Caring for the future can turn tragedy into comedy for long-term collective action under risk of collapse

Wolfram Barfuss[a,b,c,d,1], **Jonathan F. Donges**[a,e], **Vítor V. Vasconcelos**[f,g,h], **Jürgen Kurths**[b,d,i], **and Simon A. Levin**[f,1]

[a]Earth System Analysis, Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, 14473 Potsdam, Germany; [b]Complexity Science, Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, 14473 Potsdam, Germany; [c]Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany; [d]Department of Physics, Humboldt University Berlin, 12489 Berlin, Germany; [e]Stockholm Resilience Centre, Stockholm University, 11419 Stockholm, Sweden; [f]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; [g]Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ 08544; [h]Princeton Institute for International and Regional Studies, Princeton University, Princeton, NJ 08544; and [i]Department of Biology, Saratov State University, Saratov 410012, Russia

Contributed by Simon A. Levin, March 19, 2020 (sent for review October 11, 2019; reviewed by Therese Lindahl and Arne Traulsen)

We will need collective action to avoid catastrophic climate change, and this will require valuing the long term as well as the short term. Shortsightedness and uncertainty have hindered progress in resolving this collective action problem and have been recognized as important barriers to cooperation among humans. Here, we propose a coupled social–ecological dilemma to investigate the interdependence of three well-identified components of this cooperation problem: 1) timescales of collapse and recovery in relation to time preferences regarding future outcomes, 2) the magnitude of the impact of collapse, and 3) the number of actors in the collective. We find that, under a sufficiently severe and time-distant collapse, how much the actors care for the future can transform the game from a tragedy of the commons into one of coordination, and even into a comedy of the commons in which cooperation dominates. Conversely, we also find conditions under which even strong concern for the future still does not transform the problem from tragedy to comedy. For a large number of participating actors, we find that the critical collapse impact, at which these game regime changes happen, converges to a fixed value of collapse impact per actor that is independent of the enhancement factor of the public good, which is usually regarded as the driver of the dilemma. Our results not only call for experimental testing but also help explain why polarization in beliefs about human-caused climate change can threaten global cooperation agreements.

social dilemma | stochastic game | tipping element | time preferences

The internal biophysical and socioeconomic feedback dynamics of the Earth system might limit its future trajectories effectively between two states: a stabilized, habitable and a hothouse Earth state, induced by a series of tipping points and separated by timescales of up to millennia (1). Collective human action is urgently required to steer the Earth system away from such potential thresholds and stabilize it in the habitable state (1, 2). The challenge is to establish successfully these Earth system stewardship policies within the next 50 y (3), in order to avoid locking in hothouse Earth for millennia (1).

In the past, the preconditions for successful collective action have often been studied using social dilemma settings within the framework of normal-form games (4–8). A social dilemma is typically defined as a situation in which any individual prefers the socially defecting choice, regardless of what the other individuals choose, yet all individuals are better off if all choose the socially cooperative option (9, 10).

Collective action under risk of collapse has been studied with threshold public goods games (11) or similarly, collective-risk social dilemmas (12), both experimentally (12–17) and theoretically, using either classical (18) or evolutionary game-theoretic models (19–26). These studies consider the risk of a catastrophic loss if contributions do not exceed a certain threshold. They converge on the finding that greater severity and likelihood of the loss occurring are beneficial for cooperation to emerge. The risk of collective failure provides escape from the tragedy by converting the social dilemma into a coordination challenge (18). Another set of experimental studies in which a common-pool resource is collectively harvested in repeated encounters confirms this general finding. There, when the resource level falls below a threshold, the resource is either destroyed, and the game ends (27–29), or experiences an unfavorable regime shift (30, 31).

However, temporal preferences on when benefits, costs, and catastrophic impacts occur have not been considered explicitly in these studies. Within game theory, so-called folk theorems (32) say that in repeated social dilemma games, cooperation can be sustained when players care enough about future rewards (33). However, folk theorems are concerned with equilibrium payoff profiles in repeated games and do not explore how different equilibria can be reached. Experimentally, temporal factors concerning cooperation, in addition to the social dilemma component, have been considered within so-called resource dilemmas (34, 35). These social traps (36) are associated with positive short-term but negative long-term consequences. It is known that people do discount in collective action settings under delayed

Significance

One of the greatest challenges in addressing global environmental and social problems is achieving cooperation, in which social and environmental processes are increasingly interlinked. Yet, most theoretical studies investigate cooperation within social dilemma settings, using normal form games with effectively only one environmental state. This paper extends the concept of a purely social to a coupled social–ecological dilemma by studying cooperation within stochastic games with multiple environmental states. The particular stochastic game we investigate enables us to study how time preferences influence long-term collective action under risk of collapse. We find that under certain conditions, caring for the future alone can transform this collective action challenge from a tragedy up to a comedy of the commons where cooperation dominates.

Author contributions: W.B., J.F.D., and S.A.L. designed research; W.B. performed research; and W.B., J.F.D., V.V.V., J.K., and S.A.L. wrote the paper.

Reviewers: T.L., Royal Swedish Academy of Sciences; and A.T., Max Planck Institute for Evolutionary Biology.

Competing interest statement: S.A.L. and T.L. are coauthors on a 2016 policy forum.

Published under the PNAS license.

Data deposition: Python code to generate all reported results is available in Zenodo (https://doi.org/10.5281/zenodo.3751564) and GitHub (https://github.com/wbarfuss/EcoPG).

[1]To whom correspondence may be addressed Email: barfuss@pik-potsdam.de or slevin@princeton.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1916545117/-/DCSupplemental.

First published May 20, 2020.

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

benefits and costs (37, 38). This holds true when, additionally, catastrophic tipping is considered (39). However, when costs and benefits are assumed to occur at the same time, time preferences seem to have little effect on cooperation (38, 40, 41). Under completely myopic conditions (i.e., the game ends after one round of play), it has been found that voting can restore cooperation with future generations (42). Experimental treatments that have explicitly studied the temporal dilemma showed that consideration of future consequences could sustain high levels of cooperation (35). Additionally, patient individuals, measured in the laboratory, are more likely to behave sustainably in the field (40). On the other hand, the opposite effect has also been found, that more impatient individuals tend to use lower extraction rates, a result that partly is due to lower investment in extraction capability (43). Thus, it has been argued that there has been insufficient attention to situations where societies act under long-term risk and that existing knowledge is highly fragmented (44).

Taken together, there exists a research gap concerning how individual time preferences influence long-term collective action under the risk of collapse. We argue that recognizing social and ecological systems as coupled social–ecological systems (45, 46) and translating this concept to game theory can enable progress along these lines.

