J Stat Phys
https://doi.org/10.1007/s10955-018-2012-2

Collective Action Problem in Heterogeneous Groups with Punishment and Foresight

Logan Perry[1] · **Mahendra Duwal Shrestha**[2] · **Michael D. Vose**[2] · **Sergey Gavrilets**[3]

Received: 28 August 2017 / Accepted: 8 March 2018

Abstract The collective action problem can easily undermine cooperation in groups. Recent work has shown that within-group heterogeneity can under some conditions promote voluntary provisioning of collective goods. Here we generalize this work for the case when individuals can not only contribute to the production of collective goods, but also punish free-riders. To do this, we extend the standard theory by allowing individuals to have limited foresight so they can anticipate actions of their group-mates. For humans, this is a realistic assumption because we possess a "theory of mind". We use agent-based simulations to study collective actions that aim to overcome challenges from nature or win competition with neighboring groups. We contrast the dynamics of collective action in egalitarian and hierarchical groups. We show that foresight allows groups to overcome both the first- and second-order free-rider problems. While foresight increases cooperation, it does not necessarily result in higher payoffs. We show that while between-group conflicts promotes within-group cooperation, the effects of cultural group selection on cooperation are relatively small. Our models

---

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10955-018-2012-2) contains supplementary material, which is available to authorized users.

✉ Sergey Gavrilets
gavrila@tiem.utk.edu

Logan Perry
lperry13@vols.utk.edu

Mahendra Duwal Shrestha
mduwalsh@vols.utk.edu

Michael D. Vose
vose@eecs.utk.edu

1 Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA

2 Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA

3 Department of Ecology and Evolutionary Biology, Department of Mathematics, National Institute for Mathematical and Biological Synthesis, Center for the Dynamics of Social Complexity, University of Tennessee, Knoxville, TN 37996, USA

Published online: 14 March 2018

predict the emergence of a division of labor in which more powerful individuals specialize in punishment while less powerful individuals mostly contribute to the production of collective goods.

Keywords Cooperation · Conflict · Theory of mind · Modeling · Cultural group selection

1 Introduction

Working as a group can be highly profitable due to economy of scale. However, there are several obstacles that must be overcome in order to obtain such a benefit. For example, group members have to be able to effectively coordinate their actions, resolve potential conflicts, and minimize free-riding. Within-group free-riding (which often comes under a rubric of "collective action problem", [71]) can undermine cooperation whenever individual effort is costly and individuals can benefit from the effort of their group-mates. In this case, all individuals are faced with the temptation to reduce or eliminate their efforts. However, if a substantial proportion of group members follow this logic, the group benefit will not be produced and everybody will suffer. Collective action problem is common in many animal and human groups [5,27,35,43,68,71,73,96]. Nevertheless, cooperation in groups is widespread [14,88,95]. Understanding its origins and maintenance is viewed as a fundamental challenge in the fields of both social and biological sciences [5,13,35,43,68,71,73].

Accordingly, an extensive body of work exists that explores the different mechanisms capable of promoting and maintaining cooperation. Classical work in economics emphasizes selective incentives (i.e., benefits and punishment) as well as institutional design as solutions to collective action problems [71,73,90]. More recent work in evolutionary biology focuses on kin selection, direct and indirect reciprocity, group selection and also punishment [5,60]. Adding to this list is a body of literature which explores the influence of within-group heterogeneity [1,20,25,27,29,61]. Intuition suggests, and experimental work has shown, that the costs and benefits of cooperation affect participants' willingness to contribute to a collective good. Additionally, the presence of influential individuals [32,98,100] and differences in the cost of punishing [17,20] can each result in cooperation being more easily established and maintained because some individuals may be more willing to contribute than others. These findings suggest that in addition to being more realistic, the inclusion of within-group heterogeneity is key to understanding how groups cooperate.

Our primary goal here is to extend recent theoretical work on voluntary provisioning of collective goods in heterogeneous groups [27,29] for the case when individuals can punish free-riders. Mathematical models have already shown that punishing can be an effective tool for promoting the evolution of cooperation [7,8,24,66,84]. This conclusion is supported by experimental studies on cooperative behavior in both humans [23,72] and non-human animals [12,63]. Despite this work, the motivations behind why any one individual would choose to punish remains a source of much debate. While punishment may discourage the future selfish behavior of others, it comes with a personal price as the act of administering punishment requires some expenditure of effort. For this reason, it is more beneficial to abstain from punishing in the same way it can be to abstain from contributing to production of a collective good, resulting in a second-order free-rider problem. Proposed solutions to this second- (and higher-) order free-rider problem include meta-punishment [6], conformism [44], signaling [36,86] and group-selection [7,34,88].

Below, we offer a different solution to the second-order free-rider problem. It is motivated by the idea that individuals often punish others in order to modify their future behavior [16]. This is also known as deterrence theory (see [65] for a review) and utilizes what Axelrod [3] referred to as "the shadow of the future." Examples of such punishment abound in modern society, from penal systems to the grounding of one's child. In line with this, laboratory experiments have shown that individuals view punishment and the threat of punishment as way to deter future bad behavior [16,22,92,93]. For example, in an experiment performed by Krasnow et al. [54], participants played a trust game with a partner, who could cooperate or defect. Following this initial decision, participants could punish partners who had defected. Then participants played another round of the trust game with the same partner. Results show that participants were equally trusting of both partners who had previously cooperated and those who had been punished following defection [54]. The participants' willingness to trust those who had been punished indicates their belief that the act of punishing would be effective at preventing future selfish behavior.

The understanding that others will change their actions in response to one's own actions is a consequence of humans' "theory of mind". The latter refers to the ability to reason about the knowledge and thought processes of others and is a well established trait in humans [77,91], and recently was even suggested in great apes [9,55]. Humans can also use the theory of mind recursively [41,76], which entails thinking about how others think about you. This ability is thought to be key in promoting cooperation within groups [91].

Here, we develop and study an agent-based model in which individuals both value future payoffs and are aware of their influence over their group-mates' future actions. We seek to explore the effects such cognitive capabilities have on cooperation in groups facing a collective action problem. Our results suggest that the capability to anticipate future actions of group-mates (i.e. foresight) results in a higher willingness to enforce cooperation via punishment. Hence, foresight effectively provides a solution to the second-order free-rider problem that does not depend on meta-punishment, signaling, conformism, or cultural group selection. While foresight increases cooperation, the ability of foresight does not necessarily lead to higher payoffs. We also show that within-group heterogeneity results in an impromptu division of labor where dominant individuals take charge of punishment and others are forced to focus on contributing to the public good production.

2 The Basic Model

Below we start by specifying a basic set of models of collective action without foresight. We describe the corresponding results for egalitarian and hierarchical groups focusing on the effects of different parameters. Later we will build on these results to show that foresight dramatically changes the resulting social dynamics.

2.1 No Punishment and No Foresight

We will use a common framework for studying collective action [2,27–30,53,94]. We consider a population of individuals living in a large collection of G groups, each with fixed size N. These groups are engaged in collective actions, potentially affecting both individual and group survival and reproduction. Such a set-up results in multi-level selection pressures at both the within- and between-group levels. Pressures at the between-group level encourage individuals to make large contributions, while at the within-group level there is a strong incentive for individuals to free-ride. We allow for groups to be either egalitarian, so that all individuals receive an equal share of produced goods, or hierarchical, so that each individual's share depends on their rank (or strength).

Collective actions. We consider two types of collective action [27,28].The first type focuses on group activities such as hunting and gathering, defense from predators and, building/maintaining shelter. The success of an individual group in these activities is largely unaffected by the action of neighboring groups. For this reason, we refer to these actions as "us vs nature" games. In contrast, limited space, resources, and mating opportunities can result in competition between groups of conspecifics. This means that as the efforts of one group rises (ceteris paribus), the resources available to neighboring groups decrease. Since the success of a particular group depends upon the actions of its neighbors we refer such games as "us vs them". Mathematically, in "us vs nature" actions, group members participate in a generalized public goods game. In "us vs them" actions, two groups compete in a contest [53,80] over a resource. The key distinction between these two types of collective action is that in the former the absolute group effort is critical for obtaining resources, while in the latter it is the group effort relative to that of the other competing group that matters.

Let $x_{i,j}$ be a nonnegative number representing the effort expended by individual i in group j. This contribution comes at a personal cost of $cx_{i,j}$, where c is a constant parameter, that is, we assume the cost grows linearly with the effort. The combined efforts of the group members, i.e., the impact function, is $X_j = \sum_i x_{i,j}$. In "us vs nature" games, the maximum available benefit is B and the share obtained by group j is defined as

$$P_j = \frac{X_j}{X_j + X_0}, \tag{1}$$

where X_0 is the half-success effort, i.e. the group proficiency required for $P_j = 50\%$. The larger X_0, the more group effort is required to secure the reward. For example, when a resource is abundant, less effort (resulting from fewer contributors or smaller average contribution) would be required to obtain a sufficient amount. However, when a resource is rare, a larger expenditure of effort may be necessary to successfully seek it out.

In "us vs them" games, the maximum available benefit is $2B$, and the share won by group j when in competition with group k is

$$P_j = \frac{X_j}{X_j + X_k}. \tag{2}$$

This can also be interpreted as the proportion of fights where group j is able to beat group k.

Individual payoffs. Individual payoffs are defined as

$$\Pi_{ij} = 1 + \tilde{B} P_j v_i - c x_{ij}, \tag{3}$$

where $\tilde{B} = B$ in "us vs nature" games and $\tilde{B} = 2B$ in "us vs them" games, and v_i is the proportion of the reward allotted to individual i (or its valuation). In egalitarian groups, everybody gets an equal share, $v_i = 1/n$. We will also consider heterogeneous groups in which individuals will differ in v_i. It is useful to define the term $b = B/n$, which is the expected benefit per individual in an egalitarian group playing an "us vs them" game against a group making an equally strong effort. In presenting our results, we will use the relative payoffs $\pi_{ij} = \Pi_{ij} / \sum_j \Pi_{ij}$.

