MDPI

Article

Individualism or Collectivism: A Reinforcement Learning Mechanism for Vaccination Decisions

Chaohao Wu [1], Tong Qiao [1], Hongjun Qiu [1,*], Benyun Shi [2] and Qing Bao [1]

1　School of Cyberspace, Hangzhou Dianzi University, Hangzhou 310018, China; wchwch1018@hdu.edu.cn (C.W.); tong.qiao@hdu.edu.cn (T.Q.); qbao@hdu.edu.cn (Q.B.)
2　School of Computer Science and Technology, Nanjing Tech University, Nanjing 211800, China; byshi@njtech.edu.cn
*　Correspondence: qiuhj@hdu.edu.cn

Abstract: Previous studies have pointed out that it is hard to achieve the level of herd immunity for the population and then effectively stop disease propagation from the perspective of public health, if individuals just make vaccination decisions based on individualism. Individuals in reality often exist in the form of groups and cooperate in or among communities. Meanwhile, society studies have suggested that we cannot ignore the existence and influence of collectivism for studying individuals' decision-making. Regarding this, we formulate two vaccination strategies: individualistic strategy and collectivist strategy. The former helps individuals taking vaccination action after evaluating their perceived risk and cost of themselves, while the latter focuses on evaluating their contribution to their communities. More significantly, we propose a reinforcement learning mechanism based on policy gradient. Each individual can adaptively pick one of these two strategies after weighing their probabilities with a two-layer neural network whose parameters are dynamically updated with his/her more and more vaccination experience. Experimental results on scale-free networks verify that the reinforcement learning mechanism can effectively improve the vaccine coverage level of communities. Moreover, communities can always get higher total payoffs with fewer costs paid, comparing that of pure individualistic strategy. Such performance mostly stems from individuals' adaptively picking collectivist strategy. Our study suggests that public health authorities should encourage individuals to make vaccination decisions from the perspective of their local mixed groups. Especially, it is more worthy of noting that individuals with low degrees are more significant as their vaccination behaviors can more sharply improve vaccination coverage of their groups and greatly reduce epidemic size.

Keywords: collectivism; reinforcement learning; vaccination

1. Introduction

The epidemics of infectious diseases pose a great threat to human health, such as early SARS [1–3], H1N1 [4–6], tuberculosis [7–9], and COVID-19 [10–13]. For preventing those diseases, vaccination has always been one of the most effective measures to be taken [14–17]. Research has testified that the vaccine coverage level is crucial to effectively prevent a large-scale epidemic of diseases such as eradicate influenza epidemics [18–20]. Thereby, designing an effective mechanism to increase the vaccine coverage level has become a significant task nowadays.

Several vaccination strategies and mechanisms [21–25] have been proposed to improve vaccine coverage levels. They emphasize various factors that affect vaccine decisions, such as infection costs, vaccine costs, social infection information that individuals received, or other social information. Xia and Liu [22] examine individuals' awareness about disease and vaccine and find that affects the ratio and speed of taking vaccines. Our previous work [21] also demonstrates that individuals' vaccination decisions can be influenced by their social neighbors as social influence plays a key role in epidemic control. Bauch's

work [26] has shown that for any perceived relative risk $r > 0$, the expected vaccine uptake level is less than the eradication threshold. Furthermore, many previous studies use game theory to study human vaccination behaviors [26–28]. They assume that individuals are rational and make vaccination decisions with respect to the costs of vaccination and risk of infection, i.e., the individualistic behavior specified in this paper. These strategies mostly consider behaviors that are beneficial to themselves, and have not yet taken into account individual altruistic motivations, which should not be ignored as Bauch pointed out in the previous study [28].

Altruism is indeed an important factor in promoting individual vaccination [29]. Reciprocity rules of cooperation confirm the existence of individuals' altruistic thoughts [30]. Individuals in reality often exist in the form of groups, and they achieve cooperation in various ways [31]. For example, an individual taking a vaccine cannot only prevent him/herself from being infected by diseases, but also improve the level of vaccine coverage and thereby indirectly protect others around him/herself. Hamilton's rule of "kin selection" [32] indicates that natural selection can favor cooperation if the donor and the recipient of an altruistic act are genetic relatives [33–35]. Meanwhile, the rule of "indirect reciprocity" [36,37] represents that people helping someone establish a good reputation can be rewarded by third parties. It is exactly this way that people consider their reputation before making a decision, and thus make decisions that benefit others. Given that, in this paper, we assume that when an individual makes a decision, he/she should consider the impact on the people associated with him/her, such as relatives, friends, or colleagues. That is, individuals prefer to take vaccines once their action benefits their associates.

In recent studies on vaccination, altruism has been studied more and more. Meanwhile, altruism has also been shown to have a positive effect on vaccination. However, in previous research, most of the research methods on altruism are questionnaire surveys, network testing, and game theory. These studies provide significant qualitative and theoretical evidence for applying altruism to vaccination [38,39]. The impacts on disease epidemiology and vaccination cost need to be further quantitatively studied. Regarding this, we introduce a specified formulation and measurement of the impact of altruism. Moreover, we proposed a collectivism strategy based on altruism, and quantitatively examining how it affects and how much impact it can have.

In reality, it is difficult for individuals to be completely altruistic or completely selfish when making decisions. Thereby, we should take the two strategies mentioned above into consideration at the same time, namely, the strategy of individualism based on selfishness and the strategy of collectivism based on altruism. We naturally have a question: how do individuals balance selfishness and altruism when making vaccination decisions? In order to solve this question, we introduce a reinforcement learning mechanism of policy gradient to combine two strategies that individuals may consider in vaccination decisions. Under the premise of being voluntary, individuals can dynamically choose the optimal vaccination strategy with an individualistic strategy or collectivist strategy based on their historical decisions with vaccination and associated payoffs [27]. Individualistic strategy namely denotes individuals only evaluate the cost and payoff of themselves during the process of making vaccination decisions. As for collectivist strategy, individuals not only consider the costs of vaccination, but also consider the impact of their own vaccination strategy on the people around them when deciding on vaccination. Moreover, we also consider the impact of the concept of "neighbors of neighbors" [40]. Here, we assume that individuals form a locally mixed "network of network" with her/his neighbors through their social connections. For a realistic social environment, an individual does not just exist in an isolated local-mixed community. As shown in Figure 1a, Alice (red individual) forms a locally-mixed group with her four neighbors. Meanwhile, she is still in the local-mixed groups of her neighbors Bob and Cindy, as illustrated in Figure 1b,c. Therefore, we should consider the impact of individuals' vaccination strategies on the vaccination coverage level of all communities they exist. In other words, the influence of individuals on their "neighbors of neighbors" should be taken into consideration. Our experiments suggest that

reinforcement learning mechanism and collectivism have significantly promoted the level of vaccine coverage. In addition, the reinforcement learning mechanism has also shown advantages in costs and payoffs. Remarkably, we find that the contribution of individuals (nodes) with low degrees to the group cannot be ignored.
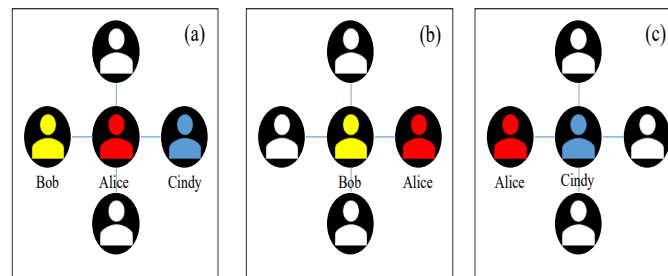
Figure 1. Locally-mixed groups in social networks. As shown in the subfigure (**a**), Alice (the red one) forms a locally-mixed group with her four neighbors. Meanwhile, she is also a member of locally-mixed groups of her neighbors Bob and Cindy, as shown in the right two figures, i.e., subfigures (**b**,**c**). Therefore, Alice's vaccination strategy can also affect other members in the groups of Bob and Cindy. We define this effect as the influence of individuals on their "neighbors of neighbors".