Therefore, we propose to refine the concept of social dilemmas to social–ecological dilemmas. While the former can be studied using (repeated) normal-form games, the latter will be studied using stochastic games. Stochastic games extend repeated normal-form games by incorporating multiple environmental states (47, 48), which can affect the actors' available actions, observations, and current rewards. Transitions between states depend on chosen actions and generally occur probabilistically. Only recently, it has been suggested to study the evolution of cooperation also in stochastic games (49).

We introduce a particular stochastic game to study the preconditions for successful intertemporal collective action under risk of collapse. The game extends the established public goods game by introducing an environmental tipping element, which is why we term it Ecological Public Good (EcoPG). The cooperative action resembles Earth system stewardship policies, such as reducing greenhouse gas emissions, enhancing or creating carbon sinks, and modifying the Earth's energy balance (1). Thus, actors are interpreted best as representatives of geopolitical units, such as states, unions of states, or cities. Cooperating actors contribute a cost c to the public good, which gets multiplied by the public goods enhancement factor f and equally distributed to all actors, reflecting the marginal benefits of avoiding gradual climate change (15, 16), as long as the environment does not collapse. If not specified otherwise, we set $c = 5$ and $f = 1.2$ throughout the paper. Defection in the EcoPG resembles business as usual policies of the participating actors of continued greenhouse gas emissions and biosphere degradation (1). It is therefore not only associated with the socially suboptimal choice but also with a probability to collapse the environment into a degraded (hothouse Earth) state, reflecting a marginal risk of collapse. The more actors defecting, the more likely a collapse becomes (with an increase in collapse probability of q_c/N per defector), regardless of whether this uncertainty in when the collapse will occur results from the current imperfect state of scientific knowledge (15) or is truly inherent in the tipping system (50). A timescale of Earth system stewardship policies to be implemented within the next $T_c = 50$ y (3) translates to a default value of the collapse leverage $q_c \approx 1/T_c = 0.02$, if not specified otherwise. In the degraded state, the actors receive a negative environmental impact payoff $m < 0$, and only the cooperation action opens the chance to recover to the prosperous state (with an increase in recovery probability of q_r/N per cooperator). A timescale of a potential lock-in in the hothouse Earth for up to millennia (1) translates to a default value of the recovery leverage $q_r = 0.0001$, if not specified otherwise. Only joint cooperation will keep the system safely in the prosperous (stabilized Earth) state (Fig. 1 and detailed description in *Materials and Methods*). A detailed comparison between different modeling choices of collapse avoidance games is in *SI Appendix*.

Results

We are interested in the preconditions under which actors decide to cooperate in the prosperous state. In order to make our results comparable with the one-shot decision studies of collective action under risk of collapse (15, 16), we assume that the actors employ stationary Markov strategies (51, 52). Thus, actors only base their decision whether to cooperate or not on the current state of the environment. They do not need to be able to observe other actors' actions. Further, we assume that they aim to maximize the sum of discounted future rewards with discount factor $0 \le \gamma < 1$. A high discount factor γ (i.e., low discount rate) denotes high caring for future rewards (53). We interpret the discount factor as an actor's normative attitude of how much the actor values the future. Based on these two assumptions, we can transform our EcoPG stochastic game into a metagame in normal-form and analyze the game equilibria of this metagame (*Materials and Methods* has details). Specifically, we are interested in the critical parameters determining 1) whether the situation is a dilemma at all or collapse avoidance is in fact suboptimal (Dilemma); 2) whether or not actors are

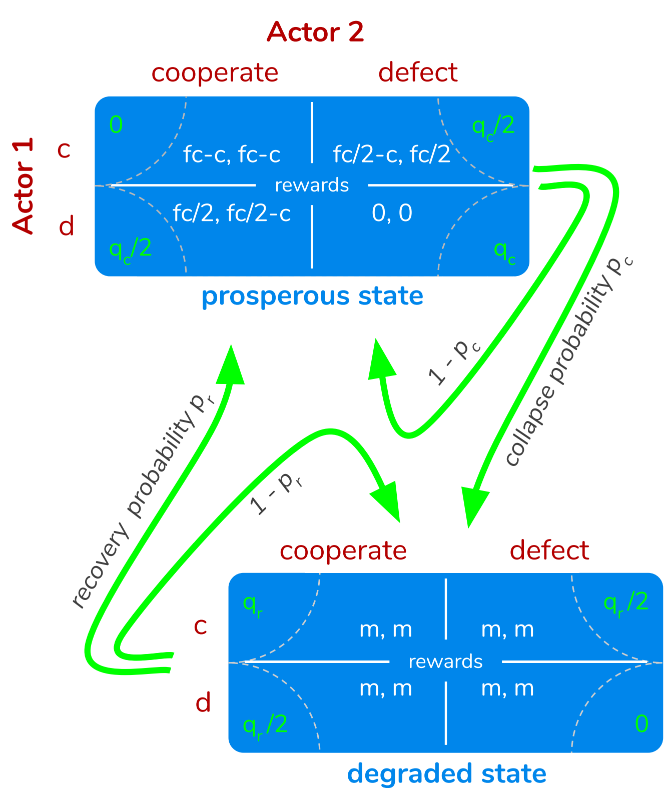

Fig. 1. EcoPG (shown for $N = 2$ actors) extends the repeated public good game to a stochastic game with two environmental states. In the prosperous state, the actors play a standard public good game, in which cooperative costly contributions c get multiplied by the enhancement factor f and equally distributed to all actors. In the degraded state, each actor has to endure an environmental collapse impact $m < 0$. State transitions depend on the joint actions, occur probabilistically, and are visualized with green arrows. Each defecting (cooperating) actor increases the collapse (recovery) probability by q_c/N (q_r/N). Throughout the paper, we set $f = 1.2$, $c = 5$, $q_r = 0.0001$, and $q_c = 0.02$, if not specified otherwise.

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

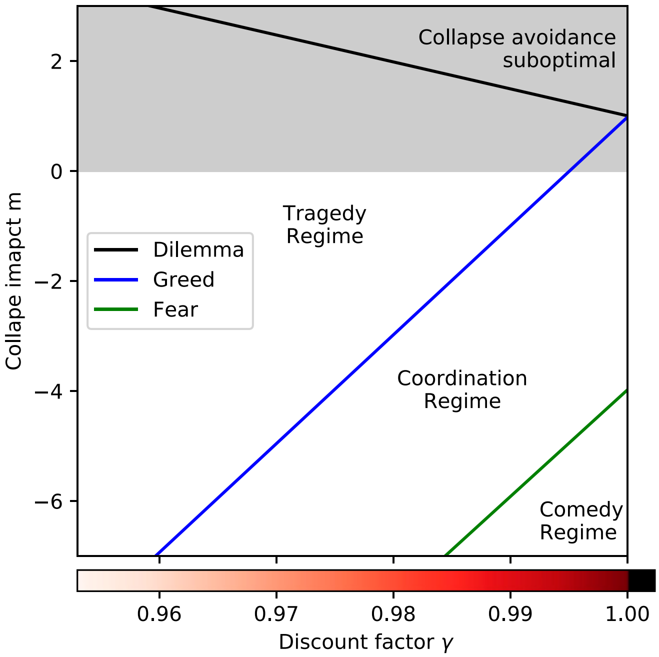

Fig. 2. Regimes in parameter space. Parameter space spanned by the collapse impact m vs. the discount factor γ. Above the Dilemma curve, collapse avoidance is suboptimal due to the positive impact of "collapse." Collapse impacts $m > 0$ are not meaningful within the interpretation of our model and thus appear shaded. Between the Dilemma curve (black) and the Greed curve (blue), the game is a tragedy, and defection dominates. Between the Greed curve and the Fear curve (green), the game requires coordination; both mutual defection and mutual cooperation are equilibria. Below the Fear curve, the game is a comedy where only cooperation prevails. Remaining parameters are $N = 2$, $c = 5$, $f = 1.2$, $q_c = 0.02$, and $q_r = 0.0001$.

