

# The Evolution of Strategic Timing in Collective-Risk Dilemmas

Christian Hilbe\*, Maria Abou Chakra, Philipp M. Altrock<sup>‡</sup>, Arne Traulsen

Evolutionary Theory Group, Max-Planck Institute for Evolutionary Biology, Plön, Germany

## Abstract

In collective-risk dilemmas, a group needs to collaborate over time to avoid a catastrophic event. This gives rise to a coordination game with many equilibria, including equilibria where no one contributes, and thus no measures against the catastrophe are taken. In this game, the timing of contributions becomes a strategic variable that allows individuals to interact and influence one another. Herein, we use evolutionary game theory to study the impact of strategic timing on equilibrium selection. Depending on the risk of catastrophe, we identify three characteristic regimes. For low risks, defection is the only equilibrium, whereas high risks promote equilibria with sufficient contributions. Intermediate risks pose the biggest challenge for cooperation. In this risk regime, the option to interact over time is critical; if individuals can contribute over several rounds, then the group has a higher chance to succeed, and the expected welfare increases. This positive effect of timing is of particular importance in larger groups, where successful coordination becomes increasingly difficult.

**Citation:** Hilbe C, Abou Chakra M, Altrock PM, Traulsen A (2013) The Evolution of Strategic Timing in Collective-Risk Dilemmas. PLoS ONE 8(6): e66490. doi:10.1371/journal.pone.0066490

**Editor:** James A. R. Marshall, University of Sheffield, United Kingdom

**Received:** March 18, 2013; **Accepted:** May 7, 2013; **Published:** June 14, 2013

**Copyright:** © 2013 Hilbe et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors have no support or funding to report.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: hilbe@evolbio.mpg.de

<sup>‡</sup> Current address: Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, United States of America

## Introduction

In joint efforts, coordination problems often arise because some individuals may question the chances of success, or the intentions of the others. In some examples, such as the prevention of climate change [1–3] or the management of global economic crises [4], a failure to coordinate on a beneficial equilibrium can endanger the whole group, or implies considerable welfare losses. In all these cases, subjects may try to alleviate the risks of collective action by sending trust-building signals over a period of time. The importance of time as a coordination device that allows individuals to interact and influence one another, has already been noted by Schelling [5]: In the context of multinational conflicts, he argues that “If each party agrees to send a million dollars to the Red Cross on condition the other does, each may be tempted to cheat if the other contributes first, and each one’s anticipation of the other’s cheating will inhibit agreement. But if the contribution is divided into consecutive small contributions, each can try the other’s good faith for a small price.” Thus, the strategic use of time may help to overcome coordination problems that would be hard to settle otherwise.

In order to explore the propensity for such strategic behaviors in humans, Milinski et. al. [6] conducted behavioral experiments for a particular coordination problem, the collective-risk dilemma. In these experiments, each subject was endowed with a fund and then asked, in each of ten consecutive rounds, to donate from this endowment into a common pool. If the group’s total contributions after ten rounds reached or surpassed a certain target amount, all group members acquired their individual withheld funds. Otherwise, if the group failed to reach the target, they lost everything

with a certain risk probability. In the experiments, a substantial fraction of groups failed to coordinate on a beneficial equilibrium with sufficient contributions, even if the risk of losing everything was as high as 90%. An analysis of the subjects’ behavior in these high-risk treatments revealed that there was a significant tendency to procrastinate contributions towards the second half of the game [6]. Such a delay of contributions could be an indicator of individual attempts to free-ride, exploiting the contributions of others. On the other hand, a temptation to wait may also arise if fearful subjects aim to avoid wasted contributions [7].

These observed temporal patterns thus call for a closer examination. However, most previous theoretical investigations for the collective-risk dilemma have neglected the impact of timing on coordination behavior [8–13]. These studies considered a one round game and assumed that individuals do not react to the contributions of their co-players over the course of the game. This means that, effectively, timing and thereby strategic behaviors were neglected. An exception is [14], which explicitly followed the setup of the experiments and considered a game with ten rounds. In computer simulations, it was observed that successful players delayed their contributions towards the later stage of the game. However, the focus was on the observable behaviors of the subjects, rather than on the underlying strategies. Moreover, as the game length was fixed to ten rounds, the impact of the duration of the game on cooperation was not analyzed. Herein, we thus add to the previous literature by systematically exploring how time and timing can promote successful coordination.

The impact of time on coordination behavior is probably best explored in the context of the volunteer’s dilemma, where a collective good is produced only if there is a volunteer who

provides it [15,16]. For this game, Weesie [17] has found that the inclusion of time greatly enhances coordination and increases the individual probability to volunteer. Moreover, in the asymmetric case where players have different costs of volunteering, time helps to select the optimal volunteer as the player with the lowest costs volunteers without delay [17]. However, there is a subtle difference between coordination in the volunteer’s dilemma and in the collective-risk dilemma: even in the symmetric volunteer’s dilemma, where players are *ex ante* indistinguishable, the *ex post* payoffs are typically asymmetric – as it takes only one player to take on the burden of volunteering. In contrast, the collective-risk dilemma allows for pure symmetric equilibria, where all players contribute equally to reach the target. Thus, in the collective-risk dilemma the question is not which of the players gives in first, but rather when and to which extent each player contributes.