This paper is organized as follows. In Section 2, we formally propose a voluntary vaccination mechanism based on reinforcement learning. In addition, the individualistic strategy and the collectivist strategy are introduced separately. In Section 3, we analyze various performances under the reinforcement learning mechanism, such as vaccine coverage level, epidemic size, payoffs, and costs, and consider the effectiveness of collectivism. The discussions on altruism are presented in Section 4. Finally, Section 5 summarizes the full text.

2. Methods

In this section, we introduce a vaccination mechanism based on reinforcement learning. For most individuals, their vaccination decisions are often made based on evaluating vaccination costs and infection costs which means individuals are selfish, and results in a low level of vaccine coverage. Here, we introduce a collectivist strategy to portray individual altruism, i.e., individuals consider the impact of their decision on their associations before making a decision. However, we believe that individuals are bounded-rational, and pure altruism may harm the individual's own interests. Thereby, we rely on a reinforcement learning mechanism to help individuals choose the optimal vaccination strategy while increasing the vaccine coverage level.

The interactive process of vaccination dynamic and transmission dynamic is modeled as an iterative two-stage process. In the first stage, each individual can make their vaccination decision with reinforcement learning mechanism via selecting individualistic strategy or collectivist strategy. We use P^S and P^A to represent the probabilities of choosing the individualistic strategy and collectivist strategy, respectively. In the second stage, the classic SIR model [41] is used to simulate the spread and recovery of the disease. The population is divided into three compartments—susceptible individuals (S), infectious individuals (I), and recovered individuals (R)—in this model.

Figure 2 describes the interactive process with the reinforcement learning mechanism of policy gradient, which has three principal components, i.e., **actions**, **observations**, and **rewards**. At first, individuals evaluate the probabilities of picking individualistic strategy and collectivist strategy, and pick the one with a higher probability (the specification of the component **actions**). According to their strategies, individuals perform vaccination behaviors or not, and become *vaccinated* or *unvaccinated* (the specification of the component **observations**). Then, after the phase of disease propagation, individuals turn to be *immune*, *infected*, and *free-riders*, and obtain corresponding payoff $1 - c$, 0, and 1 (the specification of the component **rewards**), respectively. Here, $c = \frac{C_V}{C_I}$ denotes the relative cost of

vaccination and infection. Those components of **actions**, **observations**, and **rewards** as feedback support individuals evaluate action probabilities in the forthcoming process. The detailed reinforcement learning process with policy gradient is shown in Figure 3. This process illustrates individuals' input(observation and reward) and output(the probability distribution of the action). The individual selects the action with the highest probability as the operation to be performed.
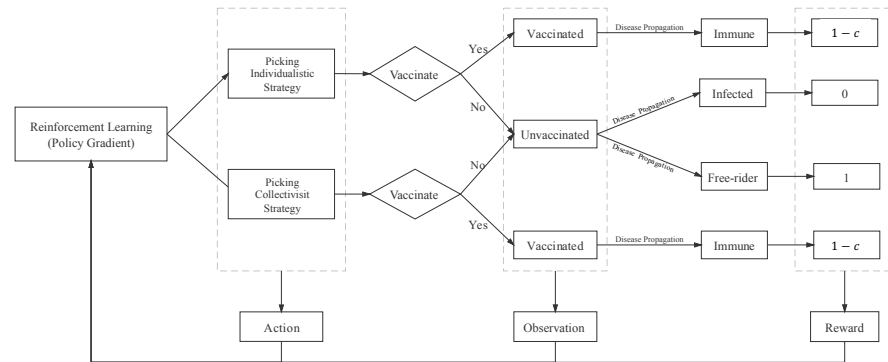
Figure 2. The decision-making process based on the reinforcement learning mechanism of policy gradient, which has three principal components, i.e., **actions**, **observations** and **rewards**. At first, individuals evaluate the probabilities of picking individualistic strategy and collectivist strategy, and pick the one with a higher probability (the specification of the component **actions**). According to their strategies, individuals perform vaccination behaviors or not, and become *vaccinated* or *unvaccinated* (the specification of the component **observations**). Then, after the phase of disease propagation, individuals turn to be *immune*, *infected*, and *free-riders*, and obtain corresponding payoff $1 - c$, 0, and 1, respectively (the specification of the component **rewards**). Here, $c = \frac{C_V}{C_I}$ denotes the relative cost of vaccination and infection. Those components of **actions**, **observations**, and **rewards** as feedback support individuals evaluate action probabilities in the forthcoming process.
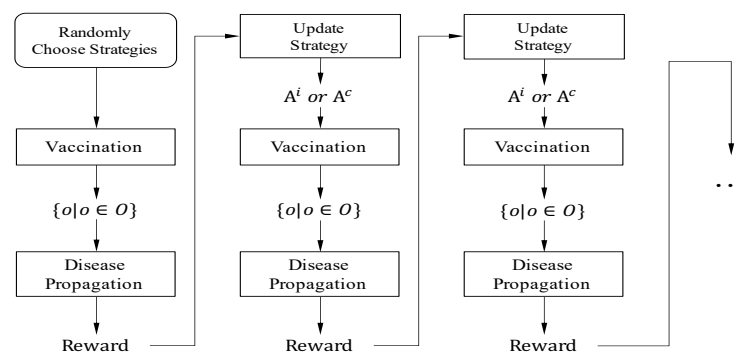
Figure 3. The reinforcement learning process with policy gradient: in the first season, individuals randomly choose vaccination strategies and enter vaccination and disease propagation to obtain the corresponding reward. Subsequently, they pick an optimal strategy according to previous choices and reward, and take or not take a vaccine for the next season of disease propagation. They repeat such operations until the process ends.

The following sections specify individuals' strategies and reinforcement learning mechanism, respectively.

2.1. Individualistic Strategy of Vaccination

Our previous study [21] introduced a memory-based vaccination mechanism in which individuals are considered to have memories of vaccination strategies and infection experiences in previous seasons. Individuals make vaccination decisions by considering their previous optimal vaccination probability and costs. Therefore, it is assumed that

individuals are essentially following individualistic nature, i.e., individuals only consider their own interests to make vaccination choices. The vaccination probability of individual i based on individualistic strategy is calculated as

$$P_i^S(n) = (1 - \varepsilon)P_i^S(n-1) + \varepsilon P_i^*(n) \tag{1}$$

Here, $P_i(n-1)$ represents the individual i's vaccination probability in season $n-1$. ε is the factor that measures the memory of the optimal vaccination probability. P^* is the optimal vaccination probability for the individual i at season n, which can be derived by achieving an equilibrium of vaccination costs and infection cost:

$$- P^* \times C_V = -r(p) \times (1 - P^*) \times C_I \tag{2}$$

where C_V and C_I denote vaccination costs and infection costs, respectively. $-r(p)$ means i's infection risk.