Events. To increase the realism of our models, we formulate them in continuous time and model social dynamics using an event-driven approach [33]. [We note also that the more commonly used synchronous updating may result in some spurious effects (e.g. cycling) which is not desirable.] There are three types of events. At each time step, with probability

E_1 a randomly chosen group participates in an "us vs. nature" game, while with probability E_2 a randomly chosen pair of groups participate in an "us vs. them" game. Lastly, with probability E_3 a randomly chosen poorly performing group copies the strategies of another randomly chosen, but well-performing group. When such a "cultural group selection" [79] event occurs, the chances to be chosen to copy another group are proportional to e^{-P_j} and those of being copied are proportional to e^{P_j}. We treat probabilities E_1, E_2, and E_3 as exogenously specified constants. Naturally, $E_1 + E_2 + E_3 = 1$. It will be useful to define the ratio $E = E_1 : E_2$ of the frequencies of "us vs. nature" to "us vs. them" games (which we will refer to as the "game frequencies ratio") for when we begin to explore the effects the frequencies of these events have on group dynamics.

Strategy revision. To increase the realism of our models, we also assume that our individuals are bounded rational (rather than, say, hard-wired genetically to behave in a particular way). After both "us vs nature" and "us vs them" events, each member of the affected group(s) independently updates their strategies with probability μ. We start by using myopic optimization (best response) with a finite number of "candidate strategies". [We note that myopic optimization and genetic selection often lead to very similar equilibria (e.g., if the underlying game is potential [46,99].)] For example, assume that the action taken by individual i was contributing effort $x_{i,0}$ to a collective action. Candidate strategies are randomly and independently drawn from a normal distribution centered at an individual's current strategy $x_{i,0}$ and with a standard deviation of σ. We draw K new "candidate strategies" $x_{i,1}, \ldots, x_{i,K}$ and then evaluate the associated $K + 1$ expected payoffs $\Pi_{i,k}$ (including the payoff for the original $x_{i,0}$ strategy), assuming that all other individuals keep their strategies from the prior round. That is, members of groups that just played consider how each of their candidate strategies would have fared in the event that just took place. Being able to assess this requires each individual to know their own effort as well as the total effort of their group (and of the competing group).

From these $K + 1$ strategies, individual i will choose one at random with probabilities proportional to $e^{\lambda \Pi_{i,k}}$, where λ is a parameter measuring the precision in evaluating the payoffs. Increasing λ increases the probability that a candidate strategy with the highest payoffs is chosen. This approach is a version of the Quantal Response Equilibrium (QRE) developed in the field of economics and serves as a way to incorporate mistakes players make when comparing potential strategies [39,62]. In our models, the standard deviation in the candidate strategies' expected payoffs is relatively small, which necessitates the use of relatively large λ values. In particular, we consider such values as $\lambda = 10, 20, 40, 80$. In order to gain additional insight, we also examine the case when $\lambda = \infty$, which corresponds to when agents make no errors always choosing myopically the best strategy available to them.

To study our model we use agent-based simulations (more details are provided in the Supplementary Information, SI). Although similar to that in Gavrilets and Fortunato [29] and Gavrilets [27], our approach differs in several important aspects. First, we assume that events happen asynchronously. Second, we allow for groups to play both types of games rather than a single type of game. Our approach introduces new parameters—the frequencies of different events E_1, E_2 and E_3. Third, we focus on evolution by myopic optimization instead of genetic evolution. This means that strategies do not evolve as a consequence of random mutation and natural selection, but rather individuals choose strategies attempting to maximize their payoffs. Our approach also introduces parameters λ, agents' precision in evaluating potential payoffs, and K, the number of new candidate strategies considered. Table 1 summarizes variables, functions and parameters for the model we introduce and study here.

Table 1 Model variables, functions and parameters

Symbols	Their meaning	Numerical values
x_i	Individual production effort ($x_i \geq 0$)	
y_i	Individual punishment threshold ($y_i \geq 0$); i punishes j if jth effort is too low, i.e., if $x_j < y_i$, and i is sufficiently strong relative to j, i.e., $s_i > s_j + \xi$, where $\xi \sim N(0, \varepsilon)$.	
X	Total group effort, $X = \sum x$	
$P(X)$	Normalized value of the resource produced or secured by the group: $P = X/(X + X_0)$ in "us vs. nature" games; $P = X/\overline{X}$ in "us vs. them" games	
v_i	Share of the reward going to i; depends on relative strength s_i: $v_i = s_i^{\beta} / \sum s_j^{\beta}$	
Π_i	Expected payoff; e.g. expected payoff from production is $\Pi_i = 1 + bP(X)v_i - cx_i$	
κ_{ij}	Cost to i of being punished by j, $\kappa_{ij} = e(y_i - x_j)$ ("graduated punishment")	
δ_{ij}	Cost to i of punishing j, $\delta_{ij} = \kappa_{ij} S_{ij}$ with $S_{ij} = S_0 \exp[\phi(s_j - s_i)]$, so that punishing a weaker opponent is relatively cheap	
U_i	Utility with foresight, $U = (1 - \omega)\Pi_i^{(1)} + \omega\Pi_i^{(2)}$	
n	Group size	4, 8, 16
b	Benefit parameter for production	0.0 : 0.25 : 2
c	Cost parameter for production	0.5, 1.0
X_0	Half-effort parameter in "us vs. nature" games	1
E	The ratio $E = E_1 : E_2$ of probabilities of participation in "us vs. nature" and "us vs. them" events	1:0, 3:1, 1:1, 1:3,0:1
E_3	Probability that the group participates in "cultural group selection" events	0.0,0.05,0.1,0.15,0.2
s_i	Individual strength, $s_i \sim U[0, 1]$	
ε	Error in strength evaluation	0.1
β	Degree of inequality; the egalitarian case corresponds to $\beta = 0$ and $v_i = 1/n$	0, 1, 2, 4
e, ϕ, S_0	Cost of punishment parameters	1, 5, 1
ω	Weight of future payoffs in the utility function	0, 0.25, 0.5
λ	Precision parameter in the QRE approach in which the probability of choosing a candidate strategy k is proportional to $\exp(\lambda\Pi_k)$ or $\exp(\lambda U_k)$	10, 20, 40, 80, ∞
K	Number of candidate strategies	1, 2, 4, 8, 16
σ	Standard deviation in innovation ($x' = x + N(0, \sigma)$, $y' = y + N(0, \sigma)$)	0.5
μ	Probability of strategy updating	0.5

Results for egalitarian groups. First we describe what happens in the case of equal division of the group reward. From earlier work focusing on games of one type only, we know that in "us vs. them" games, the equilibrium group contribution $X^* = b/(2c)$ and is always positive.

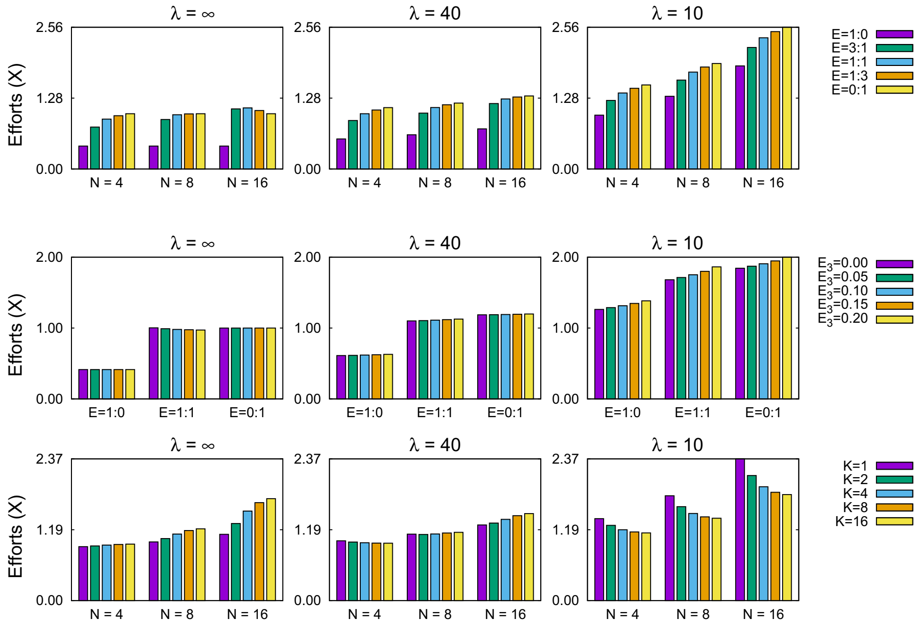

Fig. 1 Group efforts X in the egalitarian case as influenced by group size N, the precision parameter λ, the game frequencies ratio E (first row of graphs, with $E_3 = 0.1$, $K = 2$), the frequency E_3 of cultural group selection events (second row of graphs, with $N = 8$, $K = 2$), and the number of candidate strategies K (third row of graphs, with $E = 1 : 1$, $E_3 = 0.1$) Results are the averages of 20 simulations using: $b = 1.0$, $c = 0.5$ and $X_0 = 1.0$

In "us vs. nature" games, $X^* = X_0(\sqrt{R} - 1)$, where $R = b/(cX_0)$, so that X^* is positive only if the benefit-to-cost ratio R is big enough [27–29]. Here to better elucidate the impact of the parameters E, λ, and K we have chosen the values of b, c and X_0 so that contributions are always positive (see the SI for more results). All graphs below report the average values evaluated over a certain number of time steps at the end of the simulations.

Figure 1 illustrates the effects on the total group effort X of precision λ and additional parameters. [Note the meaning of each color differs between the rows.] In this figure, the graphs (from left to right) correspond to the cases of $\lambda = \infty$, 40, 10 meaning the precision in evaluating payoffs decreases as we move to the right. We see that decreasing precision results in increased contributions. This happens because errors in strategy evaluation bias individual contributions towards over-production (due to the fact that negative contributions are not possible). We also see moderate increases in the total group effort in response to group size N. This is because in larger groups, more individuals make errors. The first row of graphs in Fig. 1 shows that increasing the frequency of "us vs. them" events (specified by the game frequency ratio E) increases the total group effort. This is in line with previous findings [27,28]. The second row of graphs in Fig. 1 shows that the frequency of cultural group selection events E_3 affects the efforts only for small λ but the overall effects are small. This implies that in this model, the effects of cultural group selection on cooperation are insignificant.