greedy to exploit others (Greed); and 3) whether or not actors fear to be exploited by others (Fear) (5).

From Tragedy to Comedy. Fig. 2 shows when these critical conditions are met in the parameter space spanned by the collapse impact m vs. the discount factor γ. Between the Dilemma curve in black and the Greed curve in blue, the game is a classical tragedy of the commons social dilemma (9). Below or right from the blue curve, no greed drives the actors toward defection. Above or left from the green Fear curve, the game becomes a coordination game with at least two Nash equilibria: mutual cooperation and mutual defection (18). Below or right of the Fear curve, there is no more fear that drives cooperating actors toward defection. Here, cooperation is the only Nash equilibrium, and the tragedy has turned into a comedy (54). Thus, our model contains the full drama of the commons, as Ostrom et al. (55) have put it, since the commons can entail tragedy and comedy, or something in between. Note that what we call comedy has also been referred to as harmony game (21).

From Fig. 2, we can reproduce previous findings that more caring for the future is beneficial for cooperation in stochastic games (49). Likewise, a more severe collapse impact is beneficial for cooperation (12, 20) by turning the tragedy into a coordination challenge (18).

Remarkably, we find that the game can be transformed even into a comedy of the commons, where cooperation prevails. Furthermore, we find that, for sufficiently severe collapse impacts, the actors' own parameter of how much they care for the future (their discount factor) alone can determine whether the game is a tragedy, coordination challenge, or a comedy of the commons.

Timescales and Discounting. We found that caring for the future can turn the tragedy into a comedy. What is the influence of the timescale parameters of collapse and recovery, q_c and q_r, on the emergence of these game regimes?

Caring for the future has a qualitatively greater effect at low collapse leverages (Fig. 3). For low collapse leverages, q_c, the discount factor can determine the game regime out of the three regimes (tragedy, coordination, comedy). For high q_c, discounting can only determine the game regime between at most two neighboring regimes. The critical value of collapse leverage $q_{c,\text{crit}} \approx 0.7$ (i.e., where the dark green and the light blue curves intersect) does not depend strongly on the recovery leverage q_r.

The emergence of these game regimes are only sensitive to the recovery leverage, q_r, when caring for the future γ is high and recovery occurs more quickly than collapse (i.e., the recovery leverage q_r is greater than the collapse leverage q_c). At no caring for the future at all ($\gamma = 0$), the timescale of recovery has no effect on game regimes, as the coinciding dotted, dashed, and straight curves in the light colors show.

For sufficiently negative collapse impact, m, it is beneficial for cooperation when collapse is likely to occur soon (i.e., collapse leverage q_c is high) and chance for recovery is distant in the future (i.e., recovery leverage q_r is low).

Size of the Collective. What is the influence of the number of participating actors N on the emergence of the game regimes? Fig. 4 shows the critical collapse impacts for all three conditions (Dilemma, Greed, and Fear), rescaled by the number of actors m/N vs. N for three different public good enhancement factors f, which are usually the driver of the dilemma in the standard normal-form public good game.

Overall, Fig. 4 shows a diffusion of responsibility (56). Interestingly, all three conditions converge to a fixed value of m/N each, independent of N in the limit of large N. With more actors participating, each actor on its own has less leverage to cause the collapse, since the marginal collapse probability q_c/N scales with N. To counterbalance this effect, the collapse impact, m, that each actor experiences must become more severe in order to cause an urge to cooperate.

Most interestingly, these values of m/N are independent of the enhancement factor f. Thus, at large N, it does not depend on the quality of the public good of the prosperous state where the different game regimes lie. This means that even collapse

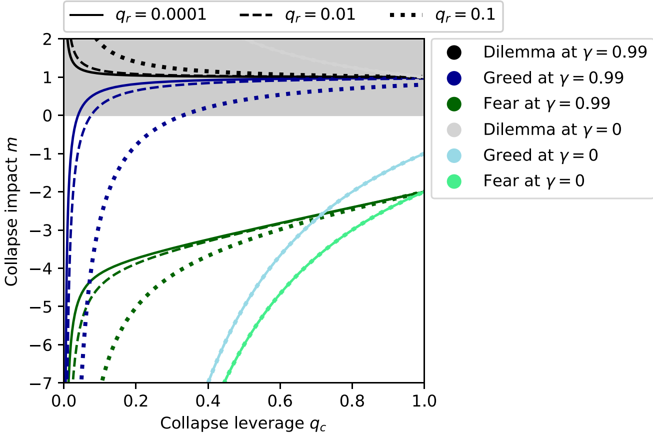

Fig. 3. Dependence on timescales. Three critical conditions (Dilemma, Greed, Fear) in the parameter space spanned by the collapse impact m vs. the collapse leverage q_c for both extremes of discounting ($\gamma = 0$ and $\gamma = 0.99$) and three different recovery leverages ($q_r = 0.0001$, $q_r = 0.01$, and $q_r = 0.1$). Remaining parameters are $N = 2$, $c = 5$, and $f = 1.2$.