2.2. Collectivist Strategy of Vaccination

In this section, we introduce the concept of collectivistic strategy, that is, individuals choose to vaccinate because they prefer to make favorable contributions to the community's vaccination coverage level p. The contribution of an individual's vaccination to his community can be defined as the sum of increased payoff due to the decreased infection risk of the community as the result of the vaccination behavior.

$$contribution = \sum^{j}(\triangle r_j \times S_j \times C_I) \tag{3}$$

It is worthy of note that the vaccinated individual pays a certain cost for taking vaccine. Then, we can further specify the contribution in the perspective of the whole community as

$$V_A = \sum^{j}(\triangle r_j \times S_j \times C_I) - C_V \tag{4}$$

where S_j is the number of individuals in the susceptible state in any community, and j is the number of communities in which individual exists. $\triangle r$ is defined as the difference value between the risk of infection of neighbors before an individual vaccination and the perceived risk of neighbors after vaccination, which can be calculated as

$$\triangle r = r(p) - r(p') = [1 - \frac{1}{R_0(1 - \frac{n_v}{N})}] - [1 - \frac{1}{R_0(1 - \frac{n_v+1}{N})}] \tag{5}$$

where N is the total number of individuals in the community and n_v is the number of vaccinated individuals in the community. In the context of epidemic modeling, R_0 is usually defined as the so-called basic reproductive rate [42].

If we consider relative cost c rather than separately considering vaccine costs C_V and infection costs C_I, we can modify Equation (4) as follows,

$$V_A = \sum^{j}(\triangle r_j \times S_j) - c \tag{6}$$

Then, we can calculate the probability that an individual taking vaccine with Fermi function [43,44], in which the contribution V_A is used as a driving force.

$$P^A = \frac{1}{1 + \exp[-\beta V_A]} = \frac{1}{1 + \exp\{-\beta[\sum^{j}(\triangle r_j \times S_j) - c]\}} \tag{7}$$

where β is the selection strength.

2.3. Reinforcement Learning Mechanism Based on Policy Gradient

In this article, we propose a reinforcement learning mechanism under the framework of policy gradient to help individuals adaptively pick a vaccination strategy to obtain optimal rewards. Policy gradient is usually modeled as an optimization function with a parameter θ [45]. It predicts the probabilities of actions to be taken next based on the current environment, and then performs the action with the highest probability. As for our issue of picking vaccination strategies, individuals use actions, observations, and rewards in the previous season to evaluate the probabilities. Here, actions denote picking individualistic strategy (A^i) or picking collectivist strategy (A^c). Observations denote the four possible states, i.e., *vaccinated* under collectivist strategy (O_v^c), *unvaccinated* under collectivist strategy (O_u^c), *vaccinated* under individualistic strategy (O_v^i), and *unvaccinated* under individualistic strategy (O_u^i), respectively. Rewards denote the possible payoff $1 - c$, 0, and 1.

Then, we set up a two-layer neural network as shown in Figure 4. Four possible state observations are being used as input. 10 neurons are in the hidden layer with a tanh activation function to achieve the conversion from the linear model to the nonlinear model. Then, the probabilities of picking individualistic strategy and collectivist strategy are set as the output.
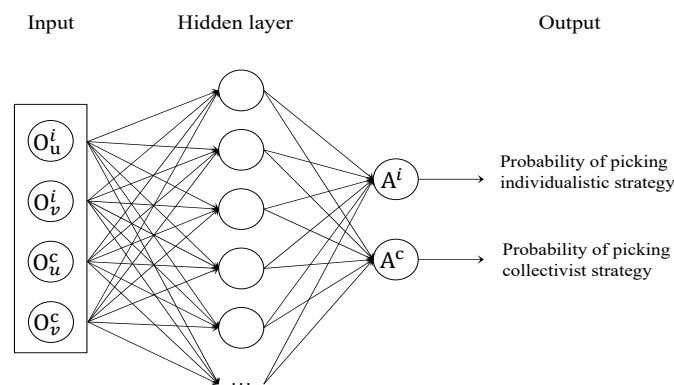
Figure 4. Our neural network with two fully connected layers. Four possible state observations are used as input and we set up 10 neurons in the hidden layer. In the hidden layer, we use the tanh activation function to achieve the conversion from the linear model to the nonlinear model, and initialize the weights randomly. Then, the probabilities of picking individualistic strategy and collectivist strategy are set as the output.

For our problem of vaccination strategy, we can consider it as an episodic environment. In the episodic environment (the simulation process from the beginning to the end is called an episode; the system learns strategies by simulating the episode again and again), the objective function measures the value calculated from the beginning state.

$$J(\theta) = V^{\pi_\theta}(o_t) = \mathbb{E}_\pi[G_t \mid O_t = o] \tag{8}$$

where J is the target policy and π_θ is a distribution over actions with given states which can be parameterized into $\pi_\theta(o, a) = P(a_t \mid o_t, \theta)$. The state-value function V could obtain expected reward, if an individual starts in the state o, and then followed the policy π at all the following time steps. G is defined as the cumulative reward that an individual can obtain after a certain time. O stands for state space, and could be any of O_v^i, O_u^i, O_v^c, O_u^i, and o.

The policy gradient theorem [45] suggests that no matter which function is adopted, the objective function can be further formulated under a multi-step MDP as follows,

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}[\nabla_\theta \log \pi_\theta(o, a) Q^{\pi_\theta}(o_t, a_t)] \tag{9}$$

where Q is the state action value function used to quantitatively evaluate the reward, once an individual takes action a in a certain state o. In actual optimization, the stochastic gradient ascent algorithm is used to perform unbiased sampling of $Q^{\pi_\theta}(o_t, a_t)$, which is recorded as v_t. Thus, the expected term can be removed, and the equation can be written by

$$\theta \leftarrow \theta + \alpha \nabla_\theta \log \pi_\theta(o, a) v_t \tag{10}$$

where α is the learning rate and a belongs to the individual's action space A which consists of A^i and A^c.

In the reinforcement learning mechanism, U will be regarded as a reward to update the neural network parameters. For our vaccination issue, we have three types of possible individuals' reward, i.e., payoff from vaccination, payoff from unvaccinated infection, and payoff from unvaccinated and uninfected (free-rider). The three corresponding forms of reward and infection status are shown in Equation (11).

$$U_i(n-1) = \begin{cases} 1 - c, & \text{if } i \text{ is vaccinated} \\ 0, & \text{if } i \text{ is infected} \\ 1, & \text{if } i \text{ is free-rider} \end{cases} \tag{11}$$

Algorithm 1 describes how our model works, including the process of disease propagation, vaccination, and reinforcement learning. Algorithm 2 demonstrates the reinforcement learning process. Moreover, Algorithm 3 specifies the vaccination process under the reinforcement learning mechanism.

Algorithm 1 Simulation process().

