the payoff matrix studied in this paper. This matrix comprises eight elements, as Player A has two options (C or D), and, for each option, there are four possibilities depending on whether the opponent, Player B, chooses C or D and whether the institution executes punishment or not.

**Table 2.** Payoff matrix  $M$  (payoffs Player A obtains in total).

Outcome Player A's Option	Cooperate Not-Punish (CN)	Cooperate Punish (CP)	Defect Not-Punish (DN)	Defect Punish (DP)
Cooperate (C)	$b - c$	$0$	$-c$	$0$
Defect (D)	$b$	$b - s$	$0$	$-s$

We call this payoff matrix  $M$ . In Table 2, for example, “CP” means that Player B helps Player A and the institution executes punishment. In this case, if A does not help B, A obtains benefit  $b$  but suffers  $s$  at the same time. The total payoff in this case is  $b - s$  ( $= M_{22}$ ).

It is worth noting that if Player A helps Player B, A will never be punished. Consequently, the element of the payoff matrix corresponding to this scenario ( $M_{12}$ ) is undefined. For the same reason,  $M_{14}$  is also undefined, as Player A’s action of choosing C will never result in punishment. We formally put value “0” for the elements  $M_{12}$  and  $M_{14}$  in the table. However, these values are never used, as these situations never occur. The cost of implementing punishment is not included in the payoff matrix, as it does not result from interactions with other players. The cost for punishment is introduced after estimating the expected payoffs on the basis of interactions with other players (See Section 3.5).

### 3.4. Strategy Switching

Players occasionally switch their strategies by social learning. As a result of the switching process, the time evolution of the proportion of strategy  $i$  ( $i \in \{1, 2, 3\}$ ), denoted by  $x_i$ , is governed by the so-called replicator equation

$$\dot{x}_i = x_i(P_i - \bar{P}),$$

where  $P_i$  is the expected payoff of strategy  $i$  and  $\bar{P} = \sum_{i=1}^3 x_i P_i$  is the expected payoff of the whole population. Since  $x_i$  is the proportion of strategy  $i$ ,  $\sum_{i=1}^3 x_i = 1$  holds.

By numerically integrating the replicator equation mentioned above, we can analyze the stability of each strategy. In the subsequent subsection, we will describe how the expected payoff  $P_i$  is calculated.

### 3.5. Probability Distribution on $M$ and Expected Payoff $P_i$

#### 3.5.1. Linear Expected Utility Theory

Expected payoff  $P_i$  reflects the long-term payoff of players using strategy  $i$ . It can be determined by calculating the probability that each outcome for payoff matrix  $M_{mn}$  occurs. The probabilities are dependent on strategies. By  $B_i$ , we denote the probability distribution on  $M$  of players using strategy  $i$ .

To determine probability distributions, it is crucial to first know the probability of whether players following a certain strategy will offer help to others. For those adhering to the AC strategy, this probability is  $1 - \varepsilon$ , while for AD, it is 0. To ascertain the probability that S-players will offer help, we define  $p$  as the frequency at which players receive help from a random player within the population. This frequency,  $p$ , can also be interpreted as the probability of a player being given help by a random participant in the game.

It is important to note that the act of receiving help in the game does not rely on the strategy used by the recipient. Therefore,  $p$  is strategy-independent and represents the chance that any given player is helped in a game. Consequently, the probability that S-players will choose action C is represented by  $(1 - \varepsilon)p$ .

The frequency  $p$  evolves over time as individuals engage in the game. We assume that the dynamics of  $p$  occur at a much faster pace than changes in strategy. In essence, players experience various interactions an infinite number of times before altering their strategies. This approach keeps the variables  $x_i$  constant while allowing  $p$  to vary, a premise similarly assumed in Sections 4.2 and 4.3 of [6].

As demonstrated in Appendix A,  $p$  reaches an equilibrium defined by the following equation:

$$p = (1 - \varepsilon)x_1 + (1 - \varepsilon)px_3.$$

Here, the first term on the right-hand side represents the probability of encountering an AC-player ( $x_1$ ), who then opts for cooperation (C) with probability  $1 - \varepsilon$ . The second term accounts for the likelihood of meeting an S-player ( $x_3$ ) who, having previously been given help, chooses C with a probability of  $(1 - \varepsilon)p$ .

By solving this equilibrium equation, we find  $p$  to be:

$$p = \frac{(1 - \varepsilon)x_1}{1 - (1 - \varepsilon)x_3}.$$

Since probability distributions on  $M$  depend on strategies, the expected payoff is dependent on strategies. Table 3(a)–(c) show probability distributions  $B_i$  with parameters  $(\varepsilon, \delta, p)$ . In the tables,  $\bar{\varepsilon} = 1 - \varepsilon$ ,  $\bar{\delta} = 1 - \delta$ ,  $\bar{p} = 1 - p$ .