ENVIRONMENTAL SCIENCES

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

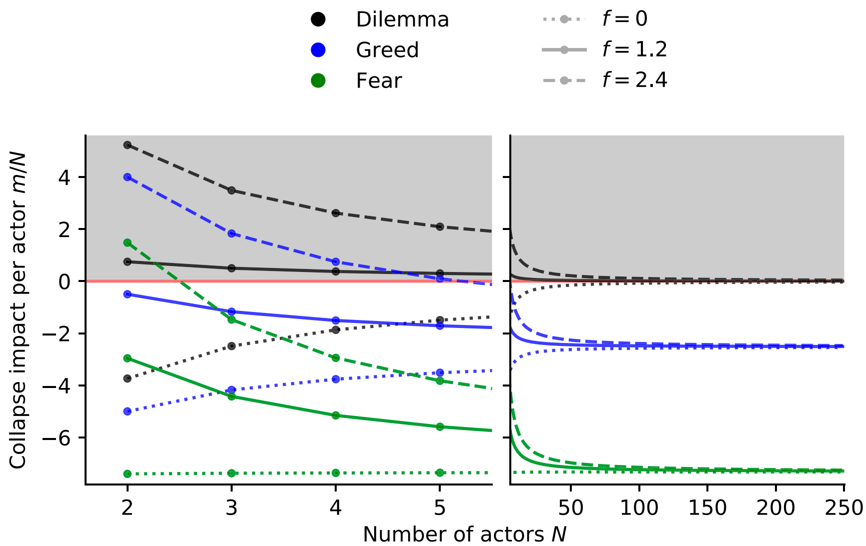

Fig. 4. Dependence on number of actors N. Three critical conditions (Dilemma: black; Greed: blue; Fear: green) in the parameter space spanned by the collapse impact per actor m/N vs. the number of actors N for discount factor $\gamma = 0.99$. Results for three different public good enhancement factors $f = 0$, $f = 1.2$, and $f = 2.4$ are shown with different line styles. Remaining parameters are $c = 5$, $q_c = 0.02$, and $q_r = 0.0001$.

avoidance games with no marginal benefit (12–14, 20–26) have the same critical collapse impacts per actor at large N. Note also that these critical m/N values are already in good agreement with each other for the different values of f at the range of N between 150 and 280 (Fig. 4). This range includes both Dunbar's number $N \approx 150$ being the cognitive limit of humans for maintaining stable groups (57) as well as the number of countries in the world. Thus, this result might be of relevance to both the local and the global level.

Learning Dynamics. So far, we have studied the preconditions under which the game presents itself as an effective social dilemma tragedy, coordination game, or comedy, where cooperation dominates. We did so by transforming our stochastic game model into a metagame in normal-form and determining the resulting equilibria. Yet, in the coordination regime, such an analysis cannot tell which equilibrium is selected.

To complement our metagame analysis, we use multiagent actor–critic reinforcement learning dynamics (58), similar to the replicator dynamics used in evolutionary game theory, with which the agents are capable of learning in stochastic games (Fig. 5).

We find that these learning dynamics fit well to the metagame equilibrium analysis. Comparing Fig. 2 at $\gamma = 0.99$ with the strategy phase space in Fig. 5, *Left* reveals the following relationship: the tragedy and comedy regimes each have a single attractor in the prosperous state of the learning dynamics—mutual defection for the tragedy, mutual cooperation for the comedy. The coordination regime corresponds to a dynamically bistable regime of the learning dynamics in the prosperous state. The degraded state on its own is strategically irrelevant as rewards are identical for all actions. There, only cooperation will eventually lead back to the more rewarding prosperous state, and actors easily learn to cooperate in the degraded state. Thus, the learning dynamics can be used as a foundation to explain the emergence of cooperation or defection, respectively.

It depends on the initial behavioral strategy whether actors will learn to cooperate or not within the parameter conditions of the coordination regime. The dashed red line in the strategy space in Fig. 5*B* indicates those points at which the learning dynamics do neither lead to more cooperative nor defective behavior. It is able to divide the strategy phase space into a cooperative and a defective basin of attraction, as exemplary trajectories show. Thus, the more cooperative the initial strategy is, the more likely it is that actors will learn to cooperate within the coordination regime.

What is the influence of the actors' attitude of how much they care for the future γ within the learning dynamics? In Fig. 5, *Right*, the critical discount factors in strategy phase space are shown. Given an initial strategy, both actors need a discount factor of at least this critical discount factor to learn to cooperate. Since the discount factor is confined by one, even the most future caring actors cannot escape the black areas in Fig. 5 and learn to cooperate. This result highlights that caring for the future is not guaranteed to be capable of resolving the tragedy within the coordination regime.

Under what conditions does such a, metaphorically speaking, black hole for the learning of cooperation exist? Observing Fig. 2, we conclude that black areas in strategy phase space exist within the bistable coordination regime as long as the collapse impact is not severe enough for a comedy regime to emerge. Thus, only the possibility of entering a comedy through sufficient caring for the future has an effect for the coordination regime. It ensures that cooperation can be obtained from anywhere in the strategy space, even the most grim initial conditions, as long as actors care enough about the future.

Discussion

The Earth system requires collective action in order to enter a long-term stabilized Earth system state and simultaneously to avoid the collapse to very unfavorable conditions for human development (1). In this article, we investigated the preconditions of success for such a collective action challenge. We did so by conceptually extending the well-studied setting of a social dilemma to a social–ecological dilemma. While social dilemmas are often studied using normal-form games, we investigated a coupled social–ecological dilemma using a stochastic game with multiple environmental states. We introduced a particular stochastic game, extending the established public goods game by an environmental tipping element. Thus, we termed this game the EcoPG. Overall, our study demonstrated that stochastic games are a suitable tool for the mathematics of sustainability (59) in order to ask and answer questions concerning social–ecological systems (45, 46).