The third row of graphs in Fig. 1 illustrates the impact of the number of candidate strategies K. When individuals are imprecise ($\lambda < \infty$) we see that K has a nonlinear effect: increasing K increases the group effort with small errors ($\lambda = 40$) but decreases it with large errors ($\lambda = 10$). The case when individuals make no errors ($\lambda = \infty$) is more nuanced and depends

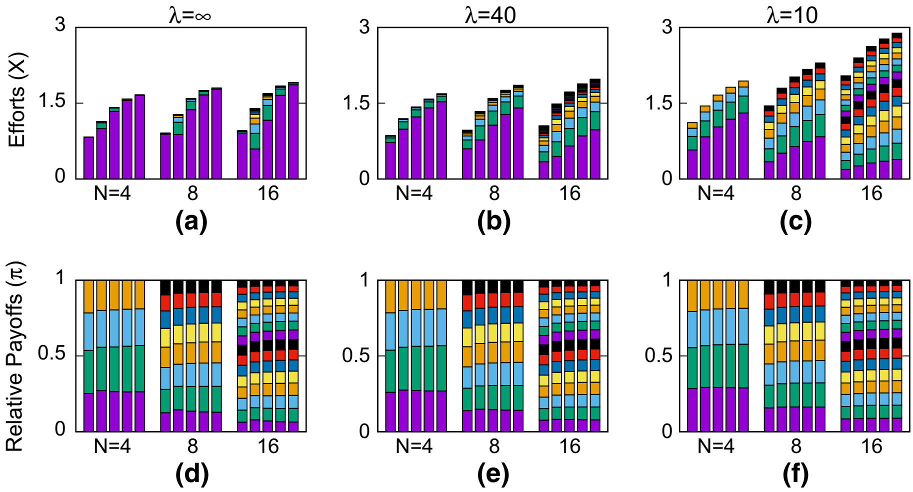

Fig. 2 Effects of group size N, the precision parameter λ, and the game frequencies ratio E on the group efforts X and relative payoffs π of individuals of different ranks in the basic hierarchical model. For each group size, the bars (from left to right) correspond to $E = 1:0, 3:1, 1:1, 1:3, 0:1$. The segments of each bar correspond to a particular individual with the strongest at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $b = 1.0$, $\beta = 1.0$, $c = 0.5$, $X_0 = 1$, $K = 2$ and $E_3 = 0.1$ (Color figure online)

upon the games frequencies ratio E. When groups only participate in "us vs. nature games ($E = 1:0$), we see that K has no effect. This is to be expected as "best responders" are expected to converge on the optimal contribution regardless of how many candidate strategies they evaluate. However, when competition with neighboring groups is introduced, increasing K unexpectedly increases group efforts. The intuition behind this effect is that with groups constantly facing different opponents, "fine tuning" of individual behavior within each group (expected with large K) prevents convergence to an equilibrium. As a result, the system exhibits nonlinear dynamics (see Figs. S2 and S3 in the SI).

Results with within-group heterogeneity. In real life, individuals are not identical in their strengths and abilities and this heterogeneity impacts social behavior. For example, stronger individuals can grab a large share of the reward or can make larger efforts because it costs them less. In our model, we postulate that group members differ in their strength s_i. An individual's strength is a constant assigned by drawing a random number from the uniform distribution over [0, 1]. We rank individuals according to their strengths. Following Gavrilets and Fortunato [29], we set individual shares v_i (or valuations of the reward) based on their relative strengths:

$$v_i = \frac{s_i^{\beta}}{\sum_k s_k^{\beta}}, \tag{4}$$

where parameter β determines the steepness of the hierarchy, i.e. the degree of inequality within the group [29]. Note if $\beta = 0$, then all shares are identical (as in the egalitarian case above), but if $\beta > 1$ stronger individuals obtain disproportionately larger shares of resources.

Figure 2 is analogous to the first row of Fig. 1 but now we show efforts of individuals of different ranks (averaged across all groups) as well as their relative payoffs π_i (i.e. shares of the reward) using different colors (the strongest individuals are at the bottom of each column). We see that heterogeneity results in the most dominant individual contributing almost all of

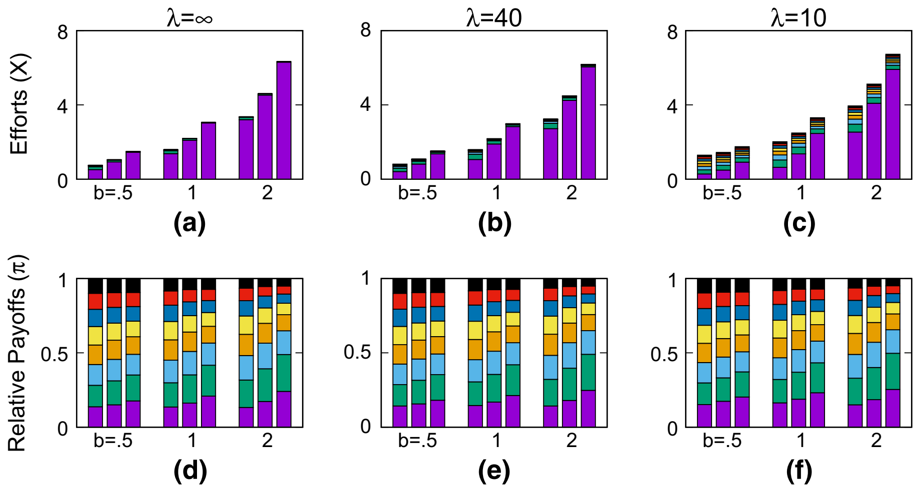

Fig. 3 Effects of the average benefit b, the precision parameter λ, and the hierarchy steepness β on the average group effort X and relative payoffs π of individuals of different ranks in the basic hierarchical model. For each value of b, the bars (from left to right) correspond to $\beta = 1, 2, 4$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $N = 8$, $c = 0.5$, $K = 2$, $E = 1 : 1$ and $E_3 = 0.1$ (Color figure online)

the effort while the rest of the group free-rides. The severity to which the rest of the group free-rides off of the efforts of the dominant individual increases in response to increased precision λ. This pattern is expected as more rational subordinate individuals would seek to fully capitalize on the dominant individual's willingness to contribute. This is an example of Olson's (1965) "exploitation of the great by the small". In some cases, the most dominant individual ends up with a smaller payoff then one or more of its subordinates in spite of receiving the biggest share of the reward, a pattern named an "altruistic bully" effect [29]. This can happen regardless of the group size N or precision parameter λ. Contributions by the dominant are driven by competition with other groups since increases in the frequency E of "us vs them" events increase efforts in all cases. Decreasing precision λ increases individual contributions of non-dominant individuals who now contribute to collective action by mistake (see also Fig S4 in the SI). In contrast, with higher precisions dominant individuals bear an even larger share of the group's burden (Fig. 2a–c). The total group effort does not depends on the group size when precision parameter λ is large and weakly increases with low λ as more individuals contribute by error in larger groups. Increasing the number of candidate strategies K results in individuals contributing less effort towards the common good (Fig. S5 in the SI). The effects of the frequency E_3 of cultural group selection events in the hierarchical model are similar to those in the egalitarian case and small (see Fig. S6 in the SI).

Figure 3 focuses on the effects of benefit b and the parameter β, which measures the steepness of a group's hierarchy, i.e. the inequality present amongst group-mates. We see increasing inequality β results in increasing total group effort [27,29] with dominant individuals contributing most of it.

2.2 Incorporating Punishment

Next we add the possibility of punishment. We utilize graduated punishment [26,42,48, 83]) as it represents the most appropriate approach for dealing with continuous production efforts (rather than discrete). Specifically, we postulate that each individual has an evolving "punishment threshold" trait y_i and is motivated to punish any other group-mate whose effort x_j is smaller than y_i. If individual i punishes individual j, j's payoff is reduced by

$$\kappa_{ij} = e(y_i - x_j), \tag{5}$$

where $e > 0$ is a scaling parameter. That is, the smaller the effort x_j of j relative to the punishment threshold y_i of i, the larger the cost κ_{ij}. We assume the act of punishing others is costly. We denote the corresponding payoff loss to i as δ_{ij}.

In standard models of graduated punishment, any individual can punish any other individual. In real life however, individuals vary greatly in their strengths and capabilities, and while strong individuals can often easily punish weak individuals, punishment of the strong by the weak is unlikely. To capture this idea, we assume that individual i is willing to punish individual j only if

$$s_i > s_j + \xi, \tag{6}$$

where ξ is a random number from a normal distribution with zero mean and standard deviation ϵ. One way to interpret this model is that i is willing to punish j only if i believes it is stronger than j with ξ being the error in evaluating relative strengths.

Moreover, we postulate that the cost to punisher i depends on the severity of punishment κ_{ij} and the difference in strengths, $s_i - s_j$:

$$\delta_{ij} = \kappa_{ij} S_{ij}, \tag{7}$$

where

$$S_{ij} = S_0 \exp[\phi(s_j - s_i)]. \tag{8}$$

Here S_0 is the baseline factor of punishing (equal to the cost of punishing an individual equal in strength) and ϕ is a scaling parameter (the larger ϕ, the faster the cost to i declines with the difference in strengths). This formulation implies that punishing of the weak by the strong is relatively cheap.

In this extension of our basic model, each individual is characterized by a pair of strategies (x_i, y_i) specifying their contribution to the collective action and punishment threshold. Correspondingly, in updating individual strategies, agents will consider K pairs of randomly generated candidate strategies $(x_{i,1}, y_{i,1}), \ldots, (x_{i,K}, y_{i,K})$ plus the original strategy pair (x_i, y_i), and select one of them with probabilities proportional to their expected payoffs as we had done previously in Sect. 2.1.

Our numerical results for this model show that punishment can happen only by error when individuals have low levels of precision λ (see below). This makes sense because of the second-order free-riding discussed above. Punishment due to error is a tenuous route to overcoming the second-order free-rider problem. Instead of depending on agents mistakenly enforcing cooperation as a result of their inability to differentiate the best payoff, we seek to develop an explanation for how both imprecise and precise agents can overcome the second-order free-rider problem. We accomplish this through the next addition to our model, which will incentivize agents to enforce cooperation despite the costs.