**Table 3.** Probability distributions for strategy 1 (a), strategy 2 (b), and strategy 3 (c).

(a) $B_1$				
Outcome Player A's option	Cooperate Not-Punish (CN)	Cooperate Punish (CP)	Defect Not-Punish (DN)	Defect Punish (DP)
Cooperate (C)	$\bar{\varepsilon}p$	0	$\bar{\varepsilon}\bar{p}$	0
Defect (D)	$\varepsilon\bar{\delta}p$	$\varepsilon\delta p$	$\varepsilon\bar{\delta}\bar{p}$	$\varepsilon\delta\bar{p}$
(b) $B_2$				
Outcome Player A's option	Cooperate Not-Punish (CN)	Cooperate Punish (CP)	Defect Not-Punish (DN)	Defect Punish (DP)
Cooperate (C)	0	0	0	0
Defect (D)	$\bar{\delta}p$	$\delta p$	$\bar{\delta}\bar{p}$	$\delta\bar{p}$
(c) $B_3$				
Outcome Player A's option	Cooperate Not-Punish (CN)	Cooperate Punish (CP)	Defect Not-Punish (DN)	Defect Punish (DP)
Cooperate (C)	$\bar{\varepsilon}p^2$	0	$\bar{\varepsilon}\bar{p}p$	0
Defect (D)	$(1 - \bar{\varepsilon}p)\bar{\delta}p$	$(1 - \bar{\varepsilon}p)\delta p$	$(1 - \bar{\varepsilon}p)\bar{\delta}\bar{p}$	$(1 - \bar{\varepsilon}p)\delta\bar{p}$

For example,  $(B_1)_{22}$  is the product of the following three probabilities: the probability that an AC-player is given help by the opponent player of a game ( $p$ ), the probability that the AC-player chooses D ( $\varepsilon$ ), and the probability that this behavior (i.e., choosing D) is punished by the external institution ( $\delta$ ). Note that  $(B_i)_{12} = (B_i)_{14} = 0$  for any  $i$ .

With these probability distributions on  $M$ , the expected payoff of players using strategy  $i$  through interactions with other players is given by

$$Q_i = \sum_{m=1}^2 \sum_{n=1}^4 (M)_{mn}(B_i)_{mn}.$$

For example,  $Q_1 = (b - c)\bar{\varepsilon}p - c\bar{\varepsilon}\bar{p} + b\varepsilon\bar{\delta}p + (b - s)\varepsilon\delta p - s\varepsilon\delta\bar{p}$ .

Considering that AD-players do not pay any cost to maintain the external punishment institution, we find the final form of the expected payoff of players following strategy  $i$ :

$$P_i = Q_i - d,$$

for  $i = 1, 3$  and

$$P_i = Q_i,$$

for  $i = 2$ .

Here,  $d$  is the cost to maintain the punishment institution. AD-players are at more of an advantage than AC-players and S-players by the value of  $d$ .

### 3.5.2. Prospect Theory

The way we include prospect theory in our evolutionary game model is essentially the same as in our previous paper [11]. Thus, we apply a value function and weighting function to each element of the payoff matrix  $M$  and to the probability distribution  $B_i$ , respectively:

$$\begin{aligned} (V)_{mn} &= v((M)_{mn}), \\ (W_i)_{mn} &= w((B_i)_{mn}), \end{aligned}$$

where  $v$  and  $w$  are the value function and the weighting function, respectively. In Appendix B, we describe the appearance of the functions.

The subjectively distorted payoff matrix  $V$  is independent of strategies. It is well-known that the nonlinearity of the value function means that the baseline of the payoff matrix, or the reference point, significantly influences the results, differing from EUT. Specifically, within the EUT framework, adding the same constant to all elements of the payoff matrix  $M$  does not alter the payoff structure. In other words, the relative superiority or inferiority of strategies remains unchanged, even if each player consistently receives the same payoff. However, under PT, adding the same constant (for instance, by assuming each player obtains a constant payoff in playing a game) can affect the outcomes. We assume the payoff obtained from the outcome (2,3) in  $M$ , which is zero, serves as the so-called reference point, with no additional constants added.

With these subjective probabilities and payoffs, the expected payoff is calculated as

$$P_i = \sum_{m=1}^2 \sum_{n=1}^4 (V)_{mn} (W_i)_{mn} + v(-d),$$

for  $i = 1, 3$  and

$$P_i = \sum_{m=1}^2 \sum_{n=1}^4 (V)_{mn} (W_i)_{mn},$$

for  $i = 2$ .

Note that we apply only the value function, not the weighting function, to the last term  $(-d)$ : the loss of payoff by paying a cost for punishment) for the following reason: it does not represent a loss suffered through interaction in a game and is therefore not a stochastic event (hence, the weighting function plays no role). By applying the value function to  $d$ , the perceived impact of the cost for maintaining the punishment institution is regarded as more significant than its actual impact. We also note that the strategy change is described by the ordinary replicator dynamics mentioned above.

## 4. Results

### 4.1. Analysis in Case of EUT

In the context of expected utility theory (EUT), the replicator dynamics can be thoroughly analyzed due to the simplicity of the expected payoffs. The expected payoffs for AC-players, AD-players, and S-players are, respectively represented as