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

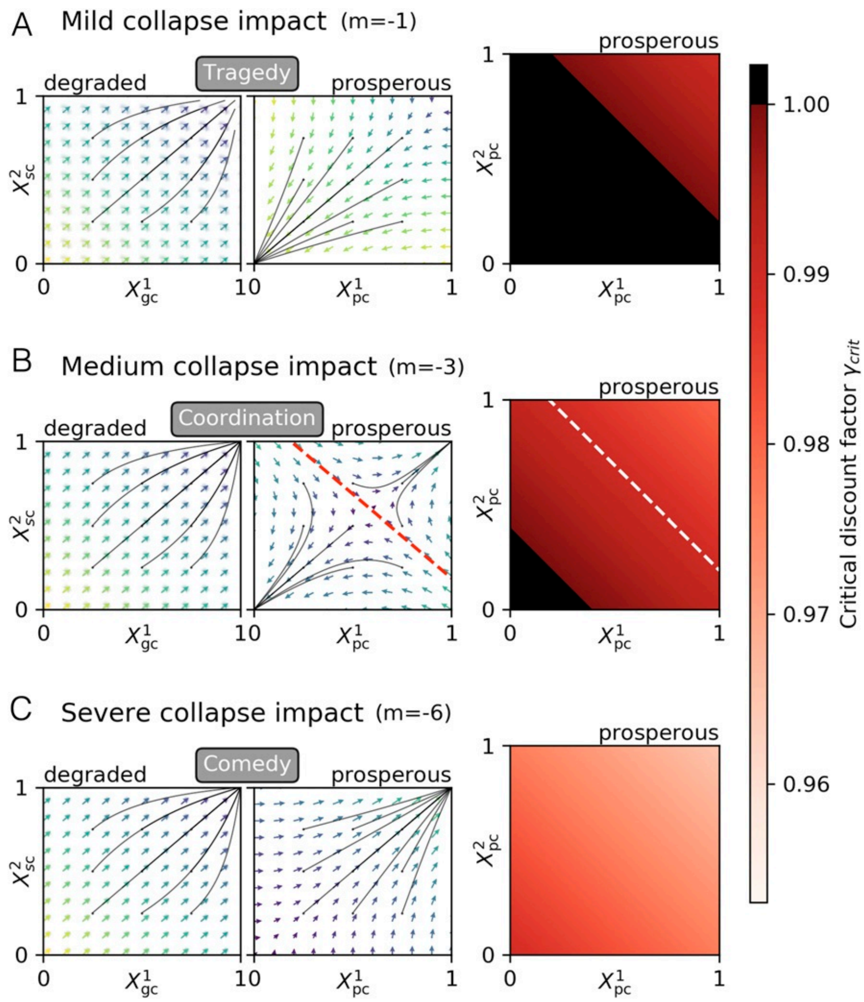

Fig. 5. Learning dynamics. Actor–critic reinforcement learning dynamics for mild (*A*; $m = -1$), medium (*B*; $m = -3$), and severe (*C*; $m = -6$) collapse impacts. On the *Left*, strategy phase spaces for exemplary discount factor $\gamma = 0.99$ are shown. X^i_{sc} denotes the probability that actor *i* plays the cooperative action in state $s \in \{p, g\}$. The arrows indicate the average direction the learning dynamics drive the actors toward for each state, averaged over all strategy space points of the other state. The dashed red curve at medium collapse impact results from the ansatz where these arrows have *x*, *y* components of equal length but different sign. It is able to divide the behavior space into a cooperation and a defection basin of attraction, as exemplary trajectories, shown in gray, indicate. On the *Right*, the critical discount factor γ_{crit} is shown in the strategy space section of the prosperous state. From an arbitrary point in strategy space, both actors need a discount factor of at least γ_{crit} to learn the cooperative solution. Thus, from areas in black, there is no such γ since the discount factor is confined by one. Remaining parameters are $c = 5$, $q_c = 0.02$, and $q_r = 0.0001$.

Specifically, we contributed to filling a research gap by theoretically exploring the interdependence of the three components of the problem; these are 1) timescales and time preferences (collapse and recovery leverage q_c and q_r, the actors' discount factor γ), 2) the magnitude of the collapse impact m, and 3) the size of the collective (the number of actors N).

For each component individually, we could reproduce previous theoretical or empirical findings. More caring for the future (49) as well as a more severe collapse impact (12, 18, 21) are beneficial for cooperation. For example, Barrett (18) showed that the impact of a collapse can transform the social dilemma into a coordination challenge, which countries are generally good at solving in international treaties.

We uncovered that the actors' normative choice of how much to care for future rewards (expressed in their discount factor γ) can determine the full drama of the commons (55) [i.e., whether the whole game is a tragedy with dominating defection (60), requires coordination between defection and cooperation (18), or even is a comedy in which cooperation dominates (54)]. The discount factor can determine which of those three regimes the game falls into only when a collapse is not expected shortly and the negative collapse impact is sufficiently severe. Thus, individual time preference can serve as a social tipping element (61, 62).

In this study, we used the most established and simple model of intertemporal choice, exponentially discounting. However, it has little empirical support as a model for intertemporal choice of humans (63). For example, discount factors increase over time and with the magnitude of rewards and are generally larger for losses than for gains (64). Our study is of qualitative nature, highlighting the possibility and tendencies for the existence of different preference regimes with respect to the actors' caring for future rewards. Future work needs to incorporate alternative models of intertemporal choice (e.g., hyperbolic discounting) as well as address their cognitive foundations.

In contrast to more complex strategies in repeated games, which take other actors' actions into account (32), we focused our analysis on Markov strategies in stochastic games that base their action only on the current environmental state (51). Thus, one can argue that the strategic social reciprocity between actors (that is required for cooperation in repeated games) has been transferred to an environmental reciprocity in the stochastic

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

game. In the former case, deviations from cooperation are punished by other deviating actors, whereas in the latter case, they are punished by a collapsing environment. Both types of reciprocity require a sufficiently large caring for future rewards. Future work is required to address the interrelation between social and environmental reciprocity.

We found a diffusion of responsibility (56) in our model with increasing number of actors. At large N, the game regime borders converge to a value of collapse impact per actor m/N each, independent of N, and most interestingly also independent of the public good enhancement factor f. Thus, marginal benefits of avoiding collapse become irrelevant for a large collective whose decision is dominated by the prospect of collapse.

Especially with respect to responsibility, a major obstacle to global cooperation in order to mitigate climate change and biosphere degradation is asymmetries between actors (13, 14, 25, 26, 65). Costs, benefits, and impacts are not equally distributed. Likewise, the impact of actors on the environment, which causes a potential collapse, is distributed heterogeneously. There exist hierarchies between actors. Within this study, we have focused on the theoretical baseline of homogeneous, symmetric actors. Future work is required to investigate our model for the preconditions of successful collective action with asymmetric actors.

We showed that multiagent actor–critic learning dynamics (58) can be used within our model as a microfoundation to explain the emergence of cooperation or defection, respectively. Parameter regions that reveal a coordination challenge from a game-theoretic point of view correspond to a bistable strategy phase space from a learning dynamics perspective. The tragedy regime corresponds to a strategy phase space with a mutually attracting defection fixed point, and the comedy regime corresponds to an all-attracting cooperating fixed point in the prosperous state. On the one hand, using these learning dynamics is a technical method to confirm and refine the equilibria analysis, as it is done often in studies that use evolutionary game theory. However, one could argue that it is not particularly realistic that the actors use immediate feedback from the environment to update their action probabilities, especially in the context of climate change. On the other hand, these learning dynamics can also be interpreted as reinforcement learning in the infinite memory limit (66). Thus, instead of immediate feedback, the actors use a model of the world and update their actions based on how they imagine the environment behaves. Future work needs to address such kinds of cognitive foundations for sustainability. Furthermore, future work is required to gain improved insights on how these learning dynamics compare with other types of adaptation dynamics under environmental feedback (67).

This learning dynamics perspective revealed that only the existence of the comedy regime has consequences for the coordination regime. As long as there is no comedy reachable through sufficient caring for the future, a black hole regime in strategy phase space exists. From within such a black hole, there is no discount factor, such that the actors would be able to learn the cooperative solution. Even the most future-caring actors prefer to collectively suffer in an environmental collapse rather than to cooperate in a prosperous environment. This interlinked social–ecological dilemma presents an interesting challenge for future research: is it possible to design nontrivial learning rules that find the cooperative solution in this regime?