To address these questions we employ evolutionary game theory [18–21]. This allows us to study the dynamics of contributions without presuming that individuals are fully rational (or that they are aware of their co-players’ rationality), as for example in [22]. In the following, we thus develop an evolutionary model to show that time has a two-fold effect in the collective-risk dilemma: on the one hand, it facilitates coordination, but on the other hand it leads subjects to delay their contributions as long as possible.

**Model**

We consider a collective-risk dilemma played among  $M$  individuals. In each of the  $R$  rounds, the players have to decide individually how much of their initial endowment  $E$  they want to contribute into a common pool. As in the experiments of Milinski et. al. [6], we assume that an individual is limited to a maximum contribution of  $E/R$  per round, such that a player contributing the maximum amount in each round expends the full endowment  $E$ . If the group collectively succeeds in investing a target sum  $T \geq E$  by the end of the game, then each player  $i$  keeps the retained portion of the endowment  $E - C_i$ , where  $C_i$  denotes the player’s total contributions over the  $R$  rounds. However, if they collectively fail to reach the target, then all the players lose everything with some exogenous probability  $p$ . Thus, player  $i$  obtains an average payoff of  $E - C_i$  when the target is reached and  $(1 - p)(E - C_i)$  when the target is missed. Overall, the individuals in such a game face a social dilemma: while everyone benefits from reaching the target, players are tempted to suppress their individual contributions.

We model the strategies in such a collective-risk dilemma as contingent rules: when deciding how much to contribute in a given round, players take into account how much their co-players contributed previously. This allows individuals to apply strategies such as Schelling’s rule and to contribute an equal amount of  $T/(MR)$  in each round, provided that their peers do the same. Inconveniently, as the number of rounds or the number of players increases, the possible number of contingent strategies increases exponentially. Moreover, collective risk dilemmas have a large set of Nash equilibria: any state in which the group members retain their endowment, or in which they meet the target exactly such that individual contributions do not exceed the expected loss upon failure,  $C_i \leq p \cdot E$ , constitutes an equilibrium. To see this, we first note that when the target is exactly met there is no benefit of a further increase of contributions. On the other hand if a player unilaterally decides to cut down her contributions, then the target is missed and the player’s payoff is at most  $(1 - p)E$ , which is below the coordination payoff  $E - C_i$  if the risk  $p$  is sufficiently high,  $p > C_i/E$ . Thus, any outcome where the target is exactly met

and where no individual contributes more than  $pE$  constitutes a Nash equilibrium, even if the costs are distributed unfairly.

To cope with the complexity due to the large number of strategies and possible equilibria, we will study such large-scale collective-risk dilemmas by performing extensive individual-based simulations. However, to provide a basic intuition, we first investigate the role of conditional strategies and timing in a simplified collective-risk dilemma between two players.

**Analysis of a simplified collective-risk dilemma between two players**

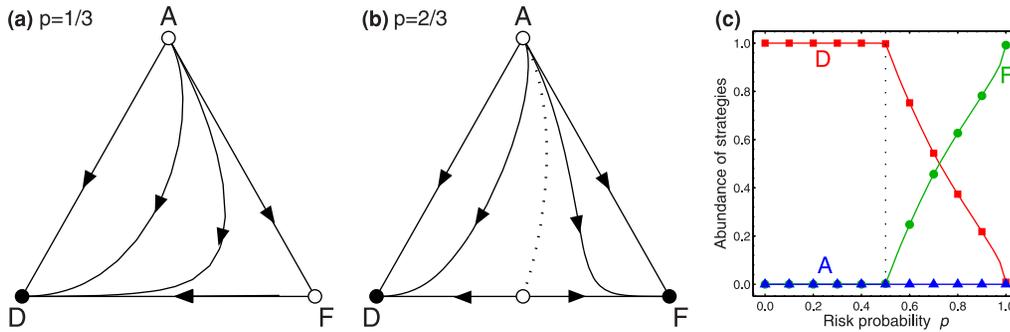
To illustrate the importance of time as a coordination device in collective-risk dilemmas, let us first explore the baseline case where players are not able to interact over multiple rounds. In such a case, only unconditional strategies are available, such as being a defector (who does not contribute,  $C_D = 0$ ), a fair-sharer (someone who contributes a proportional share of the target,  $C_F = T/M$ ), or an altruist (contributing the full endowment,  $C_A = E$ ). In the simplest case of a pairwise game where the target is equal to one player’s endowment,  $T = E$ , a collective-risk dilemma with these three strategies is represented by the payoff matrix:

	$D$	$F$	$A$	
$D$	$(1 - p)E$	$(1 - p)E$	$E$	(1)
$F$	$(1 - p)E/2$	$E/2$	$E/2$	
$A$	$0$	$0$	$0$	