3 The Model with Foresight

3.1 Strategy Revision with Foresight

We now extend our strategy revision protocol to make use of foresight. Let (x_i', y_i') be a candidate strategy of i. Let $\Pi_i^{(1)}$ be the expected payoff to i at the next collective action if i chooses (x_i', y_i') while all its group-mates keep their current strategies. Individual i may expect that its group-mate j will chose a new strategy $(\tilde{x}_j, \tilde{y}_j)$ in response to i using (x_i', y_i'). Predicting these strategies for all group-mates allows i to estimate the expected payoff $\Pi_i^{(2)}$ at the collective action two steps forward. In the standard myopic optimization (or the "best response" approach) which we have used so far, i attempts to maximize $\Pi_i^{(1)}$. In contrast, we will now postulate myopic optimization with one-step foresight where i attempts to maximize the utility function

$$\Pi_f = (1 - \omega)\Pi_i^{(1)} + \omega\Pi_i^{(2)}, \tag{9}$$

where the foresight parameter $0 \leq \omega \leq 1$ measures the weight of the future payoff in the utility function. We note that predicting how group-mates will react to a change in one's strategy should not be too difficult cognitively. Basically, one just has to answer the simple question "what would I do if I were in their place?" In numerical implementation, i generates $K + 1$ pairs of candidate strategies for each j and attempts to picks the best one from the j's perspective. The actual choice happens with probabilities proportional to $\exp(\lambda\Pi_f)$ as before.

Note that in our implementation, individuals only seek to predict their potential influence over members of their own group. In particular, agents that participate in "us vs them" events do not attempt to predict how the opposing group's agents will alter their future behavior. This is a reasonable assumption because of the lack of familiarity between members of different groups, and the fact that the two groups are not guaranteed to interact again.

3.2 Results for the Model with Foresight

We begin our investigation of the full model by first considering the effects of the foresight parameter ω and the game frequency ratio E on individual and group behavior. Figure 4 illustrates these effects for groups with perfect precision ($\lambda = \infty$, first row) and imperfect precision ($\lambda = 40$, second row). In addition to the efforts and relative payoffs, we now show the payoff loss due to being punished (labeled "punishment incurred" q in the graphs) and the payoff lost by other individuals due to the punishment by the focal individual (labeled "punishment inflicted" p in the graphs). Formally, for individual i, the punishment incurred $q_i = \sum_j \kappa_{ki}$ while the punishment inflicted $p_i = \sum_j \kappa_{ik}$. (See the SI for figures summarizing punishment thresholds y_i and the costs of punishing others δ_{ij}.)

The case of $\omega = 0$ corresponds to cooperation and punishment in groups without foresight as discussed above. In this case, most of the group effort is delivered by the most dominant individual and punishment happens only by error. We see foresight significantly increases the total group effort and that this happens in spite of the most dominant individuals greatly decreasing their efforts (Fig. 4a, e). Rather than contributing to production of the collective goods, the dominant individuals dramatically increase punishment delivering most of it (Fig. 4c, g). Punishment is mostly directed towards weak individuals (Fig. 4d, h) who greatly increase their contributions (Fig. 4a, e). As a result of these factors, dominant individuals increase their payoffs at the expense of subordinates (Fig. 4b, f). Our results thus suggest

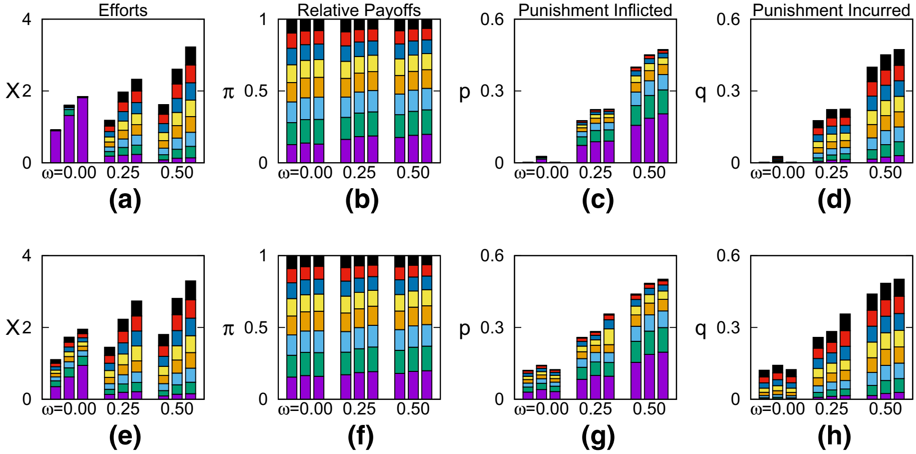

Fig. 4 Effects of the foresight parameter ω and the game frequencies ratio E on the group efforts X and relative payoffs π, the punishment inflicted p, and the punishment incurred q for individuals of different ranks in the full model with perfect ($\lambda = \infty$, first row) and imperfect ($\lambda = 40$, second row) precision. For each value of ω, the bars (from left to right) correspond to $E = 1:0, 1:1, 0:1$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $N = 8$, $b = 1.0$, $c = 0.5$, $\beta = 1$, $K = 2$ and $E_3 = 0.10$ (Color figure online)

that foresight does indeed provide a viable route to overcoming the second-order free-rider problem.

In general, dominant individuals make lower contributions and obtain larger proportions of the reward than their group-mates, but these asymmetries are lessened as group size increases (see the SI). In all cases, foresight greatly increases the efforts of subordinate group members (Fig. 4a, e) because of increased punishment of free-riders. Foresight results in dominant members severely decreasing their contributions toward production because punishing is cheaper to these individuals, and hence it is easier to increase their group's total output by promoting cooperation of others than it is to do so by increasing their own contributions. This results in a division of labor, where the dominant individual focuses on promoting cooperation through punishment while the rest of the group contributes to the common good. In this way the dominant individual acts as a coercive leader for the group.

As in the basic model, we see that increased competition with other groups (i.e. increased E) results in a higher group efforts. However, because of the aforementioned reasons, these increases are not due to the dominant individual contributing more, but instead a result of cooperation being more strictly enforced (Fig. 4a, c, e, g). Hence, the extent of a group's division of labor is driven by both the emphasis its members place on future payoffs and the amount of competition they face from other groups. These results imply that the emergence of a coercive leader would be more likely in rational groups that commonly engage in conflicts with neighboring groups.

Another factor that determines the extent of the division of labor is the number of candidate strategies generated, K. Figure 5 shows that for the case of $N = 8$ as the number of candidate strategies an individual considers grows, dominant individuals contribute less and punish more (see the SI for additional numerical results). The fact that the extent of the observed division of labor increases with the number K of candidate strategies considered implies that the option to develop such a division is payoff-maximizing.

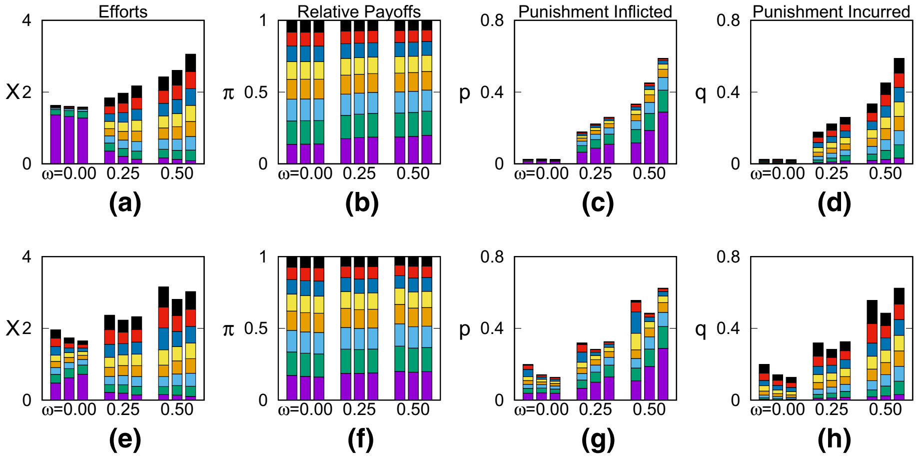

Fig. 5 Effects of the foresight parameter ω and the number of candidate strategies K on the group efforts X and relative payoffs π, the punishment inflicted p, and the punishment incurred q for individuals of different ranks in the full model with perfect ($\lambda = \infty$, first row) and imperfect ($\lambda = 40$, second row) precision. For each value of ω, the bars (from left to right) correspond to $K = 1, 2, 4$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $N = 8$, $b = 1.0$, $c = 0.5$, $E = 1:1$, and $E_3 = 0.1$ (Color figure online)

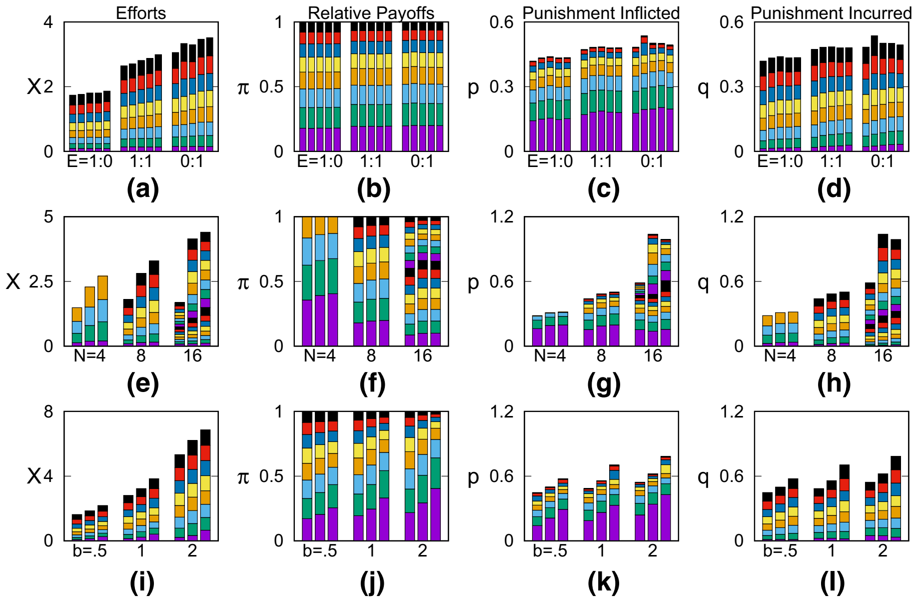

Fig. 6 Further results on punishment with foresight. First row of graphs: effects of the game frequencies ratio E and the frequency E_3 of cultural group selection events. For each value of E, the bars (from left to right) correspond to $E_3 = 0.00, 0.05, 0.10, 0.15, 0.20$. Second row of graphs: effects of the group size N and the game frequencies ratio E. For each value of N, the bars (from left to right) correspond to $E = 1:0, 1:1, 0:1$. Third row of graphs: effects of the average benefit b and the hierarchy steepness β. For each value of b, the bars (from left to right) correspond to $\beta = 1, 2, 4$. Unless specified otherwise: $N = 8$, $b = 1.0$, $c = 0.5$, $\beta = 1.0$, $K = 2$, $\omega = 0.5$ and $\lambda = 40$. Results are the averages of 20 simulations