Input: N: the number of network nodes; *Seasons*: the number of seasons;
Output: epidemic size; vaccination coverage level; number of unvaccinated individuals;

  for $n = 1 \rightarrow Seasons$ **do**
    if $n == 1$ **then**
      // first season
      initialize θ arbitrarily;
      for $i = 1 \rightarrow N$ **do**
        make vaccination decision randomly;
      end for
    else
      // other seasons
      Reinforcement learning();
      Vaccination();
    end if
    randomly set I_0 unvaccinated individuals as infectious;
    set time step $t = 0$;
    $I(t) \leftarrow I_0$;
    while $I(t) > 0$ **do**
      disease spreading using Gillespie algorithm;
      $t++$;
      $I(t) \leftarrow$ the number of infected number;
    end while
    $p \leftarrow$ vaccination coverage level among i's neighbors;
    $S_j \leftarrow$ number of unvaccinated individuals of i's neighbors and "neighbors of neighbors";
    Output epidemic size;
    Output vaccination coverage level;
    Output number of unvaccinated individuals;
  end for

Algorithm 2 Reinforcement learning().

Input: o, a, v_t;
Output: *action*;

  $\theta \leftarrow \theta + \alpha \nabla_\theta \log \pi_\theta(o, a) v_t$;
  action $\leftarrow$ Neural Networks with parameter θ;
  Output action;
  return θ

Algorithm 3 Vaccination().

Input: p, S_j, R_0, relative cost c, *action*
Output: individuals' vaccination decision $S_i(n)$

  for $i = 1 \rightarrow N$ **do**
    if $action_i == picking\ individualistic\ strategy$ **then**
      $r_i(p) = 1 - \frac{1}{R_0(1-p)}$;
      $P_i^* = \frac{r_i(p)}{r_i(p)+c}$;
      $P_i = (1 - \varepsilon)P_i(n-1) + \varepsilon P_i^*(n)$;
    else
      //picking collectivist strategy
      $r_i(p') = 1 - \frac{1}{R_0(1-\frac{n_v+1}{N})}$;
      $\triangle r_i = r_i(p) - r_i(p')$;
      $V_A = \sum^j(\triangle r_{ij} \times S_j) - c$;
      $P_i = \frac{1}{1+\exp[-\beta V_A]}$;
    end if
    if a random number $< P_i(n)$ **then**
      $S_i(n) = 1$; // vaccinate
    else
      $S_i(n) = 0$; // do not vaccinate
    end if
  end for

---

3. Experimental Results

In this section, we perform a series of simulation experiments to verify the effectiveness of our method. The experimental settings are specified as follows:

- Network structure: simulation experiments are conducted in scale-free networks. Each network has $N_v = 1000$ nodes whose average degree are equal to four $<k> = 4$.
- Transmission parameters: disease transmission rate $r = 0.55$, disease recovery rate $g = 1/3$, reinforcement learning learning rate $\alpha = 0.05$, and selection strength $\beta = 1$ [46].
- Initial vaccine coverage rate: in the first season, each individual decides whether to be vaccinated with a probability of 0.5. Therefore, the initial season vaccine coverage rate is around 50%.

For robustness, each experiment is performed in randomly generated 50 networks, and runs 50 seasons which have the above-mentioned two stages: vaccination decision and disease propagation. Specially, we use the Gillespie Algorithm to simulate disease propagation in 2000 steps. Finally, the average values of vaccine coverage and infection scale in a steady (convergent) state are taken as experimental results. It is worth noting that the comparison of different strategies is conducted in the same network to eliminate the effect of the experimental setting.

3.1. Effectiveness of Reinforcement Learning Mechanism on Vaccination

With our reinforcement learning mechanism, individuals choose vaccines based on individualistic strategy or collectivist strategy according to the observations and rewards obtained in previous seasons. As Figure 5 illustrates, with the increase of relative cost

c, the vaccine coverage levels of three strategies are all decreasing while epidemic sizes increase. However, the vaccine coverage level and epidemic size of the reinforcement learning mechanism are in second-optimal performance, better than individualistic strategy. With such a mechanism, individuals choose strategies adaptively according to observed state and reward. This finding indicates that individuals can optimize their earnings under the reinforcement learning mechanism to adaptively make vaccination decisions. In addition, we can find that with the collectivist strategy, experiments have the highest level of vaccine coverage and the lowest epidemic size. This is because when individuals make vaccination decisions by collectivist strategy, they are more concerned with the contribution of their vaccination to the community rather than relative cost *c*. However, in the face of epidemic diseases, it is unrealistic for individuals to only consider the contribution of their vaccination behaviors to their community.
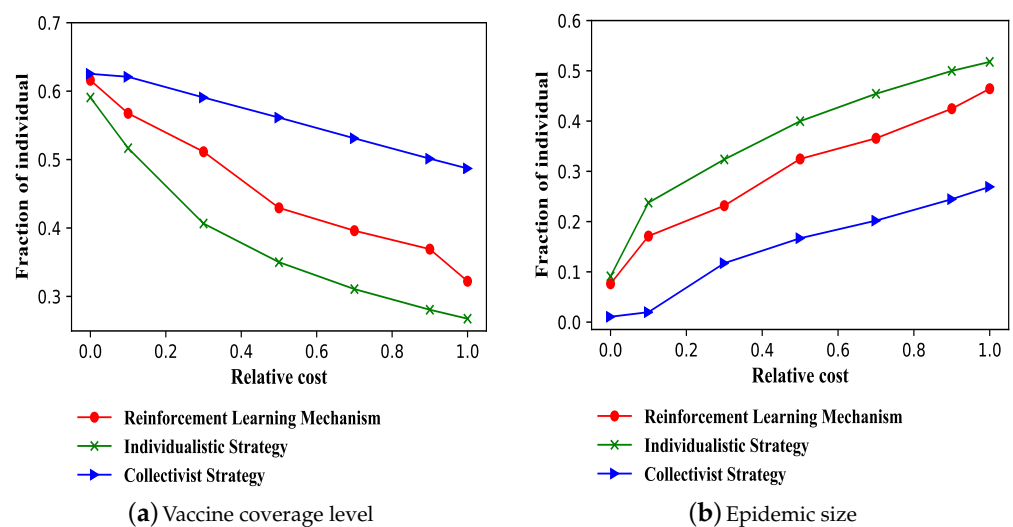
Figure 5. The comparison of results between reinforcement learning mechanism and two strategies. The vaccine coverage level is inversely proportional to *c*, and the epidemic size is directly proportional to *c*. In addition, we can find that collectivist strategy can always maintain the optimal result. Compared with the individualistic strategy, the reinforcement learning mechanism that taking collectivism into consideration has obtained better results under the same *c*.

3.2. Payoffs and Costs under Reinforcement Learning Mechanism

3.2.1. Payoffs and Costs of Population

In this section, we will conduct a more in-depth study on the impact of the reinforcement learning mechanism, and in particular, the impact of collectivist strategy with respect to the community-level costs and payoffs. Specially, we would like to further determine whether the higher vaccine coverage of the collectivist strategy is caused by high costs. If high costs are paid, the payoffs will be decreased. Accordingly, we study the total payoffs and total cost of the population under different mechanisms. For quantitative measurement, we define the total payoffs of the population as

$$
\begin{aligned}
\text{payoffs} &= num_v \times (1 - c) + num_f \times 1 + num_i \times 0 \\
&= num_v \times (1 - c) + num_f
\end{aligned}
\tag{12}
$$

where *num* is the sum of people in the states of *immune*, *free − rider* and *infected*. The individual payoffs of each state can refer to Equation (11).