In this game, irrespective of the strategy of their co-player, altruists end up with a payoff of zero. Therefore, altruism is a dominated strategy and we may expect that altruistic acts occur at very low frequencies, reducing the collective-risk dilemma to a game between defectors and fair sharers. Individuals strictly prefer defection for all  $p < 1/2$ , as the expected loss for missing the target  $E \cdot p$  is below the fair share contribution  $E/2$ . This prediction is confirmed by replicator dynamics [19,23], see Figure 1a: irrespective of the initial distribution of strategies in the population, individuals learn to stop contributing. This qualitative behavior changes as the risk of collective loss exceeds  $1/2$ . In this case, there are three possible Nash equilibria: all players withholding their contributions, all individuals doing their fair share, and a mixed population of defectors and fair-sharers. In this mixed equilibrium, the fraction of defectors is given by

$$x_D^* = \frac{2p - 1}{p} \tag{2}$$

However, since the mixed equilibrium is not evolutionary stable (Figure 1b), the dynamics either leads to a homogeneous population of defectors or to a homogeneous population of fair sharers. Which of these two possible outcomes is reached, depends on the initial behavior of the individuals: populations with a sufficient initial number of fair sharers eventually succeed in coordinating on the beneficial fair-share equilibrium, whereas populations mostly consisting of defectors end up in the detrimental equilibrium. In general, it depends on the risk of collective loss, whether or not a given initial population succeeds to coordinate on the fair-share equilibrium. To estimate the basins of attraction of each equilibrium, we have recorded the results of the evolutionary dynamics for different randomly chosen initial



**Figure 1. Replicator dynamics of the simplified collective-risk dilemma without timing.** For three strategies, the state space takes the form of a triangle, the simplex  $S_3$ . The corners of this triangle correspond to homogeneous populations, where all individuals use the same strategy, whereas points in the interior correspond to mixed populations. (a) If the risk probability  $p < 1/2$ , then all interior initial populations eventually converge to a population of defectors. (b) For  $p > 1/2$  a bistable situation emerges: If the initial frequency of fair sharers is sufficiently high, then the subjects learn to coordinate on the beneficial fair share equilibrium. (c) The fraction of initial populations that converge towards the fair share equilibrium increases with  $p$ , reaching 100% in the limit of full risk. For this graph, we have simulated the replicator dynamics for 20,000 randomly chosen initial populations.  
doi:10.1371/journal.pone.0066490.g001

populations (Figure 1c). According to these simulations, an increasing risk of collective loss stimulates attempts to reach the fair-share equilibrium. Nevertheless, even for high risk values, a substantial proportion of initial populations fails to coordinate on the beneficial equilibrium. For instance, even for  $p = 90\%$ , roughly a quarter of all initial states lead to a non-cooperative population of defectors. Increasing the risk of collective loss has therefore a two-fold effect on the achieved welfare: on the one hand, a higher  $p$  decreases the expected payoff if the target is missed, but on the other hand, high values of  $p$  make it more likely that the players cooperate. As a consequence, intermediate, and not high values of  $p$  represent the worst-case scenario for the average payoffs (Figure 1c).

To investigate the impact of time, let us now consider a collective-risk dilemma with two rounds. Again, we assume that each agent has an initial endowment  $E$ , and that each agent can either contribute 0 or  $E/2$  to the common pool in each round. The target is reached if total contributions sum up to a player's endowment, that is  $T = E$ . Obviously, this setting allows more than the previous three strategies of defectors, fair sharers, and altruists, as in the two-round game players may condition their behavior in the second round on their co-player's contribution in the first round. We can write the players' strategies as a 3-tuple  $(i; j, k)$  with  $i, j, k \in \{0, 1\}$ . The first variable  $i$  determines whether the player cooperates in the first round: If  $i = 1$ , then this player contributes  $E/2$  to the common pool, whereas for  $i = 0$ , the player does not contribute. The second variable  $j$  determines whether the focal player cooperates in the second round, given that the opponent cooperated in the first round, whereas the third variable  $k$  corresponds to the focal player's action in the second round, given that the opponent did not cooperate in the first round.

Therefore, this pairwise collective-risk dilemma allows eight possible strategies, which include the previous three strategies of the game without timing: For example, the strategy  $(0; 0, 0)$  corresponds to a defector who does not contribute to the common pool, whereas players with strategy  $(1; 1, 1)$  are altruists who contribute their full endowment, independent of the opponent's contribution behavior. The two strategies  $(1; 0, 0)$  and  $(0; 1, 1)$  are fair sharers, unconditionally contributing half of their endowment, either in the first period or in the second period, respectively. However, the collective-risk dilemma with timing allows additional strategies of interest: For instance, one may interpret a player with strategy  $(0; 1, 0)$  as a conditional cooperator, who is cooperative in