Figure 6 provides additional information on the effects of different parameters with $\omega = 0.5$ and $\lambda = 40$. The first row of graphs in this figure explores the effects on the game frequencies ratio E and the frequency of cultural group selection events E_3. Increasing E increases both contributions and punishment. Effects of E_3 are somewhat similar but relatively small. We see that weaker individuals increase their contributions towards the group's good because dominant individuals punish more with increasing E_3. The second row of graphs in Fig. 6 illustrates the effects of the group size N. We see that increasing N increases the amount of payoff spent on punishment and also decreases the payoff of dominant individuals. Our results correspond to the well-established fact that group cohesion is more difficult to maintain as group sizes increase. The third row of graphs in Fig. 6 explores the effects on the benefit b and the hierarchy steepness β. Increasing either of these parameters increases punishment and contributions as expected.

Figures S8–S10 in the SI provide additional numerical results. Among other things they show that although foresight increases production of collective goods, its effect on the absolute payoffs Π can be negative (due to the costs of punishment).

4 Discussion

Here we have studied the impact of punishment on the ability of heterogeneous groups to overcome collective action problems. The latter seriously undermine cooperation across a wide range of situations in many animal and human groups as demonstrated by numerous studies in primatology [15,52,68,69,96,97], behavioral biology [4,56,57,74], anthropology [10,11,38,40,58,59], and economics [71,73,90]. Earlier work has shown that within-group heterogeneity and inequality can, under some conditions, increase the group's productivity as dominant or stronger individuals make larger effort overcompensating free-riding of subordinate or weak group-mates [27,29]. Earlier work has also established that punishment is a powerful force for the *maintenance* of cooperation [7,8,24,66,84]. However the role of punishment in the *origin* of cooperation is less clear, partially because of the second-order free-rider problem [6,44].

Earlier models were studied under the assumptions that individuals are genetically hardwired to behave in a particular way or that they use selective imitation or myopic optimization to modifying their strategy [82]. The first approach is hardly realistic (at least in the case of humans), while the last two run into the second order free-rider problem. Selective imitation would also be difficult to use if groups are heterogeneous as a strategy successful for somebody would not necessarily be good for everybody else. These challenges forced us to look for alternative approaches.

Our novel solution focuses on a one-step foresight where players, first, care about their future payoff and, second, attempt to account for how their group-mates will react to their new strategies. These are very intuitive assumptions. Our main result is that inclusion of a limited foresight dramatically changes groups' social dynamics. First, it greatly simplifies cooperation by making punishment an utility increasing action. This is a robust result, which is not expected to crucially depend on modeling details such as the presence of heterogeneity in strengths or the specific form of punishment used. Second, foresight can result in a division of labor in which more dominant or stronger individuals specialize in enforcing cooperation through punishment while the rest of the group focuses on production of collective goods. Lastly, this route to promoting cooperation works for a wide range of the number of candidate strategies K and precision λ. The potency of this new method of strategy revision was made clear by the examining the effect of the parameter ω, which measures the emphasis individuals place on forecasted payoffs. We saw that as the importance of future payoffs increased, individuals became more willing to punish uncooperative group-mates and the extent of the division of labor deepened. While previous agent-based simulations have shown that possessing theory of mind is beneficial when participating in dyadic games [18,19], our work serves as one of the first steps towards understanding how it affects behavior in a group setting. Our findings suggest that a theory of mind is an important aspect of how groups are able to solve collective action problems.

In our paper, we started by accepting the existence of foresight as given and then explored its evolutionary consequences. An interesting and important question is how it evolved in the first place. In principle, foresight could have evolved for a variety of reasons, e.g. because it increased individual performance in hunting, building shelters, making fire, mating, increasing dominance status, or in coalition building. Exploring these possibilities is outside of the scope of our paper.

Earlier work has already emphasized the importance of strategy revision protocols in evolutionary dynamics [82]. While the first and second-order free-rider problems exist independently of the strategy revision protocol, people's decision-making process can affect the ability of groups to overcome these problems. For example, selective imitation is not able to solve the first or second-order free-rider problems without invoking additional mechanisms (e.g., group selection or reduced migration and inbreeding). In contrast, foresight is able to overcoming the second-order free-rider problem as punishment gets established and hence cooperation increases.

Our approach for making predictions of how group-mates will act is related to the so-called "level-k models" [64,89]. The latter assume players adhere to a particular model of reasoning which forms a hierarchy. At the bottom are those individuals who change their behavior completely randomly, referred to as level-0 types. If a player assumes all others adhere to level-0 reasoning and adapt his or her behavior accordingly, then they are categorized as level-1 type. Similarly, a level-2 type believes all others to be either level-0 or level-1. These levels of reasoning continue indefinitely. In our model, level-0 individuals do not change their strategies and level-1 types use standard myopic optimization. Our limited foresight updating then corresponds to level-2 reasoning types as they consider how others will react to their actions. This assumption is both quite natural and not too taxing on the players' cognitive abilities.

Here, our groups were faced with both "us vs nature" events in which a single group was faced with the task of producing a collective good, and "us vs them" events which saw two groups competing over goods. We saw that the type of collective action a group was participating in greatly affected the strategies group members chose. In particular, we saw that both production and punishment rose in response to the frequency of "us vs them" events as observed earlier [27,28]. Cooperation and punishment also increased with increasing the intensity of cultural group selection (measured in our model by parameter E_3) but the overall effect was small. Our models show that groups of rational self-interested individuals capable of limited foresight can converge on cooperative solutions even without cultural group selection (cf. [85]). Such a convergence will also happen much faster than if by cultural group selection as it does not require information exchange between groups.

The division of labor observed in our simulations is a result of the heterogeneity of our groups. Individuals in our model were assumed to differ in strengths, which impacted their shares of the reward, ability to punish others, and resiliency against punishments. These differences resulted in dominant or stronger individuals being able to more easily exert influence over their peers through punishment. The inability to efficiently retaliate against the punishment inflicted by the strong left the weaker group members with no choice but to increase their production efforts. [Otherwise, retaliation can easily destroy cooperation [67].] Hence, through the use of the threat of punishment, dominant individuals were able to enforce cooperation within their groups. Since the dominant individuals did not fear being punished and maintained cooperation largely through the threat of punishment alone, we saw them have disproportionately large payoffs. While the models without punishment predict "exploitation of the great by the small" [27,29,71], those with punishment depict "the exploitation of the small by the great".

Of particular interest are the trends noted in the payoffs of individuals. We saw that foresight coupled with heterogeneity resulted in dominant individuals reaping larger shares of rewards. The benefit to the strong provided by an ability of foresight was shown to be undermined by larger group sizes. Dominant individuals' share of the payoffs also increased when "us vs them" events became more common, the number of candidate strategies increased, or agents became more precise in their strategy selection. Also as one would expect, dominant individuals obtained larger portions of the groups' benefits when hierarchies were steeper.

The group sizes used in our simulations ($N = 4, 8,$ and 16) cover the range of the number of adult males observed in chimpanzees and extant hunter-gatherer groups [45,75]. Our model applies to situations where adult males pay the costs of contributing to collective actions but then share the benefits they have secured with their own females and offspring. [For males the share of benefits to go to their women and children are a part of their utility.] The emergence of a division of labor was highly sensitive to group size. Increases in group size resulted in growing costs of punishment and in individuals being less prone to punish group-mates. This in turn resulted in decreases in group efforts and in payoffs being more evenly distributed (since it was no longer cost effective for dominant individuals to specialize in enforcing cooperation). Large group sizes make cooperation more difficult also in many related models [27,29,30].

Two additional parameters of our models the expected benefit b and the hierarchy steepness β. As one would expect, increases in b resulted in increased efforts and punishment thresholds. While in the basic model without foresight increasing β resulted in the efforts of dominant individuals rising while their relative payoffs decreased, in the full model we observed increases in both the dominant individuals' thresholds and relative payoffs. Hence, the presence of foresight effectively removed the "altruistic bully" effect observed by Gavrilets and Fortunato [29]. Here, we observed that increasing within-group inequality increased group's production. In general, however, this is expected to depend on modeling details [27].

The models we have developed have more direct implications towards the theory of leadership [31,37,47,51,87]. In particular, one of the roles leaders often take is the role of punisher of defectors in collective actions [47,70]. Experimental evidence has shown that inter-group conflicts increases the punishment of defectors [81] and that when individuals are heterogeneous in the cost of administering punishment, the punishment is carried out by those with the lowest cost [21,78]. Our models show that these behaviors naturally emerge if there is within-group heterogeneity and foresight. Impromptu leaders are more likely to emerge in groups that experience conflict with neighbors frequently. These results are supported by earlier experimental findings [81].