$$P_1 = pb - (1 - \epsilon)c - \epsilon\delta s - d,$$

$$P_2 = pb - \delta s,$$

$$P_3 = pb - (1 - \epsilon)pc - [1 - (1 - \epsilon)p]\delta s - d.$$

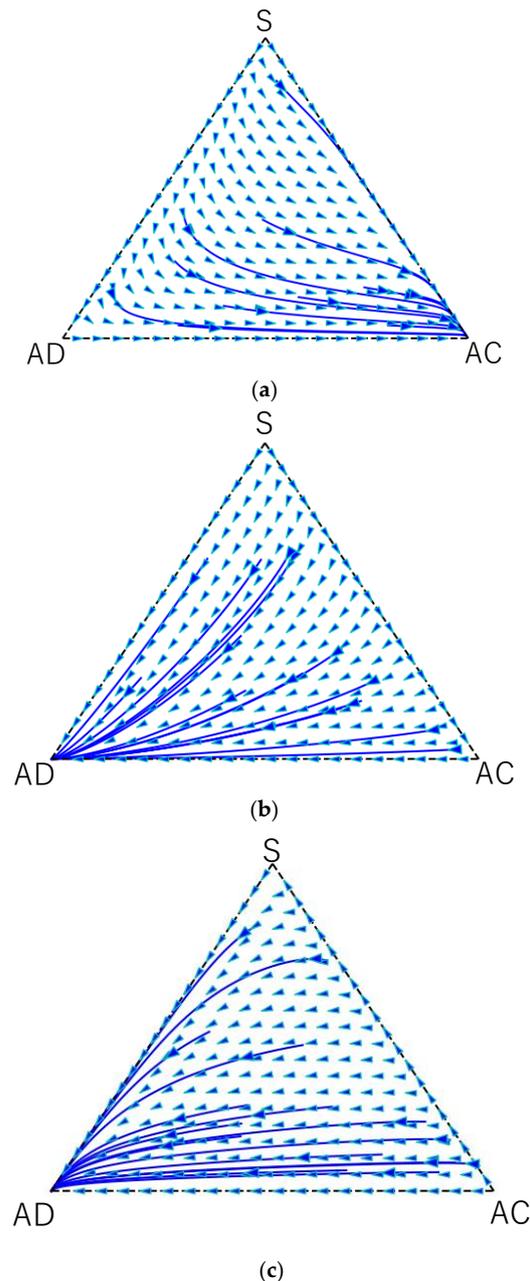
From these equations, it can be demonstrated that AD-players become extinct through the replicator dynamics, while AC-players evolve and are sustained, provided the punishment is effective or substantial enough (that is, if the parameters  $s$  or  $\delta$  surpass certain threshold values).

Specifically,  $P_1 > P_2$  suggests that  $\delta s > c + d / (1 - \epsilon)$ , while  $P_1 > P_3$  implies  $\delta s > c$ . If these conditions are met, the AC strategy emerges as the unique evolutionarily stable

strategy (ESS), leading all players to eventually adopt it, resulting in  $x_1 = 1$ . Conversely, when  $P_1 < P_2$ , the AD strategy becomes the unique ESS.

The S strategy can never be an ESS because a population consisting solely of S-players is vulnerable to invasion by an AD-player. Indeed, at the point where  $x_3 = 1$ , which implies  $p = 0$ , we find that  $P_3 = -\delta s - d < -\delta s = P_2$ . This indicates the inability of the S strategy to sustain itself as a stable strategy in the population if players follow EUT.

In the following numerical investigations, we fix parameter values as  $b = 5.0$ ,  $c = 1.0$ ,  $\varepsilon = 0.05$ ,  $d = 0.05$ . With these values, the conditions  $P_1 > P_2$  and  $P_1 > P_3$  hold if  $\delta s > 20/19 \approx 1.05$ . Figure 1a–c display trajectories in the state space obtained by numerically integrating the replicator dynamics with  $\delta = 0.2$  and with  $s = 5.6$ ,  $5.1$ ,  $4.6$ , respectively.



**Figure 1.** Trajectories and vector fields yielded by replicator dynamics for case of linear expected utility theory, with  $s = 5.6$  (a),  $s = 5.1$  (b) and  $s = 4.6$  (c). Initial conditions are random. State space is simplex defined by  $\{(x_1, x_2, x_3) | 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq 1, x_1 + x_2 + x_3 = 1\}$ , which is represented as an equilateral triangle. Arrows in trajectories show in which direction state  $(x_1, x_2, x_3)$  evolves. Other parameters:  $b = 5.0$ ,  $c = 1.0$ ,  $\varepsilon = 0.05$ ,  $d = 0.05$ ,  $\delta = 0.2$ .

The replicator dynamics analyzed to produce the figure are given by the following coupled equation system:

$$\begin{aligned}\frac{dx_1}{dt} &= x_1 \left( P_1 - \sum_{i=1}^3 x_i P_i \right), \\ \frac{dx_2}{dt} &= x_2 \left( P_2 - \sum_{i=1}^3 x_i P_i \right), \\ \frac{dx_3}{dt} &= x_3 \left( P_3 - \sum_{i=1}^3 x_i P_i \right).\end{aligned}$$

The initial conditions are set randomly. The state space is represented as the simplex  $\{(x_1, x_2, x_3) | 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq 1, x_1 + x_2 + x_3 = 1\}$ , visualized as an equilateral triangle. The vertex AC corresponds to  $(x_1, x_2, x_3) = (1, 0, 0)$ , AD to  $(0, 1, 0)$  and S to  $(0, 0, 1)$ . The state of the population  $(x_1, x_2, x_3)$  evolves within this space according to the replicator equation. The figures also illustrate vector fields generated by the replicator dynamics.

If  $s$  is sufficiently large, for example, 5.6, all trajectories converge to the vertex AC, as predicted by the above analysis (see Figure 1a). Consequently, all players ultimately adopt strategy AC, consistently aiming to help others.