Our results highlight the importance of individual time preferences for successful cooperation under risk of collapse. Since behavioral experiments have shown that people do discount (37–39), this calls for experimental designs to investigate how human time preferences influence the challenge to avoid collapse collectively. It is important to test, for example, whether one can distinguish with human players between a coordination challenge and a comedy and if so, do they experience a black hole; whether the critical collapse impact at the game regime borders indeed scales with N at large N, and whether these critical N/m are independent of the public good enhancement factor f.

As the discount factors are individual preferences, our results highlight that actors' individual attitudes can determine whether or not cooperation can emerge. Actors' individual attitudes not only consist of how much they care for the future but likewise, how much they believe that collapse is likely and severe. In that sense, our model offers a possible explanation for how the increased polarization with respect to the beliefs about human-made climate change (68) poses a threat to the stability of global cooperation agreements, such as the Sustainable Development Goals and the Paris Accord (69).

Materials and Methods

EcoPG. In a classical normal-form public good game with two actions, each actor $i \in \{1, \ldots, N\}$ can choose to either cooperate (c) or defect (d). Cooperators contribute with a cost c to the public good. All contributions get multiplied by an enhancement factor f and subsequently distributed to all actors. Thus, defectors receive a reward of $r_\mathrm{d} = N_\mathrm{c} fc/N$ and cooperators $r_\mathrm{c} = N_\mathrm{c} fc/N - c$, where N_c denotes the total number of cooperators. If not specified otherwise, we set $f = 1.2$ and $c = 5$.

Our EcoPG substantially extends this normal-form public good game to a stochastic game (70). It consists of an environment with two environmental states s: a prosperous (p) and a degraded (g) one. Actors interact repeatedly with each other and the environment, which changes its state probabilistically.

The probability of moving to the next state s' depends on the current state s and the current joint action $\mathbf{a}$ of all actors. Each defector increases the collapse probability p_c from the prosperous to the degraded state by q_c/N, with $0 \leq q_c \leq 1$. Thus, $p_c = N_\mathrm{d}/N \cdot q_c$ with $N_\mathrm{d} = N - N_\mathrm{c}$ being the total number of defectors. Consequently, the probability of remaining in the prosperous state reads $1 - p_c$. From the degraded state, we assume the probability of recovery to the prosperous state p_r increases with each cooperator by q_r/N, with $0 \leq q_r \leq 1$. Thus, $p_r = N_\mathrm{c}/N \cdot q_r$, and the probability of remaining in the degraded state reads $1 - p_r$. Assuming that one time step of our model corresponds to 1 y, a time frame of 50 y of possible Earth system stewardship and 10,000 y of a possible hothouse Earth make $q_c = 0.02$ and $q_r = 0.0001$ natural choices of these parameters.

In a stochastic game, rewards generally depend on the current environmental state s, the current joint action $\mathbf{a}$, and the next state s'. We assume that the actors receive the rewards from the classical public good only when they remain in the prosperous state: $R^i_\mathrm{pap} = r_\mathrm{c}$ if $a^i = \mathrm{c}$ and $R^i_\mathrm{pap} = r_\mathrm{d}$ if $a^i = \mathrm{d}$. When state transitions involve the degraded state g, the actors only receive an environmental collapse impact $m < 0$: $R^i_\mathrm{pag} = R^i_\mathrm{gag} = R^i_\mathrm{gap} = m$ for all $a^i \in \{\mathrm{c}, \mathrm{d}\}$. The EcoPG model can be regarded as the multiactor version the Markov decision model used in ref. 71. Fig. 1 illustrates the model for $N = 2$ actors.

Metagame Analysis. We focus our analysis on stationary Markov strategies (52). Actors choose their actions based only on the current state: X^i_{sa} denotes the probability that actor i plays action $a \in \{\mathrm{c}, \mathrm{d}\}$ when the environment is in state $s \in \{\mathrm{p}, \mathrm{g}\}$.

As it is commonly done, we also assume that actors want to maximize the sum of exponentially discounted future rewards, called return $G^i(t) = \sum_{k=0}^{\infty} \gamma^k r^i(t+k)$ with $0 \leq \gamma < 1$ being the discount factor. Thus, given a joint Markov strategy $\mathbf{X}$, the value of a state s for an actor i is defined as the expected return, given that the environment started in state s: $V^i_s(\mathbf{X}) = \mathbb{E}_\mathbf{X}[G^i(t) \mid s(t) = s]$.

If we interpret the Markov strategies X^i_{sa} as actions of a metagame in normal-form, the metaaction set of actor i becomes $\{(X^i_{\mathrm{p}a} = 1, X^i_{\mathrm{g}a'} = 1) \mid a, a' \in \{\mathrm{c}, \mathrm{d}\}\}$ and consists of in total four actions. If we now restrict our analysis further to only those metaactions that employ the cooperative action in the degraded state, we are left with only two metaactions: either cooperate or defect in the prosperous state. We do so because the degraded state on its own has no strategic relevance and from there, only the cooperative action is associated with the possibility to recover to the more rewarding prosperous state, given that m is sufficiently negative. Thus, the cooperative metaaction reads $\mathsf{C} = (X^i_\mathrm{pc} = 1, X^i_\mathrm{gc} = 1)$, and the defective metaaction reads $\mathsf{D} = (X^i_\mathrm{pd} = 1, X^i_\mathrm{gc} = 1)$. As the payoffs of this metagame, we take the values of the prosperous state V^i_p.

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

From the perspective of one actor, we denote the payoff of mutual cooperation as the reward $R = V^i_{\mathrm{p}}(\mathrm{C}, \text{ all other C})$; the payoff of unilateral defection as temptation $T = V^i_{\mathrm{p}}(\mathrm{D}, \text{ all other C})$; the payoff of unilateral cooperation as sucker $S = V^i_{\mathrm{p}}(\mathrm{C}, \text{ all other D})$; and the payoff of mutual defection as punishment $P = V^i_{\mathrm{p}}(\mathrm{D}, \text{ all other D})$ using the Prisoner's Dilemma nomenclature of ref. 5.

Whenever $R > P$ and $R > S$ plus one of the following conditions holds, a social dilemma is present: 1) actors are greedy to exploit others (i.e., they prefer unilateral defection over mutual cooperation [Greed: $T > R$]) or 2) actors fear to be exploited (i.e., they prefer mutual defection over unilateral cooperation [Fear: $P > S$]).

We analytically compute three critical curves in parameter space (*SI Appendix*): 1) the condition at which collapse avoidance becomes collectively optimal (Dilemma: $R = P$), 2) the condition at which the actor becomes indifferent to greed (Greed: $R = T$), and 3) the condition where the actor becomes indifferent to fear (Fear: $P = S$). Without greed, the game becomes a coordination game between two pure equilibria of mutual cooperation and mutual defection. Without greed and fear, the only Nash equilibrium left is mutual cooperation.