There are several important directions the current model may be extended. First, additional heterogeneity between individuals could be introduced into the precision of individuals and the number of candidate strategies they can generate. A possible route would be to associate the precision of an individual to their strength or payoffs. This extension would mean that individuals who previously obtained greater amounts of a good have the luxury to more fully consider the consequences of their actions. Second, here we assumed that all individuals updated their strategies using the same protocol. This could be contrasted with the case where individuals use different strategy revision protocols. For example, selective copying could result in groups where the extent to which individuals enforce cooperation is more equal. Third, the weight of the level-2 expected payoffs, ω, could be taken to be a dynamic trait. This means that as rounds progress individuals could alter the emphasis they place on future payoffs. Fourth, we studied only one-step foresight generalizing myopic optimization. However, individuals may be using higher-order theories of mind and care about longer term payoffs [18,19,49,50]. Models capturing these features would be much more complex, but potentially more realistic. Finally, we assumed that individuals attempt to maximize their material payoff and neglected any effects of past history. In reality, individuals are often motivated by social norms and normative values [30] and their current actions may be dependent on what happened to them or their groups in the past [94]. All these are important factors that need to be considered in future work.

As societies continue to develop, the importance of cooperation steadily grows. However, the intricacies of human motivations leave collective actions a deep and complex area of research. In particular, balancing immediate versus delayed gratifications—undertaken by all people daily—can be especially nuanced. By incorporating foresight into our model, we have taken steps towards better understanding how our desires for the future shape the society in which we live. The approach presented here is built upon concepts fundamental to group living, and hence will serve as an integral part of completing our understanding of collective action problems.

Acknowledgements We thank K. Rooker and anonymous reviewers for comments and suggestions. Supported by the U. S. Army Research Office Grants W911NF-14-1-0637 and W911NF-17-1-0150, the National Institute for Mathematical and Biological Synthesis through NSF Award #EF-0830858, with additional support from The University of Tennessee, Knoxville, and by the NIH Grant GM56693.

References

1. Andreoni, J.: Privately provided public goods in a large economy: the limits of altruism. J. Public Econ. **35**(1), 57–73 (1988)
2. Archetti, M., Scheuring, I.: Coexistence of cooperation and defection in public goods games. Evolution **65**(4), 1140–1148 (2011)
3. Axelrod, R.: The Evolution of Cooperation. Basic Books, New York (1984)
4. Bonanni, R., Valsecchi, P., Natoli, E.: Pattern of individual participation and cheating in conflicts between groups of free-ranging dogs. Anim. Behav. **79**(4), 957–968 (2010)
5. Bowles, S., Gintis, H.: A Cooperative Species: Human Reciprocity and Its Evolution. Princeton University Press, Princeton (2011)
6. Boyd, R., Richerson, P.J.: Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethol. Sociobiol. **13**(3), 171–195 (1992)
7. Boyd, R., Gintis, H., Bowles, S., Richerson, P.J.: The evolution of altruistic punishment. Proc. Natl Acad. Sci. U.S.A. **100**(6), 3531–3535 (2003)
8. Brandt, H., Hauert, C., Sigmund, K.: Punishment and reputation in spatial public goods games. Proc. R. Soc. Lond. B Biol. Sci. **270**(1519), 1099–1104 (2003)
9. Call, J., Tomasello, M.: Does the chimpanzee have a theory of mind? 30 years later. Trends Cogn. Sci. **12**(5), 187–192 (2008)
10. Carballo, D.M.: Cooperation and Collective Action: Archaeological Perspectives. University Press of Colorado, Boulder (2012)
11. Carballo, D.M., Roscoe, P., Feinman, G.M.: Cooperation and collective action in the cultural evolution of complex societies. J. Archaeol. Method Theory **21**(1), 98–133 (2014)
12. Clutton-Brock, T.H., Parker, G.A.: Punishment in animal societies. Nature **373**(6511), 209–216 (1995)
13. Colman, A.M.: Game theory and Its Applications in the Social and Biological Sciences. Butterworth-Heinemann, London, United Kingdom (2013)

14. Crespi, B.J.: The evolution of social behavior in microorganisms. Trends Ecol. Evol. **16**(4), 178–183 (2001)
15. Crofoot, M.C., Gilby, I.C.: Cheating monkeys undermine group strength in enemy territory. Proc. Natl Acad. Sci. U.S.A. **109**(2), 501–505 (2012)
16. Cushman, F.: Punishment in humans: from intuitions to institutions. Philos. Compass **10**(2), 117–133 (2015)
17. De Weerd, H., Verbrugge, R.: Evolution of altruistic punishment in heterogeneous populations. J. Theor. Biol. **290**, 88–103 (2011)
18. De Weerd, H., Verbrugge, R., Verheij, B.: How much does it help to know what she knows you know? An agent-based simulation study. Artif. Intell. **199–200**, 67–92 (2013)
19. de Weerd, H., Verbrugge, R., Verheij, B.: Higher-order theory of mind in the Tacit Communication Game. Biol. Inspired Cogn. Archit. **11**, 10–21 (2015)
20. Diekmann, A.: Cooperation in an asymmetric volunteer's dilemma game theory and experimental evidence. Int. J. Game Theory **22**(1), 75–85 (1993)
21. Diekmann, A., Przepiorka, W.: Punitive preferences, monetary incentives and tacit coordination in the punishment of defectors promote cooperation in humans. Sci. Rep. **5**, 10321 (2015)
22. Ellsworth, P.C., Ross, L.: Public opinion and capital punishment: a close examination of the views of abolitionists and retentionists. Crime Delinq. **29**(1), 116–169 (1983)
23. Fehr, E., Gächter, S.: Altruistic punishment in humans. Nature **415**(6868), 137–140 (2002)
24. Fowler, J.H.: Altruistic punishment and the origin of cooperation. Proc. Natl Acad. Sci. U.S.A. **102**(19), 7047–7049 (2005)
25. Frank, S.A.: A general model of the public goods dilemma. J. Evol. Biol. **23**(6), 1245–1250 (2010)
26. Gao, J., Li, Z., Cong, R., Wang, L.: Tolerance-based punishment in continuous public goods game. Physica A **391**(16), 4111–4120 (2012)
27. Gavrilets, S.: Collective action problem in heterogeneous groups. Proc. R. Soc. Lond. B **370**, 20150016 (2015)
28. Gavrilets, S.: Collective action and the collaborative brain. J. R. Soc. Interface **12**(102), 20141067 (2015)
29. Gavrilets, S., Fortunato, L.: A solution to the collective action problem in between-group conflict with within-group inequality. Nat. Commun. **5**, 3526 (2014)
30. Gavrilets, S., Richerson, P.J.: Collective action and the evolution of social norm internalization. Proc. Natl Acad. Sci. U.S.A. **114**, 6068–6073 (2017)
31. Gavrilets, S., Auerbach, J., van Vugt, M.: Convergence to consensus in heterogeneous groups and the emergence of informal leadership. Sci. Rep. **6**, 29704 (2016). https://doi.org/10.1038/srep29704
32. Gilby, I.C., Machanda, Z.P., Mjungu, D.C., Rosen, J., Muller, M.N., Pusey, A.E., Wrangham, R.W.: Impact hunters catalyse cooperative hunting in two wild chimpanzee communities. Philos. Trans. R. Soc. B **370**(1983), 20150005 (2015)
33. Gillespie, D.T.: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. **22**(4), 403–434 (1976)
34. Gintis, H.: Strong reciprocity and human sociality. J. Theor. Biol. **206**(2), 169–179 (2000)
35. Gintis, H.: Moral sentiments and Material Interests: The Foundations of Cooperation in Economic Life. MIT, Cambridge (2005)
36. Gintis, H., Smith, E.A., Bowles, S.: Costly signaling and cooperation. J. Theor. Biol. **213**(1), 103–119 (2001)
37. Glowacki, L., Von Rueden, C.: Leadership solves collective action problems in small-scale societies. Philos. Trans. Roy. Soc. London. **370**, 20150010 (2015)
38. Glowacki, L., Wrangham, R.W.: The role of rewards in motivating participation in simple warfare. Hum. Nat. **24**(4), 444–460 (2013)
39. Goeree, J.K., Holt, C.A., Palfrey, T.R.: Quantal Response Equilibrium. Princeton University Press (2016)
40. Hawkes, K.: Sharing and collective action. In: Smith, E.A., Winterhalder, B. (eds.) Foundations of Human Behavior: Evolutionary Ecology and Human Behavior, pp. 269–300. Aldine de Gruyter, Hawthorne (1992)
41. Hedden, T., Zhang, J.: What do you think i think you think? Strategic reasoning in matrix games. Cognition **85**(1), 1–36 (2002)
42. Helbing, D., Szolnoki, A., Perc, M., Szabó, G.: Punish, but not too hard: how costly punishment spreads in the spatial public goods game. New J. Phys. **12**(8), 083005 (2010)
43. Henrich, J.P.: Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies. Oxford University Press on Demand, Oxford (2004)
44. Henrich, J., Boyd, R.: Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. J. Theor. Biol. **208**(1), 79–89 (2001)