With lower values of  $s$ , as depicted in Figure 1b,c, all trajectories converge towards the AD strategy. At these parameter values, rational individuals opt for the AD strategy because the punishment executed by the institution is insufficiently severe (i.e.,  $s$  is small) and the probability of being punished is too low (i.e.,  $\delta$  is small). Consequently, the punitive institution fails to function effectively in this scenario.

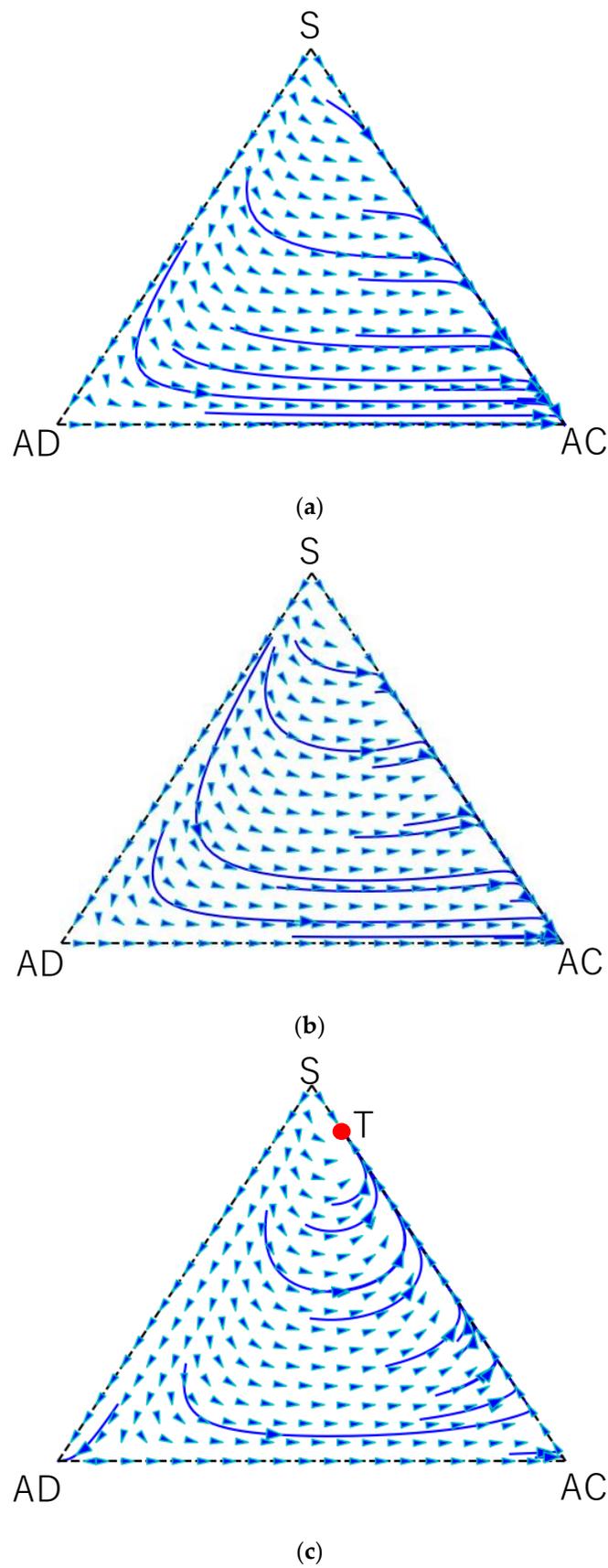
A notable difference between Figure 1b ( $s = 5.1$ ) and Figure 1c ( $s = 4.6$ ) is observed in the vector directions on the edge between AC and S. In the former scenario, AC is able to invade S, whereas, in the latter, S invades AC. However, it is important to note that S can never achieve evolutionary stability as it is susceptible to invasion by AD. Consequently, we can deduce that the pay-it-forward strategy, S, does not significantly contribute to the evolution of cooperation with the institutional punishment within the EUT framework.

#### 4.2. Numerical Results in Case of PT

Figure 2a–c illustrate the trajectories and vector fields when prospect theory is applied, using the same parameter values as in Figure 1. We highlight the following two observations: (1) Figure 2b demonstrates that, even with parameter values for which AD is the unique evolutionarily stable strategy under EUT, cooperation evolves when players adhere to PT; that is, AC is the ESS. (2) In Figure 2c, a new stable fixed point, called T, emerges on the edge between S and AC. This fixed point represents a mixed state mainly comprising S-players, with AC-players also being present. The system exhibits bi-stability, as some trajectories converge towards the vertex AD, depending on their initial conditions.

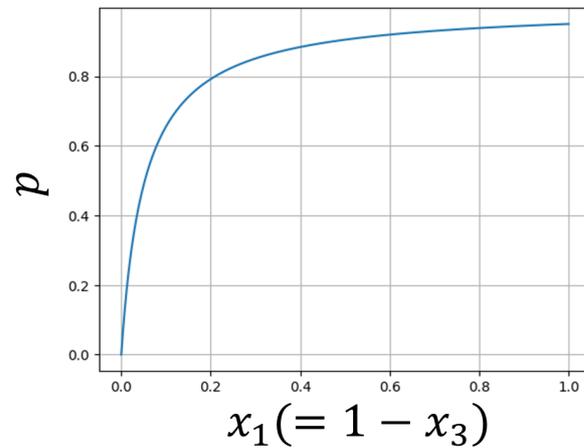
In the case of Figures 1 and 2c in particular, the AD strategy has several advantages. First, the condition  $b > s$  indicates that the punishment is not severe enough to offset the benefits gained from free-riding. In addition, with  $\delta$  being relatively small, choosing the D strategy entails an 80% chance of avoiding punishment. As such, AD players face a low probability (0.2) of being penalized despite consistently choosing D. Moreover, AD players receive an additional payoff from  $d$ , as they incur no costs to support the institution.