The total cost of the population is defined as

$$
\text{cost} = num_v \times C_V + num_i \times C_I
\tag{13}
$$

Because C_I is an unknown constant for all individuals, and we only consider relative cost $c = \frac{C_V}{C_I}$; therefore,

$$\text{cost} = num_v \times c + num_i \tag{14}$$

As Figure 6 illustrates, it can be seen that the payoffs are inversely proportional to the relative cost and the cost is proportional to the relative cost. That is, the total payoffs decrease with the increase of the relative cost, while the total cost increases with the increase of the relative cost. Surprisingly, the total payoffs of the population under the reinforcement learning mechanism are always higher than that of the individualistic strategy when the relative cost c is less than 0.7. Moreover, the total cost of the population is also smaller than the individualistic strategy. When the c is higher than 0.7, the total payoffs under the reinforcement learning mechanism are nearly close to that of individualism. Such a result is gratifying. With our reinforcement learning mechanism, the population cost significantly decreases, while the population payoffs distinctly increase. In addition, our reinforcement learning mechanism has improved the overall community vaccine coverage level on the basis of individualism and effectively reduced the final epidemic size. Those results shed light on the necessity of encouraging individuals to make vaccination decisions based on collectivism for public health and government.

(**a**) Changes in crowd payoffs (**b**) Changes in crowd costs
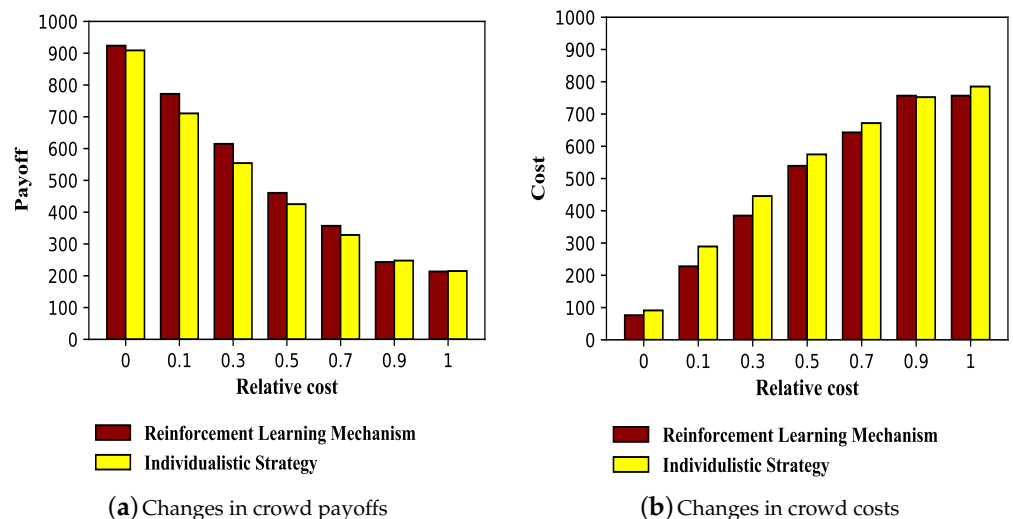
Figure 6. Relationship between the overall payoffs and costs of the population and the relative cost. (**a**) The change of population payoffs with a relative cost. (**b**) The change of population cost with the relative cost. With the increase of the relative cost, the payoffs of the crowd are decreasing, whereas the costs of the crowd are rising. Basically, the total payoffs of the population under the reinforcement learning mechanism are always higher than that of the individualism mechanism, and the total cost of the population is always lower than that of the individualism mechanism.

3.2.2. Dynamic of Long-Term Payoffs and Costs

In the previous section, we know that the vaccination strategy based on the reinforcement learning mechanism can reduce the overall community cost of vaccination. In this section, we will further study the dynamics of population costs and payoffs with the reinforcement learning mechanism. As we mentioned above, each experiment includes 50 seasons. In order to study the dynamic of long-term payoffs, we divide the 50 seasons into two parts: the shocking season and the stationary season. In the case of different c, the dynamic of community vaccine coverage level under the reinforcement learning mechanism is similar. Moreover, there is a process of oscillation and stability. Therefore, we can choose the case of $c = 0.5$ as an example. As Figure 7 shows, we denote the season in which the vaccine coverage oscillates dramatically as the shock season, i.e., season 0 to season 10. Furthermore, the season in which the vaccine coverage converges to a steady state is called the stationary season, i.e., season 11 to season 50.
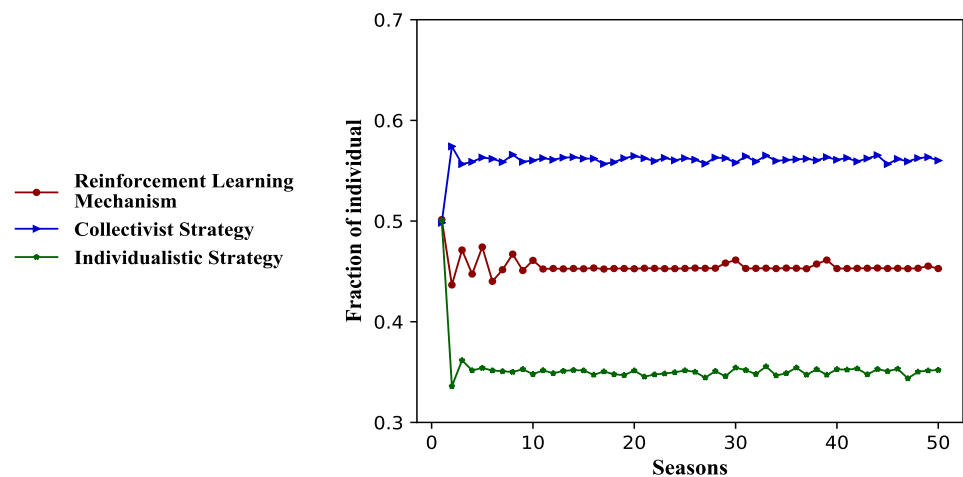
Figure 7. The dynamics of community vaccine coverage under the reinforcement learning mechanism when $c = 0.5$. The level of vaccine coverage fluctuated in the first 10 seasons and then stabilized. The first 10 seasons are called the shocking season, and the following stable seasons are called the stationary season.

As Figure 8 shows, as c gradually increases from 0 to 1, individuals' payoffs in stationary seasons become higher than that in the shock seasons, and their costs become lower. This is because when the relative cost is relatively small, most individuals tend to be vaccinated even in shock seasons. Therefore, most seasons have a relatively high vaccine coverage level and individuals' average payoffs keep higher. On the contrary, when the relative cost is relatively high, most shock seasons are at the low level of vaccine coverage, which makes individuals' payoffs are a bit low.