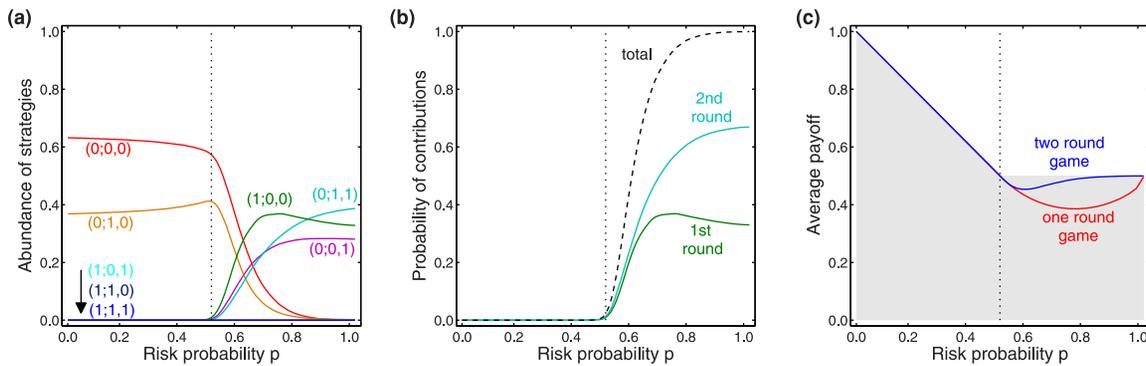
the second round, given that the co-player was cooperative in the first round. In contrast, a player using  $(0; 0, 1)$  applies a wait & see strategy, by awaiting the first round and by cooperating in the second round if there were no contributions in the first round. We can summarize the eight possible strategies' payoffs in a matrix:

	(0; 0, 0)	(0; 0, 1)	(0; 1, 0)	(0; 1, 1)	(1; 0, 0)	(1; 0, 1)	(1; 1, 0)	(1; 1, 1)
(0; 0, 0)	$(1-p)E$	$(1-p)E$	$(1-p)E$	$(1-p)E$	$(1-p)E$	$E$	$(1-p)E$	$E$
(0; 0, 1)	$(1-p)E/2$	$E/2$	$(1-p)E/2$	$(1-p)E/2$	$(1-p)E$	$E$	$(1-p)E$	$E$
(0; 1, 0)	$(1-p)E$	$(1-p)E$	$(1-p)E$	$(1-p)E$	$E/2$	$E/2$	$E/2$	$E/2$
(0; 1, 1)	$(1-p)E/2$	$E/2$	$(1-p)E/2$	$E/2$	$E/2$	$E/2$	$E/2$	$E/2$
(1; 0, 0)	$(1-p)E/2$	$(1-p)E/2$	$E/2$	$E/2$	$E/2$	$E/2$	$E/2$	$E/2$
(1; 0, 1)	0	0	0	0	$E/2$	$E/2$	$E/2$	$E/2$
(1; 1, 0)	$(1-p)E/2$	$(1-p)E/2$	$E/2$	$E/2$	0	0	0	0
(1; 1, 1)	0	0	0	0	0	0	0	0

(3)

Similar to the case without timing, the defector's strategy  $(0; 0, 0)$  leads to the highest payoff for all risk values  $p < 1/2$ . As the risk of collective loss exceeds  $1/2$ , there are three additional pure Nash equilibria, two fair-share strategies  $(1; 0, 0)$  and  $(0; 1, 1)$  and the wait & see strategy  $(0; 0, 1)$ . To estimate the robustness of these equilibria, we have again performed simulations with randomly chosen initial populations (see Figure 2a). As expected, defection is the most abundant strategy for low values of  $p$ . However, for  $p > 1/2$ , the three cooperative equilibrium strategies  $(1; 0, 0)$ ,  $(0; 1, 1)$  and  $(0; 0, 1)$  are soon applied by a substantial share of initial populations, leading to complete coordination on a beneficial equilibrium with sufficient contributions as  $p$  approaches one. Remarkably, for a risk of collective loss of  $p = 80\%$ , more than 95% of all initial populations learn to coordinate on an equilibrium with sufficient contributions in this game with timing, while only 63% reach the target in the game without timing. The opportunity to interact and influence one another thus indeed proves as a powerful means to reach cooperation in the collective-risk dilemma.

To analyze the timing of contributions, we recorded the fraction of players who contribute in the first and in the second round, respectively (Figure 2b). Depending on the risk of collective loss, one can roughly distinguish three different parameter regions: for  $p < 1/2$ , contributions are neither made in the first nor in the second round, whereas in the interval  $1/2 < p < 2/3$ , the fraction of contributions increases considerably in both rounds. In this



**Figure 2. Replicator dynamics of the collective-risk dilemma with timing.** a) Monte-Carlo simulations for 20,000 randomly chosen initial populations confirm that individuals are most likely to adopt the defector's strategy if  $p < 1/2$ , whereas subjects tend to use cooperative strategies for higher risk values. (b) An analysis of the timing of contributions reveals that for high risk values, individuals tend to make their contributions in the second rather than in the first round. (c) Average payoffs in the game with timing are above the payoffs in the game without timing (the grey shaded area represents the set of all possible average payoffs). doi:10.1371/journal.pone.0066490.g002

region of intermediate risk, early contributors with strategy (1;0,0) make up the majority of the population (Figure 2a), as they benefit from the presence of conditional cooperators (0;1,0). As  $p$  exceeds approximately  $2/3$ , individuals tend to delay their contributions towards the second round. In the limit of full risk,  $p=1$ , contributions in the second round are twice as likely as early contributions, which is in line with the experimental observation that subjects tend to procrastinate their contributions towards the second half of the game [6].

The positive effect of time on coordination is reflected in the achieved average payoff (Figure 2c); especially for moderate risk values, the two-round game results in substantially higher payoffs than the one-round game. In particular, the minimum payoff increases by more than 17% if individuals have the option to interact over time. Again, this minimum payoff is not attained at maximum risk,  $p=1$ , but rather at an intermediate risk value.