It should be noted that, if all players were to adopt strategy S, cooperation would fail to materialize. This is due to the spread of action D across the population as a result of implementation errors. Consequently, at vertex S, the rate of cooperation is zero. This phenomenon can be understood by examining the equation for the variable  $p$ , as discussed earlier. Importantly,  $p$  represents the cooperation rate of the entire population, which equals zero when  $x_1 = 0$  (i.e., in the absence of AC-players).



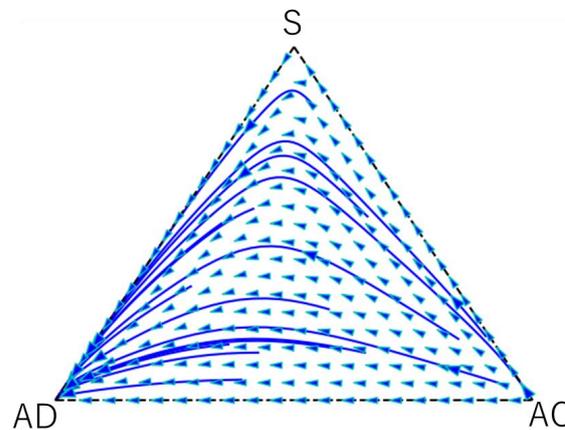
**Figure 2.** Trajectories and vector fields yielded by replicator dynamics for the case of prospect theory, with  $s = 5.6$  (a),  $s = 5.1$  (b) and  $s = 4.6$  (c). Parameter values are the same as in Figure 1.

Plotting the value of  $p$  (the “cooperation rate”) as a function of  $x_1$  on the edge between S and AC reveals that  $p$  increases sharply with even small increments in  $x_1$  (see Figure 3). Since  $x_1$  and  $x_3$  are related by  $x_3 = 1 - x_1$  along this edge,  $p$  effectively becomes a function of a single variable. At the stable fixed-point T, where  $x_1$  is approximately 0.1, the value of  $p$  is around 0.6, indicating that more than half of the population engages in helping others.



**Figure 3.** Value of  $p$  as a function of  $x_1$  on edge between S and AC, with the same parameter values used in Figure 2. Note that  $p$  is equivalent to cooperation rate in population.  $x_1 = 0$  corresponds to cooperation rate at vertex S, and  $x_1 = 1$  to vertex AC. Note that  $x_1$  and  $x_3$  are not independent, as  $x_3 = 1 - x_1$  holds. Therefore,  $p$  is a function with one variable ( $x_1$ ). We see  $p = 0$  at  $x_1 = 0$ . However, the value of  $p$  sharply increases as  $x_1$  increases, as far as  $x_1$  is small.

Figure 4 demonstrates that, when the punishment strength  $s$  becomes smaller ( $s = 3$  is chosen for this figure), all trajectories converge to the vertex AD. In this case, qualitative differences between EUT and PT are eliminated.



**Figure 4.** Trajectories and vector field yielded by replicator dynamics in the case of prospect theory with  $s = 3$ . Other parameter values are the same as in Figures 1 and 2.

**5. Conclusions**

We have seen that there is a set of parameter values with which prospect theory leads to the evolution of pay-it-forward behaviors and with which linear expected utility theory does not. Ironically, punishment is more effective when people are irrational. One reason why even minimal punishment can be highly effective lies in the fact that players disproportionately fear the small possibility of punishment, perceiving it as more likely to occur than it actually does. Cognitive distortions described by PT significantly contribute to mitigating the risk of catastrophic population collapse due to free-riding.

As mentioned in the Introduction, in [11] it was demonstrated that combining prospect theory with punishment enables it to evolve and sustain the unconditional strategy AC,

thus preventing free-riders from infiltrating a population of cooperators. In this paper, we applied a similar logic to pay-it-forward and have shown that prospect theory, coupled with slight punishment, facilitates the evolution of pay-it-forward. Given that the fixed-point  $T$  is located near vertex  $S$ , we can assert that the pay-it-forward strategy is “risk-dominant” over the unconditional strategy  $AC$ , making it more advantageous than unconditional altruistic behaviors.

We noted that the probability of an  $S$ -player helping another player  $X$  does not rely on  $X$ 's strategy, indicating that  $S$ -players do not discriminate among other players. However, an  $S$ -player's decision to help another is contingent upon whether the  $S$ -player received help in the previous game. Thus, while pay-it-forwarders do not discriminate, their actions are conditional. Taking these facts into account, we suggest that strategy  $S$  serves as an intermediate strategy between unconditional altruistic strategies, such as  $AC$ , and conditional cooperative strategies like “stern-judging”, “simple-standing”, or “image-scoring” in downstream indirect reciprocity.