45. Hill, K.R., Walker, R.S., Bozicevic, M., Eder, J., Headland, T., Headland, B., Helwett, B., Hurtado, A.M., Marlowe, F., Wiessner, P., Wood, B.: Co-residence patterns in hunter-gatherer societies show unique human social structure. Science **331**(6022), 1286–1289 (2011)
46. Hofbauer, J., Sandholm, W.: On the global convergence of stochastic fictitious play. Econometrica **70**(6), 2265–2294 (2002). https://doi.org/10.1111/j.1468-0262.2002.00440.x
47. Hooper, P.L., Kaplan, H.S., Boone, J.L.: A theory of leadership in human cooperative groups. J. Theor. Biol. **265**, 633–646 (2010)
48. Iwasa, Y., Lee, J.-H.: Graduated punishment is efficient in resource management if people are heterogeneous. J. Theor. Biol. **333**, 117–125 (2013)
49. Jehiel, P.: Limited horizon forecast in repeated alternate games. J. Econ. Theory **67**(2), 497–519 (1995)
50. Jehiel, P.: Limited foresight may force cooperation. Rev. Econ. Stud. **68**(2), 369–391 (2001). https://doi.org/10.1111/1467-937X.00173
51. King, A.J., Johnson, D.D., Van Vugt, M.: The origins and evolution of leadership. Curr. Biol. **19**(19), 911–916 (2009)
52. Kitchen, D.M., Beehner, J.C.: Factors affecting individual participation in group-level aggression among non-human primates. Behaviour **144**(12), 1551–1581 (2007)
53. Konrad, K.A.: Strategy and Dynamics in Contests. Oxford University Press, Oxford (2009)
54. Krasnow, M.M., Cosmides, L., Pedersen, E.J., Tooby, J.: What are punishment and reputation for? PLoS ONE **7**(9), 45662 (2012). https://doi.org/10.1371/journal.pone.0045662
55. Krupenye, C., Kano, F., Hirata, S., Call, J., Tomasello, M.: Great apes anticipate that other individuals will act according to false beliefs. Science **354**, 110–114 (2016)
56. MacNulty, D.R., Smith, D.W., Mech, L.D., Vucetich, J.A., Packer, C.: Nonlinear effects of group size on the success of wolves hunting elk. Behav. Ecol. **23**(1), 75–82 (2011)
57. MacNulty, D.R., Tallian, A., Stahler, D.R., Smith, D.W.: Influence of group size on the success of wolves hunting bison. PLoS ONE **9**(11), 112884 (2014)
58. Mathew, S., Boyd, R.: Punishment sustains large-scale cooperation in prestate warfare. Proc. Natl Acad. Sci. U.S.A. **108**(28), 11375–11380 (2011)
59. Mathew, S., Boyd, R.: The cost of cowardice: punitive sentiments towards free riders in turkana raids. Evol. Hum. Behav. **35**(1), 58–64 (2014)
60. McElreath, R., Boyd, R.: Mathematical Models of Social Evolution. A Guide for the Perplexed. Chicago University Press, Chicago (2007)
61. McGinty, M., Milam, G.: Public goods provision by asymmetric agents: experimental evidence. Soc. Choice Welfare **40**(4), 1159–1177 (2013)
62. McKelvey, R.D., Palfrey, T.R.: Quantal response equilibria for normal form games. Games Econ. Behav. **10**, 6–38 (1995)
63. Monnin, T., Ratnieks, F.L.: Policing in queenless ponerine ants. Behav. Ecol. Sociobiol. **50**(2), 97–108 (2001)
64. Nagel, R.: Unraveling in guessing games: an experimental study. Am. Econ. Rev. **85**(5), 1313–1326 (1995)
65. Nagin, D.S.: Deterrence and incapacitation. In: Tonry, M.H. (ed.) The Handbook of Crime and Punishment, pp. 345–368. Oxford University Press, Oxford (1998)
66. Nakamaru, M., Iwasa, Y.: The evolution of altruism by costly punishment in lattice-structured populations: score-dependent viability versus score-dependent fertility. Evol. Ecol. Res. **7**(6), 853–870 (2005)
67. Nikiforakis, N.: Punishment and counter-punishment in public good games: can we really govern ourselves? J. Public Econ. **92**, 91–112 (2008)
68. Nunn, C.L.: Collective benefits, free-riders, and male extra-group conflict. In: Kappeler, P.M. (ed.) Primate Males: Causes and Consequence of Variation in Group Composition, pp. 192–204. Cambridge University Press, Cambridge (2000)
69. Nunn, C.L., Deaner, R.O.: Patterns of participation and free riding in territorial conflicts among ringtailed lemurs (*Lemur catta*). Behav. Ecol. Sociobiol. **57**(1), 50–61 (2004)
70. O'Gorman, R., Henrich, J., Van Vugt, M.: Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation. Proc. R. Soc. Lond. B Biol. Sci. **276**(1655), 323–329 (2009)
71. Olson, M.: The Logic of Collective Action: Public Goods and the Theory of Groups. Harvard University Press, Cambride (1965)
72. Ostrom, E.: Collective action and the evolution of social norms. J. Econ. Perspect. **14**(3), 137–158 (2000)
73. Ostrom, E.: Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, Cambride (2015)
74. Packer, C., Ruttan, L.: The evolution of cooperative hunting. Am. Nat. **132**(2), 159–198 (1988)

75. Patterson, S.K., Sandel, A.A., Miller, J.A., Mitani, J.C.: Data quality and the comparative method: the case of primate group size. Int. J. Primatol. **35**(5), 990–1003 (2014). https://doi.org/10.1007/s10764-014-9777-1
76. Perner, J., Wimmer, H.: John thinks that Mary thinks that Attribution of second-order beliefs by 5-to 10-year-old children. J. Exp. Child Psychol. **39**(3), 437–471 (1985)
77. Premack, D., Woodruff, G.: Does the chimpanzee have a theory of mind? Behav. Brain Sci. **1**(4), 515–526 (1978)
78. Przepiorka, W., Diekmann, A.: Individual heterogeneity and costly punishment: a volunteer's dilemma. Proc. R. Soc. Lond. B Biol. Sci. **280**(1759), 20130247 (2013)
79. Richerson, P.J., Boyd, R.: Not by Genes Alone. How Culture Transformed Human Evolution. The University of Chicago Press, Chicago (2005)
80. Rusch, H., Gavrilets, S.: The logic of animal intergroup conflict: a review. J. Econ. Behav. Org. **25**, 373–389 (2018)
81. Sääksvuori, L., Mappes, T., Puurtinen, M.: Costly punishment prevails in intergroup conflict. Proc. R. Soc. Lond. B Biol. Sci. **278**(1728), 3428–3436 (2011)
82. Sandholm, W.H.: Population Games and Evolutionary Dynamics. MIT, Cambridge (2010)
83. Shimao, H., Nakamaru, M.: Strict or graduated punishment? Effect of punishment strictness on the evolution of cooperation in continuous public goods games. PLoS ONE **8**(3), 59894 (2013)
84. Sigmund, K., Hauert, C., Nowak, M.: Reward and punishment. Proc. Natl Acad. Sci. U.S.A. **98**(19), 10757–10762 (2001)
85. Singh, M., Wrangham, R., Glowaki, L.: Self-interest and the design of rules. Hum. Nat. **28**, 457–480 (2017)
86. Smith, E., Bird, R.B.: Turtle hunting and tombstone opening: public generosity as costly signaling. Evol. Hum. Behav. **21**(4), 245–261 (2000)
87. Smith, J.E., Gavrilets, S., Mulder, M.B., Hooper, P.L., El Mouden, C., Nettle, D., Hauert, C., Hill, K., Perry, S., Pusey, A.E., et al.: Leadership in mammalian societies: emergence, distribution, power, and payoff. Trends Ecol. Evol. **31**(1), 54–66 (2016)
88. Sober, E., Wilson, D.S.: Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge (1999)
89. Stahl, D.O., Wilson, P.W.: On players models of other players: theory and experimental evidence. Games Econ. Behav. **10**(1), 218–254 (1995)
90. Taylor, M.: Anarchy and Cooperation. Wiley, New York (1976)
91. Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H.: Understanding and sharing intentions: the origins of cultural cognition. Behav. Brain Sci. **28**(05), 675–691 (2005)
92. Vidmar, N.: Retributive and utilitarian motives and other correlates of Canadian attitudes toward the death penalty. Can. Psychol. **15**(4), 337–356 (1974)
93. Vidmar, N., Ellsworth, P.: Public opinion and the death penalty. Stanf. Law Rev. **26**(6), 1245–1270 (1973)
94. Whitehouse, H., Jong, J., Buhrmester, M.D., Gómez, Ángel, Bastian, B., Kavanagh, C.M., Newson, M., Matthews, M., Lanman, J.A., McKay, R., Gavrilets, S.: The evolution of extreme cooperation via shared dysphoric experiences. Sci. Rep. **7**, 44292 (2017)
95. Wilkinson, G.S.: Recipricał food sharing in the vampire bat. Nature **308**(5955), 181–184 (1984)
96. Willems, E.P., van Schaik, C.P.: Collective action and the intensity of between-group competition in nonhuman primates. Behav. Ecol. **26**(2), 625–631 (2015)
97. Willems, E.P., Hellriegel, B., van Schaik, C.P.: The collective action problem in primate territory economics. Proc. R. Soc. Lond. B Biol. Sci. **280**(1759), 20130081 (2013)
98. Wilson, M.L., Kahlenberg, S.M., Wells, M., Wrangham, R.W.: Ecological and social factors affect the occurrence and outcomes of intergroup encounters in chimpanzees. Anim. Behav. **83**(1), 277–291 (2012)
99. Xu, Z.: Convergence of best-response dynamics in extensive-form games. J. Econ. Theory **162**, 21–54 (2016). https://doi.org/10.1016/j.jet.2015.12.001
100. Zhuang, Q., Wang, D., Fan, Y., Di, Z.: Evolution of cooperation in a heterogeneous population with influential individuals. Physica A **391**(4), 1735–1741 (2012)

Supporting Information for Perry et al. "Collective action problem in heterogeneous populations with punishment and foresight"

Additional details of numerical simulations

Parameters. In addition to the parameters specified in the text, all simulations were run using 100 groups, a half-success effort of $X_0 = 1$, a cost parameter of $c = 0.5$, and the scaling parameters $e = 1.0, \phi = 5.0, S_0 = 1$. All simulations were run for 30,000 time steps, meaning on average each group underwent $300(1 + E_2)$ events (this is because "us vs them" events include two groups while us vs. nature and cultural group selection events included only one group affected). Individuals updated their strategies with probability $\mu = 0.5$, which means on average each individual updated their strategies $150(1 + E_2)$ times. Candidate strategies were generated by perturbing an individual's current strategy by Gaussian noise with a standard deviation of $\sigma = 0.5$. The standard deviation of the error ξ in evaluating strengths was $\varepsilon = 0.10$. We performed 20 independent runs for each parameter combination.

Initial conditions. At the start of a run, individuals' initial efforts were drawn from a uniform distribution over the interval $[0, 0.05]$. When applicable, initial punishment thresholds were assigned in the same way. Individuals were also assigned non-evolvable strengths s_i at the beginning of each run (via a uniform distribution over the interval $[0, 1]$) and then sorted according to their strengths.