Due to the stochastic game, the values V^i_{p} are nonlinear functions of the action probabilities. We explicitly ensure that no mixed equilibria exist at the defection-dominated tragedy and the cooperation-dominated comedy regime by checking that $dV^i_{\mathrm{p}}/dX^i_{\mathrm{pc}} = 0$ has no solutions.

Learning Dynamics. Concerning the learning dynamics, we build upon ref. 58. There, a deterministic limit of so-called temporal difference reinforcement learning is derived to yield difference equations similar to the well-known replicator dynamics but capable of learning in a stochastic game where the environment changes depending on the actors' actions (ref. 67 has a related approach to ecoevolutionary dynamics). In essence, the strategy profile X^i_{sa} is updated by a temporal difference error $\mathcal{TD}^i_{sa}(\mathbf{X})$, which effectively functions as a gradient toward more valuable actions.

To estimate the critical minimum discount factor necessary for the actors to learn to cooperate in strategy space, we assume that the actors will end up cooperating if an initial change of behavior will increase overall cooperation behavior. For the special case of $N = 2$, this can be expressed as

$$\mathcal{TD}^1_{\mathrm{pc}}(\mathbf{X}) - \mathcal{TD}^1_{\mathrm{pd}}(\mathbf{X}) \stackrel{!}{=} -\left(\mathcal{TD}^2_{\mathrm{pc}}(\mathbf{X}) - \mathcal{TD}^2_{\mathrm{pd}}(\mathbf{X})\right). \quad [1]$$

ACKNOWLEDGMENTS. This work was developed in the context of the project on Coevolutionary Pathways in the Earth System at the Potsdam Institute for Climate Impact Research (PIK), within the Princeton–Humboldt Cooperation and Collective Cognition Network, and the Princeton–Stockholm Resilience Centre–PIK Resilience Network. This work emerged from one chapter of the doctoral dissertation of W.B. W.B. acknowledges funding from the Heinrich Böll Foundation. J.F.D. acknowledges funding from Earth League's EarthDoc program, European Research Council Advanced Grant Project Earth Resilience in the Anthropocene, and Leibniz Association Project DominoES (Domino effects in the Earth system). V.V.V. acknowledges funding from the Princeton Institute for International and Regional Studies, Rapid Switch Community. S.A.L. and V.V.V. acknowledge funding by the US Defense Advanced Research Projects Agency (D17AC00005) and the NSF Grant GEO-1211972. We thank Jobst Heitzig for insightful discussions.

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.

1. W. Steffen *et al.*, Trajectories of the earth system in the Anthropocene. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8252–8259 (2018).
2. W. Steffen *et al.*, Planetary boundaries: Guiding human development on a changing planet. *Science* **347**, 1259855–1259855 (2015).
3. J. Rockström *et al.*, A roadmap for rapid decarbonization. *Science* **355**, 1269–1271 (2017).
4. R. Axelrod, W. Hamilton, The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
5. M. W. Macy, A. Flache, Learning dynamics in social dilemmas. *Proc. Natl. Acad. Sci. U.S.A.* **99** (suppl. 3), 7229–7236 (2002).
6. M. A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
7. A. Szolnoki, M. Perc, G. Szabó, Phase diagrams for three-strategy evolutionary prisoner's dilemma games on regular graphs. *Phys. Rev.* **80**, 056104 (2009).
8. A. J. Bladon, T. Galla, Learning dynamics in public goods games. *Phys. Rev.* **84**, 041132 (2011).
9. G. Hardin, The tragedy of the commons. *Science* **162**, 1243–1248 (1968).
10. R. M. Dawes, Social dilemmas. *Annu. Rev. Psychol.* **31**, 169–193 (1980).
11. R. T. Croson, M. B. Marks, Step returns in threshold public goods: A meta- and experimental analysis. *Exp. Econ.* **2**, 239–259 (2000).
12. M. Milinski, R. D. Sommerfeld, H. J. Krambeck, F. A. Reed, J. Marotzke, The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2291–2294 (2008).
13. M. Milinski, T. Röhl, J. Marotzke, Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic Change* **109**, 807–814 (2011).
14. A. Tavoni, A. Dannenberg, G. Kallis, A. Löschel, Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11825–11829 (2011).
15. S. Barrett, A. Dannenberg, Climate negotiations under scientific uncertainty. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17372–17376 (2012).
16. S. Barrett, A. Dannenberg, Sensitivity of collective action to uncertainty about climate tipping points. *Nat. Clim. Change* **4**, 36–39 (2013).
17. A. Dannenberg, A. Löschel, G. Paolacci, C. Reif, A. Tavoni, On the provision of public goods with probabilistic and ambiguous thresholds. *Environ. Resour. Econ.* **61**, 365–383 (2014).
18. S. Barrett, Climate treaties and approaching catastrophes. *J. Environ. Econ. Manag.* **66**, 235–250 (2013).
19. J. Wang, F. Fu, T. Wu, L. Wang, Emergence of social cooperation in threshold public goods games with collective risk. *Phys. Rev.* **80**, 016101 (2009).
20. F. C. Santos, J. M. Pacheco, Risk of collective failure provides an escape from the tragedy of the commons. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10421–10425 (2011).
21. F. C. Santos, F. L. Pinheiro, T. Lenaerts, J. M. Pacheco, The role of diversity in the evolution of cooperation. *J. Theor. Biol.* **299**, 88–96 (2012).
22. M. Abou Chakra, A. Traulsen, Evolutionary dynamics of strategic behavior in a collective-risk dilemma. *PLoS Comput. Biol.* **8**, e1002652 (2012).
23. C. Hilbe, M. Abou Chakra, P. M. Altrock, A. Traulsen, The evolution of strategic timing in collective-risk dilemmas. *PLoS One* **8**, e66490 (2013).
24. V. V. Vasconcelos, F. C. Santos, J. M. Pacheco, A bottom-up institutional approach to cooperative governance of risky commons. *Nat. Clim. Change* **3**, 797–801 (2013).
25. M. Abou Chakra, A. Traulsen, Under high stakes and uncertainty the rich should lend the poor a helping hand. *J. Theor. Biol.* **341**, 123–130 (2014).
26. V. V. Vasconcelos, F. C. Santos, J. M. Pacheco, S. A. Levin, Climate policies under wealth inequality. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2212–2216 (2014).
27. J. M. Walker, R. Gardner, Probabilistic destruction of common-pool resources: Experimental evidence. *Econ. J.* **102**, 1149–1161 (1992).
28. A. Botelho, A. Dinar, L. M. C. Pinto, A. Rapoport, Time and uncertainty in resource dilemmas: Equilibrium solutions and experimental results. *Exp. Econ.* **17**, 649–672 (2014).
29. E. Blanco, M. C. Lopez, J. M. Walker, The opportunity costs of conservation with deterministic and probabilistic degradation externalities. *Environ. Resour. Econ.* **64**, 255–273 (2016).
30. C. Schill, T. Lindahl, A. S. Crépin, Collective action and the risk of ecosystem regime shifts: Insights from a laboratory experiment. *Ecol. Soc.* **20**, 48 (2015).
31. T. Lindahl, A. S. Crépin, C. Schill, Potential disasters can turn the tragedy into success. *Environ. Resour. Econ.* **65**, 657–676 (2016).
32. D. Fudenberg, J. Tirole, *Game Theory* (MIT Press, Cambridge, MA, 1991).
33. R. Axelrod, The emergence of cooperation among egoists. *Am. Polit. Sci. Rev.* **75**, 306–318 (1981).
34. L. Hendrickx, W. Poortinga, R. van der Kooij, Temporal factors in resource dilemmas. *Acta Psychol.* **108**, 137–154 (2001).
35. K. V. Kortenkamp, C. F. Moore, Time, uncertainty, and individual differences in decisions to cooperate in resource dilemmas. *Pers. Soc. Psychol. Bull.* **32**, 603–615 (2006).
36. J. Platt, Social traps. *Am. Psychol.* **28**, 641–651 (1973).
37. C. Deck, S. Jahedi, An experimental investigation of time discounting in strategic settings. *J. Behav. Exp. Econ.* **54**, 95–104 (2015).
38. F. Kölle and T. Lauer, Cooperation, discounting, and the effects of delayed costs and benefits. http://dx.doi.org/10.2139/ssrn.3260626 (28 October 2018).
39. J. Jacquet *et al.*, Intra- and intergenerational discounting in the climate game. *Nat. Clim. Change* **3**, 1025–1028 (2013).
40. E. Fehr, A. Leibbrandt, A field study on cooperativeness and impatience in the tragedy of the commons. *J. Publ. Econ.* **95**, 1144–1155 (2011).
41. D. Davis, A. Ivanov, O. Korenok, Individual characteristics and behavior in repeated games: An experimental study. *Exp. Econ.* **19**, 67–99 (2016).
42. O. P. Hauser, D. G. Rand, A. Peysakhovich, M. A. Nowak, Cooperating with the future. *Nature* **511**, 220–223 (2014).
43. A. Javaid, M. M. Kulesz, A. Schlüter, A. Ghosh, N. S. Jiddawi, Time preferences and natural resource extraction behavior: An experimental study from artisanal fisheries in Zanzibar. *PLoS One* **11**, e0168898 (2016).
44. R. Shwom, R. E. Kopp, Long-term risk governance: When do societies act before crisis?. *J. Risk Res.* **22**, 1374–1390 (2019).
45. F. Berkes, C. Folke, *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience* (Cambridge University Press, 1998).
46. S. Levin *et al.*, Social-ecological systems as complex adaptive systems: Modeling and policy implications. *Environ. Dev. Econ.* **18**, 111–132 (2013).
47. L. S. Shapley, Stochastic games. *Proc. Natl. Acad. Sci. U.S.A.* **39**, 1095–1100 (1953).
48. A. Neyman, S. Sorin, *Stochastic Games and Applications* (Springer, 2003).
49. C. Hilbe, Š. Šimsa, K. Chatterjee, M. A. Nowak, Evolution of cooperation in stochastic games. *Nature* **559**, 246–249 (2018).
50. J. Rockström *et al.*, Planetary boundaries: Exploring the safe operating space for humanity. *Ecol. Soc.* **14**, 32 (2009).
51. E. Maskin, J. Tirole, Markov perfect equilibrium. I. Observable actions. *J. Econ. Theor.* **100**, 191–219 (2001).