Indeed, the bi-stable structure depicted in Figure 2c resembles the structure commonly observed in studies on downstream indirect reciprocity, where it has been found that, if players adopt strategies that consider so called “second-order information”, such as stern-judging or simple-standing in downstream reciprocity, the system exhibits bi-stability. Conversely, we have demonstrated that, when players adhere to prospect theory, upstream reciprocity exhibits a similar structure.

Image-scoring, a strategy in downstream reciprocity that relies solely on first-order information, is insufficient for the evolution of cooperation within the framework of EUT. Consequently, image-scoring fails to sustain stable cooperation in traditional settings [6], even though experimental evidence suggests that real humans only consider first-order information. Investigations into how the cognitive distortions described by prospect theory influence the stability of conditional strategies that use only first-order information in downstream indirect reciprocity present an intriguing research avenue.

**Author Contributions:** Conceptualization, S.U. and H.Y.; methodology, S.U.; formal analysis, S.U.; writing—original draft preparation, S.U.; writing—review and editing, S.U, T.S., H.Y. and I.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by JSPS KAKENHI, Grant Numbers JP23K05943(TS), JP23K21017 (IO, HY), JP21KK0027(IO, HY), and JP23K25160(HY, IO).

**Data Availability Statement:** The authors confirm that the article has no data.

**Conflicts of Interest:** The authors declare no conflicts of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

## Appendix A

The frequency  $p$  changes over time  $\tau$ . It is important to note that we use a distinct symbol,  $\tau$ , for the time during which  $p$  changes, as opposed to the symbol  $t$ , which is usually used to describe changes in  $x_i$ . The proportion of players participating in a game within the short time interval  $[\tau, \tau + \Delta\tau]$  is proportional to  $\Delta\tau$ , denoted as  $a\Delta\tau$ , with  $a$  being some positive constant. The probability that these players receive help in a game is given by  $(1 - \varepsilon)x_1 + (1 - \varepsilon)p x_3$ . The first term represents the probability of encountering  $AC$ -players who do not make an implementation error, while the second term pertains to encounters with  $S$ -players who were previously given help and, consequently, will offer support, assuming they do not make an implementation error. As a result, the frequency of players who have received help at time  $\tau + \Delta\tau$  is given by

$$p(\tau + \Delta\tau) = (1 - a\Delta\tau)p(\tau) + a\Delta\tau[(1 - \varepsilon)x_1 + (1 - \varepsilon)p(\tau)x_3].$$

Taking the limit as  $\Delta\tau \rightarrow 0$ , we derive the following time evolution equation for  $p$ :

$$\frac{dp}{d\tau} = -ap + a[(1 - \varepsilon)x_1 + (1 - \varepsilon)px_3],$$

the equilibrium point of which is  $p^* = \frac{(1-\varepsilon)x_1}{1-(1-\varepsilon)x_3}$ . This equilibrium is stable since the derivative of  $p$  with respect to  $\tau$  (i.e.,  $dp/d\tau$ ) is positive if  $p < p^*$  and negative if  $p > p^*$ .

### Appendix B

We briefly outline the essence of prospect theory utilized in this paper, focusing on the value function and the weighting function [12–14]. It is important to note that the following description closely mirrors Section 2.2.2 of reference [11].

As mentioned in the main text, each element  $(B_i)_{mn}$  of the probability distribution  $B_i$  gives the probability that the outcome  $(m, n)$  occurs in a game. In the framework of PT, players calculate this probability distribution subjectively. The subjective probability distribution is given by applying a nonlinear function, called a weighting function, to the original probability  $(B_i)_{mn}$ :

$$(W_i)_{mn} = w((B_i)_{mn}),$$

with

$$w(x) = \frac{x^\gamma}{[x^\gamma + (1 - x)^\gamma]^{\frac{1}{\gamma}}},$$

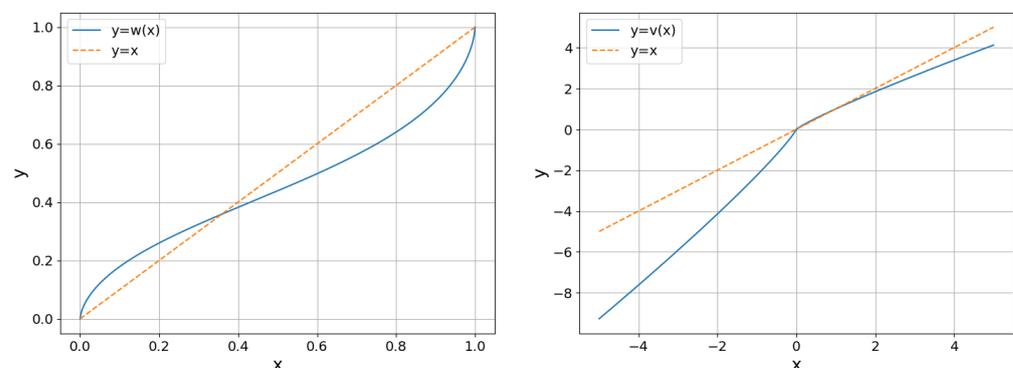
where  $x$  is an objectively given probability.

We set the value of parameter  $\gamma$  to 0.65, which is an often-used value in prospect theory. The shape of the function with this parameter value is shown in the left panel of Figure A1 (solid curve) and, together in the case with  $\gamma = 1$ , drawn as a dashed line, which corresponds to objectively given probabilities. According to this function, small objective probabilities are estimated to be greater than they are because  $w(x) > x$  for small  $x$ .