(a) Dynamic of payoffs
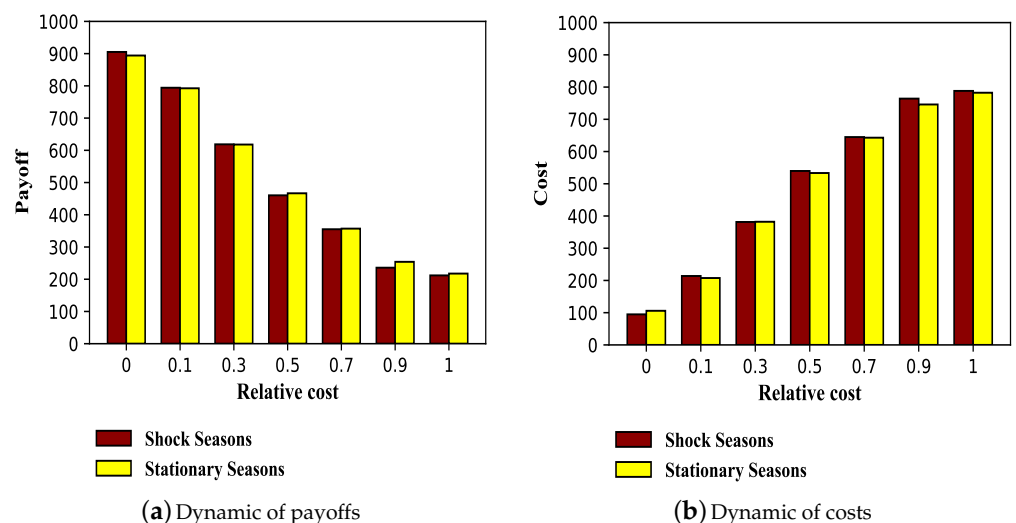
(b) Dynamic of costs

Figure 8. The dynamic of payoffs and costs. As c gradually increases from 0 to 1, the payoffs in the stationary states become gradually higher than the payoffs in the shock seasons, and the corresponding individuals' costs become gradually lower.

In addition, as c increases from 0 to 1, the payoffs in the stationary states become gradually higher than the payoffs in the shock seasons, and the corresponding individuals' costs become gradually lower. Specifically, when $c = 0$, communities reach the state of herd immunity more quickly, and result in the existence of lots of free-rider individuals. Consequently, they get a higher payoff at shock seasons. However, when $c \geq 0.5$, the average payoffs in stationary seasons are significantly higher. Therefore, we can draw the conclusion that in the case of small c, the reinforcement learning mechanism has little effect

on increasing individuals' long-term payoff. However, as *c* increases, the reinforcement learning mechanism can significantly enable individuals to obtain high payoffs in stationary seasons, and produce long-term high payoffs.

3.3. Effectiveness of Collectivist Strategy

From the perspective of public health, we expect more and more persons can make vaccination decisions with collectivist strategy, as it is verified that with the increase of *c*, the "free riders" phenomenon leads to a very low level of vaccine coverage in the population with the individualistic strategy. To systematically study the impact of collectivist strategy, we manually adjust the weight of individualism and collectivism in a mixed strategy with a parameter of α ($0 \leq \alpha \leq 1$). The individual's vaccination probability is formulated as

$$P_i(n) = \alpha P_i^S(n) + (1 - \alpha) P_i^A \tag{15}$$

where α denotes the weight of individualistic strategy.

As we can see from Figure 9a, the strategy with a small value α always gets a higher level of vaccine coverage no matter which value the relative cost *c* has. Moreover, when the individual's strategy is closer to a pure individualistic strategy, i.e., $\alpha = 1$, the vaccine coverage drops to the lowest, comparing that with the same relative cost. In addition, it is worthy of noting that when c = 0, the colors in the range of $0 < \alpha < 0.7$ are very close as shown in Figure 9b, which means the epidemic sizes are almost the same. This is because the vaccine coverage is basically above 0.6 at that time and reaches the critical coverage level which can prevent disease propagation among populations [47].
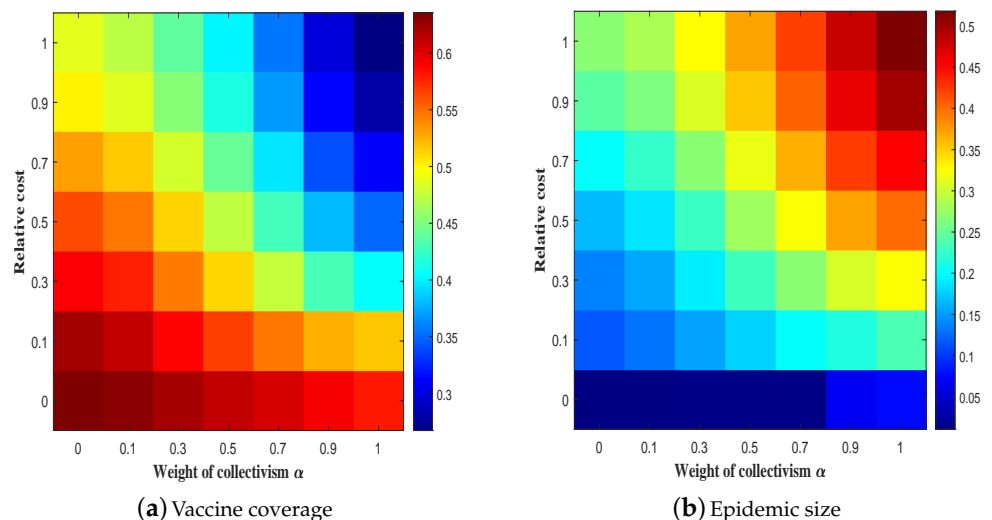

Figure 9. The relationship between vaccine coverage and epidemic size with relative cost *c* and the weight α. (**a**) The change of vaccine coverage. (**b**) The change of epidemic size. When *c* is fixed, vaccine coverage is inversely proportional to α while the epidemic size is proportional to α. Meanwhile, we can also find that vaccine coverage is inversely proportional to *c* while the epidemic size is proportional to *c* when α is fixed.

3.4. Effectiveness of Vaccination Mechanism with Respect to Individuals' Degree

Until now, we have testified that people following the collectivist strategy can get a higher vaccine coverage level and a lower epidemic size. However, as vaccination is mostly voluntary, we cannot force or incentive all individuals to act with collectivism. In view of this, we further study the effectiveness of vaccination mechanisms with respect to their degrees (the number of their associations) in this section, in order to find those with optimal performance. We take the 10% nodes with the highest degree and the other 10% nodes with the lowest degree for example.

As Figure 10 shows, the proportion of picking collectivistic strategy is much higher when individuals with low degrees, comparing that with high degrees. With the increase of relative cost *c*, the fraction of individuals picking the collectivistic strategy declines to 20% more or less among those with low degrees. However, the fraction remains around 20% among individuals with high degrees, regardless of relative cost *c*. Individuals with low degrees have fewer neighbors. When one person takes the vaccine, the proportion of vaccinated individuals in their local mixed groups would increase largely, i.e., the vaccination coverage level increase greatly. On the contrary, individuals with high degrees have more neighbors, and the action of taking vaccine affect little on vaccine coverage in their groups. In reality, most people have relatively few associations. Therefore, our findings testify for public health the necessity and significance of motivating individuals to vaccinate with the collectivist strategy as much as possible.