### Collective-risk-dilemmas with multiple players and multiple rounds

Real-world coordination problems often involve a large group of individuals and multiple interactions over time. It is therefore natural to explore these more general cases. However, games with multiple players and multiple rounds are considerably more complex, and the size of the payoff matrix increases exponentially in both variables. To investigate such large-scale collective-risk dilemmas, we have performed extensive individual-based simulations. Simulations were conducted using the same setup as in [14], which allows for a comparison with previous work. This setup is similar to the two-round case: each of the  $M$  individuals has an initial endowment  $E$  and may contribute at maximum  $E/R$  per round to the collective pool, in order to reach the group target  $T$ , which is set to  $T=EM/2$  (i.e., the target is reached if all players give their half endowment). For the multi-round case, however, we assume that individuals base their decisions on the collective pool so far, rather than on co-players' individual decisions. That is, for every round  $r \in \{1, \dots, R\}$  a player defines an individual threshold  $\tau_r \in [0, T]$  on the total contributions up to round  $r$ . A player's strategy is then a set  $(\tau_r; j_r, k_r)$ , such that the player contributes an amount  $j_r \in \{0, E/R\}$  if the common pool satisfies the threshold  $\tau_r$ , whereas the player contributes  $k_r \in \{0, E/R\}$  if the individual threshold is not satisfied.

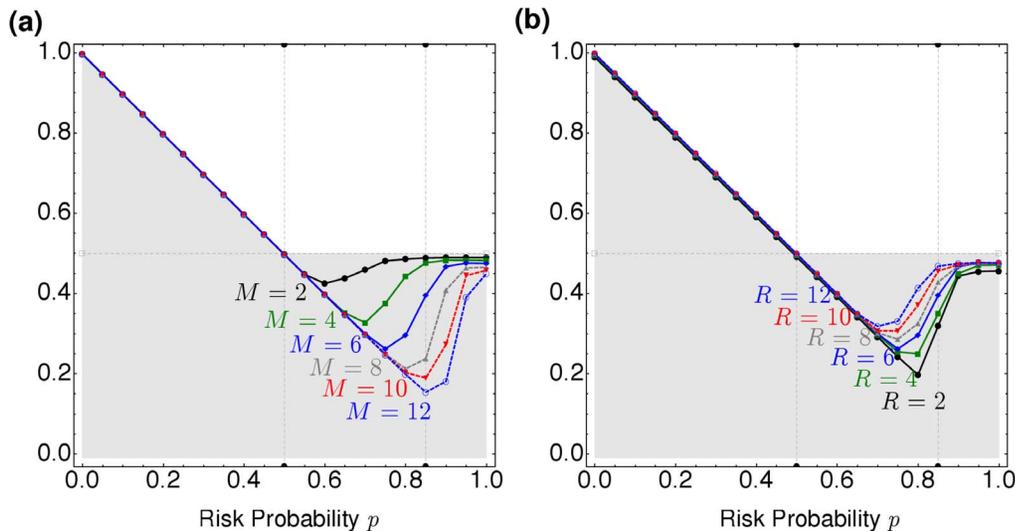
To model the evolutionary dynamics, we use a mutation-selection process in a population of finite size  $N=100$ . In each

generation, individuals participate in several collective-risk dilemmas. Thereafter, the individuals' fitness is calculated as an exponential function of their payoffs,  $f(i) = \exp(\beta \cdot \pi_i)$ , where the strength of selection parameter  $\beta \geq 0$  measures the importance of a player's payoff for its fitness. Individuals are then selected in proportion to their fitness to give rise to the next generation [24,25]. Offspring inherits the strategy of the parent with probability  $1-\mu$ ; with the remaining probability  $\mu$  a player explores a randomly chosen new strategy. In case of such a mutation event, we assume that changes in the thresholds and in the investments of each round occur independently, and that changes in the thresholds are normally distributed around a mean of  $\tau_r$  with a variance of  $\sigma$ . We use this evolutionary game setup to explore the impact of group size,  $M$ , and round number,  $R$ , on coordination in collective-risk dilemmas.

As one may expect, group size has no effect for low risk values,  $p < 1/2$ , where withholding contributions is a weakly dominant strategy (see Figure 3a). However, for  $p > 1/2$ , small groups obtain higher payoffs, due to the higher probability to coordinate on a beneficial equilibrium with sufficient contributions. For larger groups, we observe a diffusion of responsibility [26], and it takes higher risks to motivate players to join the collaborative effort. For example, for a risk of collective loss  $p=80\%$ , groups of 12 players typically fail to reach the target (resulting in a low payoff of approximately  $E/5$ ), whereas two-player groups almost always reach it (leading to the maximum attainable payoff  $E/2$ ). This group effect is quenched when game length increases (see Figure 3b): for  $p > 1/2$ , an increase in the number of rounds leads to a higher probability to coordinate on an equilibrium where the target is reached, resulting in a higher average payoff for all players. This positive influence of time is especially pronounced for intermediate risks, such as  $p=80\%$ . However this quenching did not eliminate the group effect; even if the risk of collective loss approaches one and players have 12 rounds to reach the target, there are still instances of collective failure (resulting in an average payoff below the optimum  $E/2$ ). Thus, while the inclusion of time in general facilitates coordination, there is no guarantee that subjects reach the target.