The value function has the form

$$v(x) = \begin{cases} x^\alpha & (x \geq 0) \\ -\lambda(-x)^\alpha & (x \leq 0) \end{cases}$$

where  $x$  represents an objectively given payoff. We choose the values of the two parameters  $\alpha$  and  $\lambda$  to be 0.88 and 2.25, respectively according to the literature on prospect theory. The right panel in Figure A1 shows the value function with these parameter values (solid curve) together with the linear function (drawn by the dashed line). Parameter  $\alpha < 1$  implies that the value function is concave. A person obeying this function is called risk-averse. Parameter  $\lambda > 1$  means that a person following this function is more sensitive to a loss than a gain (“loss-averse”).



**Figure A1.** (Left panel) Weighting function (solid curve)  $w(x)$  with parameter  $\gamma = 0.65$ . Horizontal axis represents objectively given probabilities  $x$ , and vertical axis represents subjective probability denoted

by  $y$ . Linear function with  $\gamma = 1$  corresponding to linear expected utility theory is also displayed (dashed line). **(Right panel)** Value function (solid curve)  $v(x)$  with parameters  $\alpha = 0.88$ ,  $\lambda = 2.25$ . Argument of the function  $x$  represents objectively given outcomes and function outputs subjective values. Linear function with  $\alpha = \lambda = 1$  is also shown (dashed line).

## References

1. Fowler, J.H.; Christakis, N.A. Cooperative behavior cascades in human social networks. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5334–5338. [[CrossRef](#)]
2. Pressman, S.D.; Kraft, T.L.; Cross, M.P. It's good to do good and receive good: The impact of a “pay it forward” style kindness intervention on giver and receiver well-being. *J. Posit. Psychol.* **2015**, *10*, 293–302. [[CrossRef](#)]
3. Nowak, M.A.; Roch, S. Upstream reciprocity and the evolution of gratitude. *Proc. R. Soc. B* **2007**, *274*, 605–610. [[CrossRef](#)]
4. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*; Cambridge University Press: Cambridge, UK, 1990.
5. Nowak, M.A. *Evolutionary Dynamics*; Harvard University Press: Cambridge, MA, USA, 2006.
6. Sigmund, K. *The Calculus of Selfishness*; Princeton University Press: Princeton, UK; Oxford, UK, 2010.
7. Nowak, M.A.; Highfield, R. *Super Cooperators*; Free Press: New York, NY, USA, 2011.
8. Bowls, S.; Gintis, H. *A Cooperative Species*; Princeton University Press: Princeton, UK; Oxford, UK, 2011.
9. Yamagishi, T. *Trust: The Evolutionary Game of Mind and Society*; Springer: New York, NY, USA, 2011.
10. Perc, M.; Jordan, J.J.; Rand, D.G.; Wang, Z.; Boccaletti, S.; Szolnoki, A. Statistical physics of human cooperation. *Phys. Rep.* **2017**, *68*, 1–51. [[CrossRef](#)]
11. Uchida, S.; Yamamoto, H.; Okada, I.; Sasaki, T. Evolution of Cooperation with Peer Punishment under Prospect Theory. *Games* **2019**, *10*, 11. [[CrossRef](#)]
12. Kahneman, D.; Tversky, A. Prospect theory: Analysis of decision under risk. *Econometrica* **1979**, *47*, 263–291. [[CrossRef](#)]
13. Tversky, A.; Kahneman, D. Loss aversion in riskless choice: A reference-dependent model. *Q. J. Econ.* **1991**, *106*, 1039–1061. [[CrossRef](#)]
14. Wakker, P.P. *Prospect Theory: For Risk and Ambiguity*; Cambridge University Press: Cambridge, UK, 2010.
15. Gilboa, I. *Theory of Decision under Uncertainty*; Cambridge University Press: Cambridge, UK, 2009.
16. Gilboa, I. *Making Better Decisions: Decision Theory in Practice*; John Wiley and Sons: Tokyo, Japan, 2010.
17. Tversky, A.; Kahneman, D. Judgement under uncertainty: Heuristics and biases. *Science* **1974**, *185*, 1124–1131. [[CrossRef](#)]
18. Machina, M.J. Expected utility analysis without the independence axiom. *Econometrica* **1982**, *50*, 277–323. [[CrossRef](#)]
19. Tversky, A.; Kahneman, D. Extensional vs. intuitive reasoning: The conjunction fallacy in probability judging. *Psychol. Rev.* **1983**, *90*, 293–315. [[CrossRef](#)]
20. Schmeidler, D. Subjective probability and expected utility without additivity. *Econometrica* **1989**, *57*, 571–587. [[CrossRef](#)]
21. Gilboa, I.; Schmeidler, D. Maxmin expected utility with a non-unique prior. *J. Math. Econ.* **1989**, *18*, 141–153. [[CrossRef](#)]
22. Starmer, C. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *J. Econ. Lit.* **2000**, *38*, 332–382. [[CrossRef](#)]
23. Yang, G.; Chen, J. Research on Enterprise Radical Technology Innovation Behavior—Evolutionary Game Analysis Based on Prospect Theory. *J. Ind. Technol. Econ.* **2020**, *39*, 57–64.
24. Sun, H.; Gao, G.; Li, Z. Evolutionary Game Analysis of Enterprise Carbon Emission Regulation Based on Prospect Theory. *Soft Comput.* **2022**, *26*, 13357–13368. [[CrossRef](#)]
25. Shen, J.; Gao, X.; He, W.; Sun, F.; Zhang, Z.; Kong, Y.; Wan, Z.; Zhang, X.; Li, Z.; Wang, J.; et al. Prospect Theory in an Evolutionary Game: Construction of Watershed Ecological Compensation System in Taihu Lake Basin. *J. Clean. Prod.* **2021**, *291*, 125929. [[CrossRef](#)]
26. Li, Y.; Zhang, J. Evolutionary Game Analysis of Low-Carbon Incentive Behaviour of Power Battery Recycling Based on Prospect Theory. *Sustainability* **2024**, *16*, 2793. [[CrossRef](#)]
27. Sigmund, K. *Games of Life*; Dover Publications: Mineora, NY, USA, 2017.
28. Ferreira, P.L.; Santos, F.C.; Pequito, S. Risk sensitivity and theory of mind in human coordination. *PLoS Comput. Biol.* **2021**, *17*, e1009167. [[CrossRef](#)]
29. Nowak, M.A. Five rules for the evolution of cooperation. *Science* **2006**, *314*, 1560–1563. [[CrossRef](#)]
30. Nowak, M.A.; Sigmund, K. Evolution of indirect reciprocity. *Nature* **2005**, *437*, 1292–1298. [[CrossRef](#)]
31. Nowak, M.A.; Sigmund, K. The dynamics of indirect reciprocity. *J. Theor. Biol.* **1998**, *194*, 561–574. [[CrossRef](#)]
32. Leimar, O.; Hammerstein, P. Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. B* **2001**, *268*, 745–753. [[CrossRef](#)]
33. Ohtsuki, H.; Iwasa, Y. How should we define goodness?—Reputation dynamics in indirect reciprocity. *J. Theor. Biol.* **2004**, *231*, 107–120. [[CrossRef](#)]
34. Ohtsuki, H.; Iwasa, Y. The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* **2006**, *239*, 435–444. [[CrossRef](#)]
35. Sasaki, T.; Okada, I.; Nakai, Y. The evolution of conditional moral assessment in indirect reciprocity. *Sci. Rep.* **2017**, *7*, 41870. [[CrossRef](#)]