(a) Low degree individuals   **(b)** High degree individuals
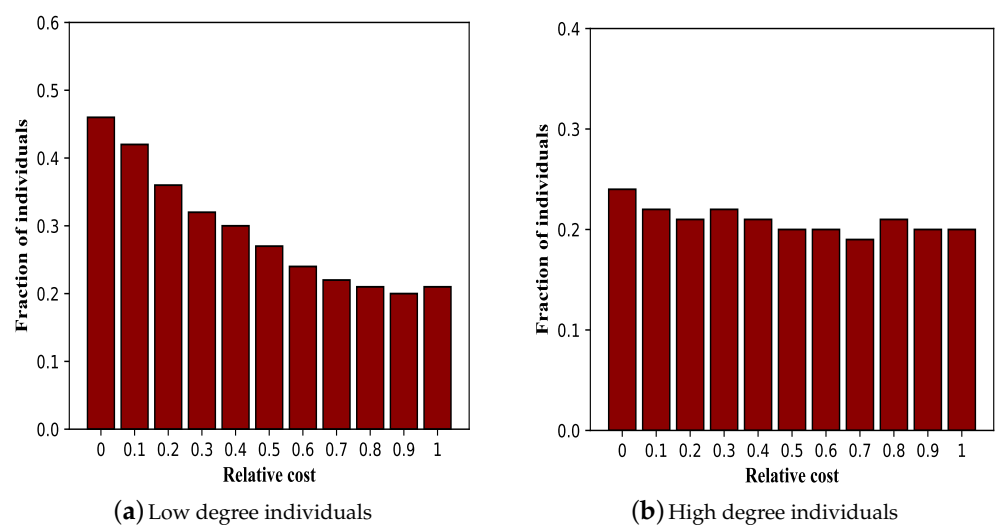
Figure 10. The nodes are sorted in descending order of degree. The top 100 nodes are the top 10%, and the bottom 100 nodes are the last 10%. With the increase of the relative cost, the collectivist strategy of individuals with low degree is declining. However, the proportion of collectivist strategy of individuals with high degree has been kept at around 20%.

4. Discussion

Previous studies on altruism are mainly based on questionnaire surveys, network data collection, and game theory. They analyze individuals' decision-making processes and provide strong theoretical support for the promotion of altruism. However, from a practical point of view, the conclusions drawn by such methods cannot directly provide a clear guide to the implementation of realistic vaccination strategies. That is because they testify that altruism can have a positive impact, but some issues, such as how it affects and how much impact it can have, remain unsolved. Considering this, we quantitatively formulate the influence of altruism, with an individuals' intelligent decision-making mechanism based on reinforcement. Our research results further testify the effects of altruism in terms of vaccine coverage level and epidemic size. In the future, we can investigate the quantitative effects of altruism with different network structures and different diseases, which has specific practical significance for the implementation of vaccination strategies in real society.

Meanwhile, we believe that cultural factors may also be very important influencing factors for studying altruism. From a realistic perspective, individuals in society often exist in multiple small communities. The "local-mixed groups" we proposed corresponded to this reality. We believe that in such a small community individuals can pay more attention to the impact of their vaccination decisions on their associations. We believe that cultural factors related to collectivism may also help individuals make better decisions. For example,

the popularity of the "family" culture drives individuals to make vaccination decisions that benefit other family members.

In this article, we roughly investigate disease propagation with the SIR model. However, the epidemiological characteristics and dynamics of infectious diseases may also affect individual vaccination decisions. Considering the diversity of infectious diseases, we may conduct more studies on different epidemiological models, such as SI, SEIR, etc., in future work.

5. Conclusions

In this paper, we have focused on investigating human voluntary vaccination through making decisions based on the perspectives of individuals themselves or individuals' local-mixed groups. Accordingly, we have presented two different strategies, i.e., (i) the selfish individualistic strategy allows individuals to take the vaccine or not only based on their perceived vaccination and infected cost/payoff, and (ii) the altruistic collectivist strategy allows individuals to act based on the perceived cost/payoff of individuals and their groups. Moreover, we propose a voluntary vaccination mechanism based on reinforcement learning. Our mechanism drives individuals adaptively picking one of these two strategies to obtain optimal vaccination payoff, and finally achieves a higher vaccine coverage and a lower epidemic size.

Our simulations on scale-free networks show that our mechanism effectively promotes the vaccine coverage level. More importantly, with such a mechanism, communities can always get higher total payoffs with fewer costs paid, compared to that of the pure individualistic strategy. Based on our numerical experiments, we can find such performance mostly stems from individuals' adaptively picking collectivist strategy.

Our findings suggest that the collectivist strategy is more effective for improving vaccine coverage level during an epidemic. Public health authorities should pay more attention to encourage individuals to make vaccination decisions from the perspective of their local mixed groups. Especially, note that individuals with low degrees are more significant as their vaccination behaviors can more sharply improve vaccination coverage of their groups and greatly reduce epidemic size.

Author Contributions: Conceptualization, H.Q. and B.S.; methodology, H.Q., B.S. and C.W.; formal analysis, H.Q. and C.W.; investigation H.Q. and C.W.; writing—original draft preparation, H.Q. and C.W.; writing—review and editing, H.Q., C.W., T.Q. B.S., and Q.B.; visualization, H.Q. and C.W. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The simulation data used to support the findings of this study are available from the corresponding author upon request.

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

References

1. Marra, M.A.; Jones, S.J.M.; Astell, C.R.; Holt, R.A.; Brooks-Wilson, A. The genome sequence of the SARS-associated coronavirus. *Science* **2003**, *300*, 1399. [CrossRef] [PubMed]
2. The Chinese SARS Molecular Epidemiology Consortium. Molecular evolution of the SARS Coronavirus, during the Course of the SARS epidemic in China. *Science* **2004**, *303*, 1666–1669. [CrossRef] [PubMed]
3. Fouchier, R.A.M.; Kuiken, T.; Schutten, M.; Amerongen, G.V.; Doornum, G.J.J.V.; Hoogen, B.G.V.D.; Peiris, M.; Lim, W.; Stoehr, K.; Osterhaus, A.D.M.E. Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* **2003**, *423*, 240. [CrossRef]
4. Small, M.; Walker, D.M.; Tse, C.K. Scale-Free distribution of Avian influenza outbreaks. *J. Math. Biol. Nat.* **2006**, *53*, 253–272. [CrossRef] [PubMed]
5. Fatimah, S.D.; Seema, J.; Lyn, F.; Michael, W.S.; Novel Swine-Origin Influenza A HN Virus Investigation Team. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* **2009**, *360*, 2605–2615.