### Timing of Contributions

The previous simulations also allow us to investigate in more detail how individuals time their contributions in games with multiple rounds. If evolution leads to a contribution scheme

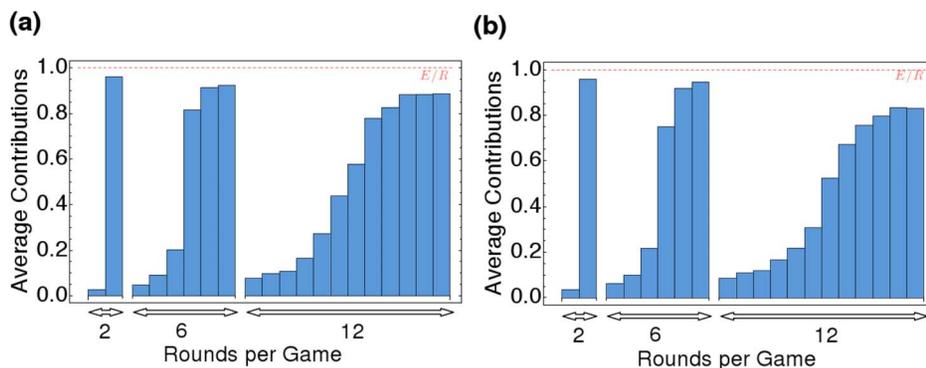


**Figure 3. Simulations of the evolutionary dynamics for the collective-risk dilemma with multiple players and multiple rounds.** Each graph depicts the average payoff for various  $p$ , measured in fractions of the initial endowment (the grey shaded area represents the set of all possible average payoffs). (a) A collective-risk game with  $R=6$  rounds and varying group size, (b) a collective-risk game with  $M=6$  players and varying round number (averages over  $10^5$  generations, number of games per generation  $G=1000$ , mutation rate  $\mu=0.03$ , and the standard deviation for mutations in the thresholds  $\tau_r$  is set to  $\sigma=0.15$ ). doi:10.1371/journal.pone.0066490.g003

comparable to Schelling's rule, then we would expect that individuals signal their willingness to contribute already in early rounds, and they would refrain from further contributions as soon as they realize that their co-players do not follow. If applied by all individuals, such a strategy would lead to overall contributions that are evenly distributed over the  $R$  rounds. Alternatively, individuals that are rational could also apply backward induction: for  $p > 1/2$ , and given that players intend to reach the target, backward induction would suggest that players contribute nothing in the first half of the game, while they would donate the maximum amount  $E/R$  in the second half. In this way, late contributions serve as a self-commitment, which allows individuals to signal credibly that they will not contribute more than their fair share. By not contributing in the beginning, they simply forego any possibility to compensate insufficient contributions, but they ensure that others will either contribute or face collective loss.

Our evolutionary simulations suggest that the timing of contributions is somewhere between these two extremes: while average contributions in the last round are typically close to the maximum amount  $E/R$ , this does not imply that all contributions are shifted towards the second half of the game. Instead, there is always a baseline level of early contributions, independent of the total number of rounds (see Figure 4a). However, the results rather seem to be in line with the backward induction outcome than with evenly distributed contributions over time. This relative abundance of awaiting strategies does not depend on the maximum contribution per round; nor does it depend on the assumption that players only have the binary option of contributing either 0 or  $E/R$  in a given round (see Figure 4b, where subjects could either contribute nothing,  $E/(2R)$  or  $E/R$ ).

A significant delay of contributions towards the second half of the game was also observed in the experiments of Milinski et. al. [6]. However, a comparison of the experimental data for a high

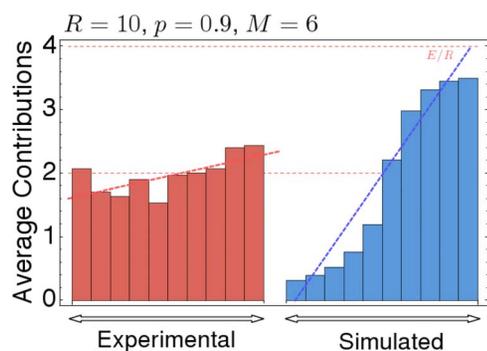


**Figure 4. Timing of contributions in the collective-risk dilemma.** Simulations of the evolutionary dynamics of a collective-risk game with  $M=6$ , and round number  $R$  of 2, 6, and 12 depicting the average contribution per round. We consider two treatments: (a) Possible contributions 0 or  $E/R$ . (b) Possible contributions 0,  $E/(2R)$ , or  $E/R$ . In both treatments we observe delayed contributions, irrespective of the total game length (averages over  $10^5$  generations,  $M=6$ ,  $N=100$ ,  $G=1000$ ,  $\beta=1$ ,  $\mu=0.03$ ,  $\sigma=0.15$ ). doi:10.1371/journal.pone.0066490.g004