36. Uchida, S.; Yamamoto, H.; Okada, I.; Sasaki, T. A theoretical approach to norm ecosystems: Two adaptive architectures of indirect reciprocity show different paths to the evolution of cooperation. *Front. Phys.* **2018**, *6*, 14. [[CrossRef](#)]
37. Yamamoto, H.; Okada, I.; Uchida, S.; Sasaki, T. Exploring norms indispensable for both emergence and maintenance of cooperation in indirect reciprocity. *Front. Phys.* **2022**, *10*, 1019422. [[CrossRef](#)]
38. Fujimoto, Y.; Ohtsuki, H. Evolutionary stability of cooperation in indirect reciprocity under noisy and private assessment. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2300544120. [[CrossRef](#)]
39. Iwagami, A.; Masuda, N. Upstream reciprocity in heterogeneous networks. *J. Theor. Biol.* **2010**, *265*, 297–305. [[CrossRef](#)] [[PubMed](#)]
40. Pfeiffer, T.; Rutte, C.; Killingback, T.; Taborsky, M.; Bonhoeffer, S. Evolution of cooperation by generalized reciprocity. *Proc. R. Soc. B* **2005**, *272*, 1115–1120. [[CrossRef](#)] [[PubMed](#)]
41. Okada, I. A review of theoretical studies on indirect reciprocity. *Games* **2020**, *11*, 27. [[CrossRef](#)]
42. Ohtsuki, H.; Iwasa, Y. Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *J. Theor. Biol.* **2007**, *244*, 518–531. [[CrossRef](#)]
43. Uchida, S.; Sigmund, K. The competition of assessment rules for indirect reciprocity. *J. Theor. Biol.* **2010**, *263*, 13–19. [[CrossRef](#)]
44. Murase, Y.; Kim, M.; Baek, S.K. Social norms in indirect reciprocity with ternary reputations. *Sci. Rep.* **2022**, *12*, 455. [[CrossRef](#)]
45. Yamagishi, T. The provision of a sanctioning system as a public good. *J. Personal. Soc. Psychol.* **1986**, *51*, 110–116. [[CrossRef](#)]
46. Boyd, R.; Richerson, P.J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* **1992**, *13*, 171–195. [[CrossRef](#)]
47. Milinski, M.; Rockenbach, B. Human behaviour: Punisher pays. *Nature* **2008**, *452*, 297–298. [[CrossRef](#)]
48. Casari, M.; Luini, L. Cooperation under alternative punishment institutions: An experiment. *J. Econ. Behav. Organ.* **2009**, *71*, 273–282. [[CrossRef](#)]
49. Kosfeld, M.; Okada, A.; Riedl, A. Institution formation in public goods games. *Am. Econ. Rev.* **2009**, *99*, 1335–1355. [[CrossRef](#)]
50. Sigmund, K.; de Silva, H.; Traulsen, A.; Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **2010**, *466*, 861–863. [[CrossRef](#)]
51. Traulsen, A.; Röhl, T.; Milinski, M. An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proc. Biol. Sci.* **2012**, *279*, 3716–3721. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.