6.   Peiris, J.S.M.; Poon, L.L.M.; Guan, Y. Emergence of a novel swine-origin influenza A virus (S-OIV) H1N1 virus in humans. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* **2009**, *45*, 169–173. [CrossRef]
7.   Alamelu, R. Immunology of tuberculosis. *Med. Clin. N. Am.* **1993**, *77*, 1235–1251.
8.   Christopher, D.; Blanc, L. Global tuberculosis control: Surveillance, planning, financing. *Wkly. Epidemiol. Rep.* **2003**, *78*, 122–128.
9.   Dye, C.; Scheele, S.; Dolin, P.; Pathania, V.; Raviglione, M.C. Estimated incidence, prevalence, and mortality by country. *JAMA* **2003**, *282*, 677. [CrossRef] [PubMed]
10.   Rothan, H.A.; Byrareddy, S.N. The epidemiology and pathogenesis of Coronavirus disease (COVID-19) outbreak. *J. Autoimmun.* **2020**, *109*, 102433. [CrossRef]
11.   Velavan, T.P.; Meyer, C.G. The COVID-19 epidemic. *Trop. Med. Int. Health* **2020**, *25*, 278. [CrossRef]
12.   Liu, Y.; Gayle, A.A.; Wilder-Smith, A.; Röcklv, J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Travel Med.* **2020**, *2020*. [CrossRef]
13.   Yin, H.; Wang, S.; Zhu, Y.; Zhang, R.; Ye, X.; Wei, J.; Hou, P.C. The development of critical care medicine in China: From SARS to COVID-19 pandemic. *Crit. Care Res. Pract.* **1999**, *282*, 677. [CrossRef] [PubMed]
14.   Paolo, B. Demographic impact of vaccination: A review. *Vaccine* **1999**, *17*, S120–S125.
15.   Elwood, J.M. Smallpox and its eradication. *J. Epidemiol. Commun. Health* **1988**, *43*, 92. [CrossRef]
16.   Fine, P.E.M.; Jacqueline, A.C. Individual versus public priorities in the determination of optimal vaccination policies. *Am. J. Epidemiol.* **1987**, *124*, 1012–1020. [CrossRef]
17.   Kristin, L.N.; April, L.; Karen, L.M.; Maureen, M.; Meri, H.; Sanne, M.; Mari, D. The effectiveness of vaccination against influenza in healthy, working adults. *Phys. Life Rev.* **1995**, *333*, 889–893.
18.   Klaus, D. Infectious diseases of humans: Dynamics and control. *Parasitol. Today* **1992**, *8*, 179.
19.   Chen, F.H. A susceptible-infected epidemic model with voluntary vaccinations. *J. Math. Biol.* **2006**, *53*, 253–272. [CrossRef]
20.   Viboud, C.; Boёlle, P.-Y.; Fabrice, C.; Alain-Jacques, V.; Antoine, F. Prediction of the spread of influenza epidemics by the method of analogues. *Am. J. Epidemiol.* **2003**, *158*, 996–1006. [CrossRef]
21.   Liu, G.; Qiu, H.; Shi, B.; Zhen, W. Imitation and memory-based self-organizing behaviors under voluntary vaccination. In Proceedings of the International Conference on Security, Pattern Analysis, and Cybernetics, Shenzhen, China, 15–17 December 2017; pp. 491–496.
22.   Xia, S.; Liu, J. A belief-based model for characterizing the spread of awareness and its impacts on individuals' vaccination decisions. *J. R. Soc. Interface* **2014**, *11*, 20140013. [CrossRef]
23.   Kaufmann, S.H.E.; Mcmichael, A.J. Annulling a dangerous liaison: Vaccination strategies against AIDS and tuberculosis. *Nat. Med.* **2005**, *11*, S33. [CrossRef]
24.   Keeling, M.J.; Woolhouse, M.E.J.; Davies, R.M.M.G.; Grenfell, B.T. Modelling vaccination strategies against foot-and-mouth disease. *Nature* **2003**, *421*, 136–142. [CrossRef] [PubMed]
25.   Tildesley, M.J.; Savill, N.J.; Shaw, D.J.; Deardon, R.; Brooks, S.P.; Woolhouse, M.E.J.; Grenfell, B.T.; Keeling, M.J. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature* **2006**, *404*, 83–86. [CrossRef]
26.   Bauch, C.T.; Earn, D.J. Vaccination and the theory of games. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13391–13394. [CrossRef]
27.   Shi, B.; Liu, G.; Qiu, H.; Wang, Z.; Ren, Y.; Chen, D. Exploring voluntary vaccination with bounded rationality through reinforcement learning. *Phys. Stat. Mech. Appl.* **2019**, *515*, 171–182. [CrossRef]
28.   Bauch, C.T.; Galvani, A.P.; Earn, D.J. Group interest versus self-interest in smallpox vaccination policy. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10564–10567. [CrossRef] [PubMed]
29.   Hershey, J.C.; Asch, D.A.; Thumasathit, T.; Meszaros, J.; Waters, V.V. The roles of altruism, free riding, and bandwagoning in vaccination decisions. *Organ. Behav. Hum. Decis. Process.* **1994**, *59*, 177–187. [CrossRef]
30.   Zhang, Q.; Buckling, A.; Ellis, R.J.; Godfray, H.C.J. Coevolution between cooperators and cheats in a microbial system. *Evolution* **2009**, *63*, 2248–2256. [CrossRef]
31.   Nowak, M.A. Five rules for the evolution of cooperation. *Science* **2007**, *314*, 1560–1563. [CrossRef]
32.   Hamilton, W.D. The genetical evolution of social behaviour. *J. Theor. Biol.* **1964**, *7*, 17–52. [CrossRef]
33.   Foster, K.R.; Wenseleers, T.; Ratnieks, F.L.W. Kin selection is the key to altruism. *Trends Ecol. Evol.* **2006**, *21*, 57–60. [CrossRef] [PubMed]
34.   Gardner, A.; West, S.A.; Wild, G. The genetical theory of kin selection. *J. Evol. Biol.* **2011**, *24*, 1020–1043. [CrossRef] [PubMed]
35.   Queller, D.C.; Strassmann, J.E. Kin selection and social insects. *Bioscience* **1998**, *48*, 165–175. [CrossRef]
36.   Ohtsuki, H.; Iwasa, Y. The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* **2006**, *239*, 435–444. [CrossRef]
37.   Ohtsuki, H.; Iwasa, Y. How should we define goodness? Reputation dynamics in indirect reciprocity. *J. Theor. Biol.* **2004**, *231*, 107–120. [CrossRef]
38.   Cucciniello, M.; Pin, P.; Imre, B.; Porumbescu, G.; Melegaro, A. Altruism and Vaccination Intentions: Evidence from Behavioral Experiments. *medRxiv* **2020**. [CrossRef]
39.   Rieger, M.O. Triggering Altruism Increases the Willingness to Get Vaccinated against COVID-19. *Soc. Health Behav.* **2020**, *3*, 78. [CrossRef]
40.   Shi, B.; Qiu, H.; Niu, W.; Ren, Y.; Ding, H.; Chen, D. Voluntary vaccination through self-organizing behaviors on locally-mixed social networks. *Sci. Rep.* **2017**, *7*, 2665. [CrossRef]

41.   Liu, G.; Peng, S.; Qiu, H.; Shi, B.; Chen, Y. Perceiving epidemic severity in social network. *Complexity* **2019**, *2019*, 13.
42.   Jalil, R.; Mehri, S.; Jorge, D.; Cristina, J.; Nuno, M. On the dynamical complexity of a seasonally forced discrete SIR epidemic model with a constant vaccination strategy. *Complexity* **2018**, *2018*. [CrossRef]
43.   Feng, F.; Daniel, I.R.; Wang, L.; Martin, A.N. Imitation dynamics of vaccination behaviour on social networks. *Proc. Biol. Sci.* **2011**, *278*, 42–49.
44.   Arne, T.; Dirk, S.; Sommerfeld, R.D.; Hans-Jrgen, K.; Manfred, M. Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2962–2966.
45.   Sutton, R.S.; Mcallester, D.; Singh, S.; Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. *Neural Inf. Process. Syst.* **2000**, *12*, 1057–1063.
46.   Zhang, H.; Shu, P.; Wang, Z.; Tang, M. Preferential imitation can invalidate targeted subsidy policies on seasonal-influenza diseases. *Appl. Math. Comput.* **2017**, *294*, 332–342. [CrossRef]
47.   Anderson, R.; May, R. *Infectious Diseases of Humans*; Oxford University Press: Oxford, UK, 1992.