52. Y. Shoham, K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations* (Cambridge University Press, 2008).
53. O. Ibe, *Markov Processes for Stochastic Modeling* (Newnes, 2013).
54. C. Rose, The comedy of the commons: Custom, commerce, and inherently public property. *Univ. Chi. L. Rev.* **53**, 711–781 (1986).
55. E. E. Ostrom *et al.*, *The Drama of the Commons* (National Academy Press, 2002).
56. J. M. Darley, B. Latané, Bystander intervention in emergencies: Diffusion of responsibility. *J. Pers. Soc. Psychol.* **8**, 377–383 (1968).
57. RI. Dunbar, Neocortex size as a constraint on group size in primates. *J. Hum. Evol.* **22**, 469–493 (1992).
58. W. Barfuss, J. F. Donges, J. Kurths, Deterministic limit of temporal difference reinforcement learning for stochastic games. *Phys. Rev.* **99**, 043305 (2019).
59. S. Levin, The mathematics of sustainability. *Not. AMS* **60**, 11 (2013).
60. G. J. Hardin, The tragedy of the commons. *Science* **162**, 1243–1248 (1968).
61. J. F. Donges *et al.*, Closing the loop: Reconnecting human dynamics to earth system science. *Anthropocene Rev.* **4**, 151–157 (2017).
62. J. F. Donges *et al.*, Taxonomies for structuring models for world-earth system analysis of the Anthropocene: Subsystems, their interactions and social-ecological feedback loops. *Earth Syst. Dyn. Discuss.* **2018**, 1–30 (2018).
63. S. Frederick, G. Loewenstein, T. O'donoghue, Time discounting and time preference: A critical review. *J. Econ. Lit.* **40**, 351–401 (2002).
64. U. Benzion, A. Rapoport, J. Yagil, Discount rates inferred from decisions: An experimental study. *Manag. Sci.* **35**, 270–284 (1989).
65. R. Kline, N. Seltzer, E. Lukinova, A. Bynum, Differentiated responsibilities and pro-social behaviour in climate change mitigation. *Nat. Hum. Behav.* **2**, 653–661 (2018).
66. W. Barfuss, "Reinforcement learning dynamics in the infinite memory limit" in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (International Foundation for Autonomous Agents and Multi-agent Systems, 2020).
67. A. R. Tilman, J. B. Plotkin, E. Akçay, Evolutionary games with environmental feedbacks. *Nat. Commun.* **11**, 915 (2020).
68. R. E. Dunlap, A. M. McCright, J. H. Yarosh, The political divide on climate change: Partisan polarization widens in the U.S. *Environment* **58**, 4–23 (2016).
69. R. O. Keohane, D. G. Victor, Cooperation and discord in global climate policy. *Nat. Clim. Change* **6**, 570–575 (2016).
70. W. Barfuss, "Learning dynamics and decision paradigms in social-ecological dilemmas," PhD thesis, Humboldt-Universität zu Berlin, Berlin, Germany (2019).
71. W. Barfuss, J. F. Donges, S. J. Lade, J. Kurths, When optimization for governing human-environment tipping elements is neither sustainable nor safe. *Nat. Commun.* **9**, 2354 (2018).

Downloaded from https://www.pnas.org by IISER PUNE on October 7, 2024 from IP address 14.139.123.36.