risk of collective loss,  $p=90\%$  with our results reveals that the evolutionary simulations overestimate the extent of procrastination (see Figure 5). There might be several reasons for this discrepancy: first, in our model individuals only consider the total contributions so far. However, the subjects in experiments may also base their actions on the outcome of the previous round, or on individual behaviors. Unfortunately, reasons behind subjects' actions are still wanting and thus we chose to focus on total contributions. Second, the subjects in the experiments were only allowed to play the game once and thus they did not have the opportunity to learn and adapt their strategies, as assumed in our evolutionary simulations. We would therefore expect that experienced subjects exhibit a behavior that is closer to the backward induction solution, as found in other economic interactions [27]. Third, our evolutionary analysis does not include any psychological motives for contributions, such as loss aversion [28], or framing effects suggesting subjects *should* reach the target by contributing a fair amount each round. The presence of such effects would also explain why subjects contributed a considerable amount of their endowment even in treatments where the risk of collective loss was only 10%, in which case non-contribution would have been the individual and the social optimum. However, if the game is played repeatedly, one might also expect that the impact of these psychological motives decreases [29], and the observed timing of contributions might reveal a stronger tendency to procrastinate.

## Discussion

Examples, such as the prevention of dangerous climate change, or donations to charities, show that many collaborative efforts do not take the form of a one-shot game. Instead, individuals often have the option to await the others' decisions, or to influence others by taking the lead. Here, we have studied how the inclusion of time affects equilibrium selection in a collective-risk dilemma. As a result, we find that time greatly enhances the probability to move towards the efficient equilibrium. This positive effect is of particular importance in larger groups, where successful coordination becomes increasingly difficult [10,30]. Moreover, we have shown that an increasing risk of a catastrophic event has a two-fold effect on the expected welfare: on the one hand, players have



**Figure 5. Comparison of the expected timing of contributions according to the simulations with the observed timing in the experiments of Milinski et al. [6].** The bold dashed lines show the linear trend, indicating that in the experiments and in our simulations contributions tend to be delayed towards later stages of the game. Parameters were chosen to fit the rules of the experiment, i.e. group size  $M=6$ , number of rounds  $R=10$ , initial endowment  $E=40$ , and individuals are allowed to contribute either 0, 2 or 4 monetary units per round. The other parameters are set to the values in the previous figures.

doi:10.1371/journal.pone.0066490.g005

a stronger incentive to coordinate on the beneficial equilibrium, on the other hand it also increases the expected loss upon failure. As a consequence, high risks do not represent the worst-case scenario; rather intermediate risks pose the biggest challenge in collective-risk dilemmas. This result recovers previous observations that severe crises may be actually beneficial for a society, since they increase the probability that necessary measures are adopted [31].

While the inclusion of time facilitates cooperation, it also promotes the evolution of procrastination (which is in line with timing models for public good games, see e.g. [32]). In the extreme case, this may result in strategies that contribute 90% of their share in the very last round (as for example in the two-round game shown in Fig. 4a). Taken together, this may come as a surprise: if players hardly contribute in the early stages of the game anyway, why does the inclusion of these stages increase the probability of successful coordination? It turns out that the fact that most evolutionary trajectories lead towards delayed contributions does not diminish the importance of the early stages. Early contributions help the group to escape from non-cooperative states by motivating conditional cooperators to join in. Once cooperation is established, individuals learn to delay their efforts, because late contributors are less prone to exploitation. Thus, even if early contributions diminish in the long run, they play an important role as a catalyst for cooperation.

Our results thus highlight the importance of time in overcoming coordination problems. Some studies take an opposite view; for example, Drazen and Grilli [31] argue that necessary economic reforms may be delayed if one party attempts to shift the burden of stabilization onto socioeconomic groups that are represented by the other party. In their model, delayed contributions come with a cost, since it prolongs the time spent in an inefficient status quo. In contrast, we have assumed that late and early efforts do not differ in their welfare implications. This may be considered as a limiting case for coordination problems where delayed actions are costly, but where the cost of procrastination is low compared to the stakes in the game (for a model that includes such a cost on late contributions, see [14]). However, *time* should not be taken literally; several instances of collective-risk dilemmas are played over a rather short period (such as efforts to build an emergency sandbag levee by neighbors to protect their community from a flood, [6]). What is crucial, though, is that each player can, directly or indirectly, observe the co-players' actions: it is the flow of information that transforms a one-shot game into a dynamic game, rather than the actual time span (this transformation of the game structure is exactly what is intended when recent donations to charities are publicly announced, instead of made privately, e.g. [33]).

Various generalizations of our model can be addressed. First, we have been considering a homogeneous group, where all individuals are affected equally, and where the quality of contributions does not differ across subjects. Recently, there has been an increasing interest in the impact of inequality [34,35], investigating the question whether "richer" players would be willing to do a bigger share of the target. While these experiments indeed find that individuals with a high endowment contribute more to the common pool, it was also shown that inequality in general reduces the chance of reaching the target. Second, in our model players could not communicate directly; they could only convey their intentions through their contributions. In contrast, some treatments of Tavoni et. al. [35] allowed subjects to make (non-binding) pledges. Despite being cheap talk, the opportunity to communicate intended contributions increased the success rate dramatically. Typical game-theoretic models have problems to reproduce such an effect of pre-play communication. However, if the game is not

considered in isolation, but if players have a reputation to lose (which may affect their performance in future interactions), then modeling the advantages of making pledges seems to be feasible. Herein, we were interested in the human's natural propensity to use time and information to overcome coordination problems, and to motivate others to cooperate. Thus, we have started from a comparably simple model, mimicking the setup of the experiments in [6]. However, we believe that additional communication possibilities will even enhance the group's ability to coordinate on a beneficial equilibrium.

## References

- Milinski M, Semmann D, Krambeck HJ, Marotzke M (2006) Stabilizing the earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences USA* 103: 3994–3998.
- Raihani N, Aitken D (2011) Uncertainty, rationality and cooperation in the context of climate change. *Climatic Change* 108: 47–55.
- Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences USA* 109: 17372–17376.
- Summers LH (2000) International financial crises: Causes, prevention, and cures. *American Economic Review* 90: 1–16.
- Schelling TC (1960) *The strategy of conflict*. Harvard University Press, Cambridge, MA.
- Milinski M, Sommerfeld RD, Krambeck HJ, Reed FA, Marotzke J (2008) The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences USA* 105: 2291–2294.
- Rapoport A, Eshed-Levy D (1989) Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes* 44: 325–344.
- Wang J, Fu F, Wu T, Wang L (2009) Emergence of social cooperation in threshold public good games with collective risk. *Physical Review E* 80: 016101.
- Wang J, Fu F, Wang L (2010) Effects of heterogeneous wealth distribution on public cooperation with collective risk. *Physical Review E* 82: 016102.
- Santos FC, Pacheco JM (2011) Risk of collective failure provides an escape from the tragedy of the commons. *Proceedings of the National Academy of Sciences USA* 108: 10421–10425.
- Greenwood G (2011) Evolution of strategies for the collective-risk social dilemma relating to climate change. *Europhysics Letters* 95: 40006.
- Chen X, Szolnoki A, Perc M (2012) Averting group failures in collective-risk social dilemmas. *Europhysics Letters* 99: 68003.
- Chen X, Szolnoki A, Perc M (2012) Risk-driven migration and the collective-risk social dilemma. *Physical Review E* 86: 036101.
- Abou Chakra M, Traulsen A (2012) Evolutionary dynamics of strategic behavior in a collective-risk dilemma. *PLoS Computational Biology* 8: e1002652.
- Diekmann A (1985) Volunteer's dilemma. *Journal of Conflict Resolution* 29: 605–610.
- Archetti M (2009) Cooperation as a volunteer's dilemma and the strategy of conflict in public goods games. *Journal of Evolutionary Biology* 11: 2192–2200.
- Weesie J (1993) Asymmetry and timing in the volunteer's dilemma. *Journal of Conflict Resolution* 37: 569–590.
- Weibull JW (1995) *Evolutionary Game Theory*. Cambridge: MIT Press.
- Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.
- Nowak MA (2006) *Evolutionary Dynamics*. Harvard University Press, Cambridge.
- Sigmund K (2010) *The calculus of selfishness*. Princeton Univ. Press.
- Marx LM, Matthews SA (2000) Dynamic voluntary contribution to a public project. *Review of Economic Studies* 67: 327–358.
- Taylor PD, Jonker L (1978) Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40: 145–156.
- Hartl D, Clark A (1997) *Principles of Population Genetics*. Sinauer Associates, Inc., Sunderland, Massachusetts, third edition.
- Imhof LA, Nowak MA (2006) Evolutionary game dynamics in a Wright-Fisher process. *Journal of Mathematical Biology* 52: 667–681.
- Ciccarelli JN & S N White (2009) *Psychology*. New Jersey: Pearson Education.
- Palacios-Huerta I, Volij O (2009) Field centipedes. *American Economic Review* 99: 1619–1635.
- Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263–291.
- Cookson R (2000) Framing effects in public goods experiments. *Experimental Economics* 3: 55–79.
- Santos FC, Vasconcelos V, Santos MD, Neves P, Pacheco JM (2012) Evolutionary dynamics of climate change under collective-risk dilemmas. *Mathematical Models and Methods in Applied Sciences* 22: 1140004.
- Drazen A, Grilli V (1993) The benefit of crises for economic reforms. *The American Economic Review* 83: 598–607.
- Gradstein M (1992) Time dynamics and incomplete information in the private provision of public goods. *Journal of Political Economy* 100: 581–597.
- Romano R, Yildirim H (2001) Why charities announce donations: a positive perspective. *Journal of Public Economics* 81: 423–447.
- Milinski M, Röhl T, Marotzke J (2011) Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic Change* 109: 807–814.
- Tavoni A, Dannenberg A, Kallis G, Löschel A (2011) Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences USA* 108: 11825–11829.

## Acknowledgments

We are grateful for Manfred Milinski's comments and suggestions and for providing us with the experimental data obtained in [6].

## Author Contributions

Conceived and designed the experiments: CH MAC PA AT. Performed the experiments: CH MAC PA AT. Analyzed the data: CH MAC PA AT. Wrote the paper: CH MAC PA AT.