Bounds. Individual's efforts x were restricted to be both nonnegative and below the quantity $\frac{1+\tilde{B}v_i}{c}$ (to ensure there are no negative payoffs). Punishment thresholds y were taken to be nonnegative, but had no upper bound in place. In the event that a candidate strategy was a negative number, we took that candidate strategy to be 0. We did not disallow repeated candidate strategies. All payoffs were artificially bounded below by 10^{-5}.

Summary statistics and graphics. During simulations the data were saved every 10th time steps. Our graphs report the average values obtained after discarding the first 10,000 time steps. That the reported values correspond to stochastic equilibria was established by visual inspection of trajectories. Note that all runs were done for a minimum of 30,000 time steps, which translates into individuals updating their strategies 150 times on average. For each run the averages of all individuals of equivalent rank are taken across all groups. For the case of hierarchal groups, stacked histograms were used to relate the measurement corresponding to each individuals while simultaneously showing the group total. In the stacked histograms, the lowest segment corresponds to the dominant individual (with the largest strength s) while the upper-most corresponds to the weakest group-member. All figures were generated using the Gnuplot software or Matlab.

Comments. Individuals were allowed to update their punishment thresholds only if they estimated they were stronger than at least one other individual in their group (that is, if $s_i > s_j + \eta$ for at least one other individual j). Moreover, if the cost of punishing exceeds an individual's expected payoff, then they simply punish as much as their expected payoff allows. Our results occasionally show the development of large thresholds by weaker individuals. These however do not translate into higher actual punishment by such individuals as they they punish very rarely and only similarly week individuals. The development of this "false bravado" in weaker individuals becomes less common as precision λ increases. Additionally, it is absent in stronger individuals since they actively implement punishment.

The model was implemented in C; some of its components were also independently implemented in Matlab (with similar results). Simulations were run on a cluster.

Additional figures

Basic egalitarian model. Figure S1 investigates the effects of benefit b, the precision parameter λ, and the games frequencies ratio E on the group effort in the egalitarian model. Note that if the groups are only engaged in "us vs. nature" games (i.e., if $E = 1 : 0$) and precision in payoff evaluation is perfect ($\lambda = \infty$), groups are predicted to contribute only if $b > cX_0$ (Gavrilets 2015a,b). For parameter values used in Figure S1, this implies $b > 0.5$. This is indeed what is observed in simulations (see the leftmost graph).

Figures S2 and S3 use longer runs (100,000 times steps) to illustrates the effects of parameter K. It shows that with large K and with groups engaged in "us. vs. them" games, the system does not settle to an equilibrium but exhibits non-equilibrium dynamics even when precision in payoff evaluation is perfect ($\lambda = \infty$).

Basic hierarchy model. Figures S4, S5, and S6 provide additional information on the effects of various parameters on the individual and group efforts and on the relative payoffs of individuals of different ranks. In these figures, we explore a wider range of parameters than in the main part. Specifically, we use five different values of the precision parameter ($\lambda = \infty, 80, 40, 20,$ and 10) in Fig. S4, five different values of the number of candidate strategies ($K = 1, 2, 4, 8,$ and 16) in Fig. S5, and five different values of the frequency of cultural group selection events ($E_3 = 0.00, 0.05, 0.10, 0.15,$ and 0.20) in Fig. S6.

Egalitarian model with foresight and punishment in "us vs. nature" games. Figure S7 shows the effect of foresight in the most basic case of our model. Namely, we assume that groups are egalitarian and only participating in "us vs nature" games. This is done to highlight the presence of the threshold effect established in prior works (Gavrilets 2015a,b), and illustrate how foresight overcomes the second-order free rider-effect. For the parameter values used, individuals without foresight should not contribute when $b \leq 1$. The leftmost graph in Fig. S7 shows that when foresight is absent ($\omega = 0$) then individuals only contribute for the cases of $b = 1.25, 1.50, 1.75,$ and 2.00. When foresight is present ($\omega = 0.50$), however, we see contributions in the cases of $b = 0.75$, and 1.00 as well. The reason for this is that foresight successfully solves the second-order free-rider effect as illustrated by the emergence of positive threshold values (centermost graph). It is important to note, that while foresight does promote cooperation, it does not necessarily increase the groups payoffs (rightmost graph). This is because of punishment costs.

Hierarchy model with foresight and punishment. In contrast to the figures in the main text, here we also show the punishment thresholds, the loss of payoffs due to punishing others and the total payoffs. Figure S8 shows the interactions of foresight

ω, group size N, and the game frequencies ratio E for $\lambda = 40$. Figure S9 shows the interactions of foresight ω, group size N, and precision λ for $E = 1:1$. Figure S10 shows the interactions of benefit b, group size N, and the hierarchy steepness β for $\omega = 0.5, \lambda = 40$, and $E = 1:1$.

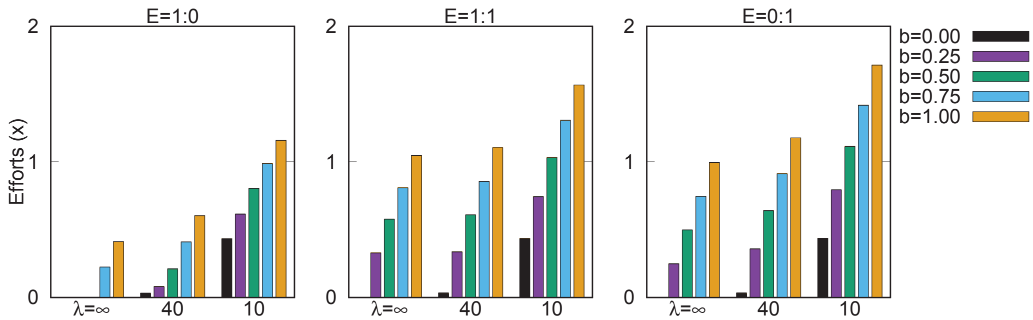

Fig. S1: Effects of the precision parameter λ, the games frequencies ratio E, and the average benefit b on the group effort X in the basic egalitarian model. For each value of E, the bars (from left to right) correspond to $b = 0.00, 0.25, 0.50, 0.75, 1.00$. Results are the averages of 20 simulations using: $N = 8, c = 0.5, X_0 = 1.0, K = 2$, and $E_3 = 0.1$.

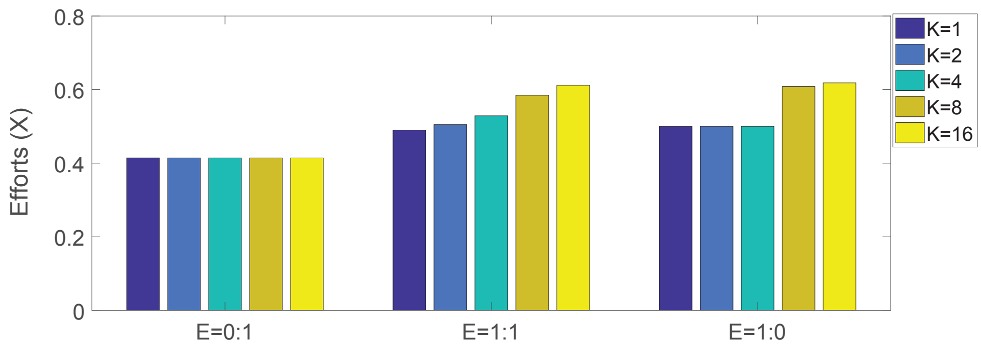

Fig. S2: Effects of the games frequencies ratio E, and the number of candidate strategies K on group efforts X in the basic egalitarian model with perfect precision, i.e. $\lambda = \infty$. For each value of E, the bars (from left to right) correspond to $K = 1, 2, 4, 8$, and 16. Results are the averages of 20 simulations using: $N = 8, b = 1.0, c = 0.5, X_0 = 1.0$, and $E_3 = 0.0$.

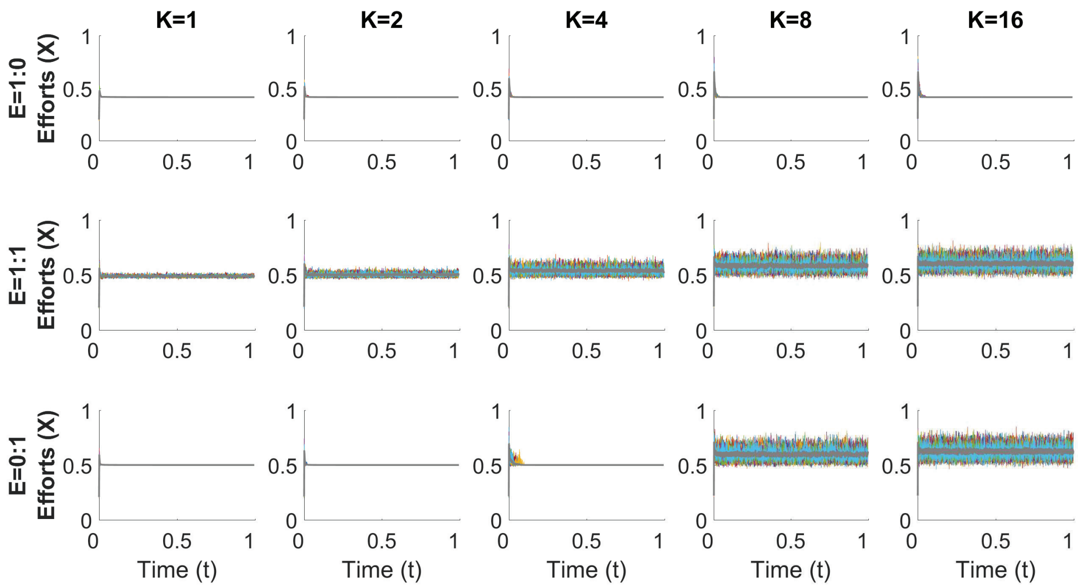

Fig. S3: Time series summarizing the effects of the games frequencies ratio E (varies by row), and the number of candidate strategies K (varies by column) on group efforts X in the basic egalitarian model with perfect precision, i.e. $\lambda = \infty$. The x-axis uses units of 100,000 timesteps. Results are the averages of 20 simulations using: $N = 8, b = 1.0, c = 0.5, X_0 = 1.0$, and $E_3 = 0.0$.

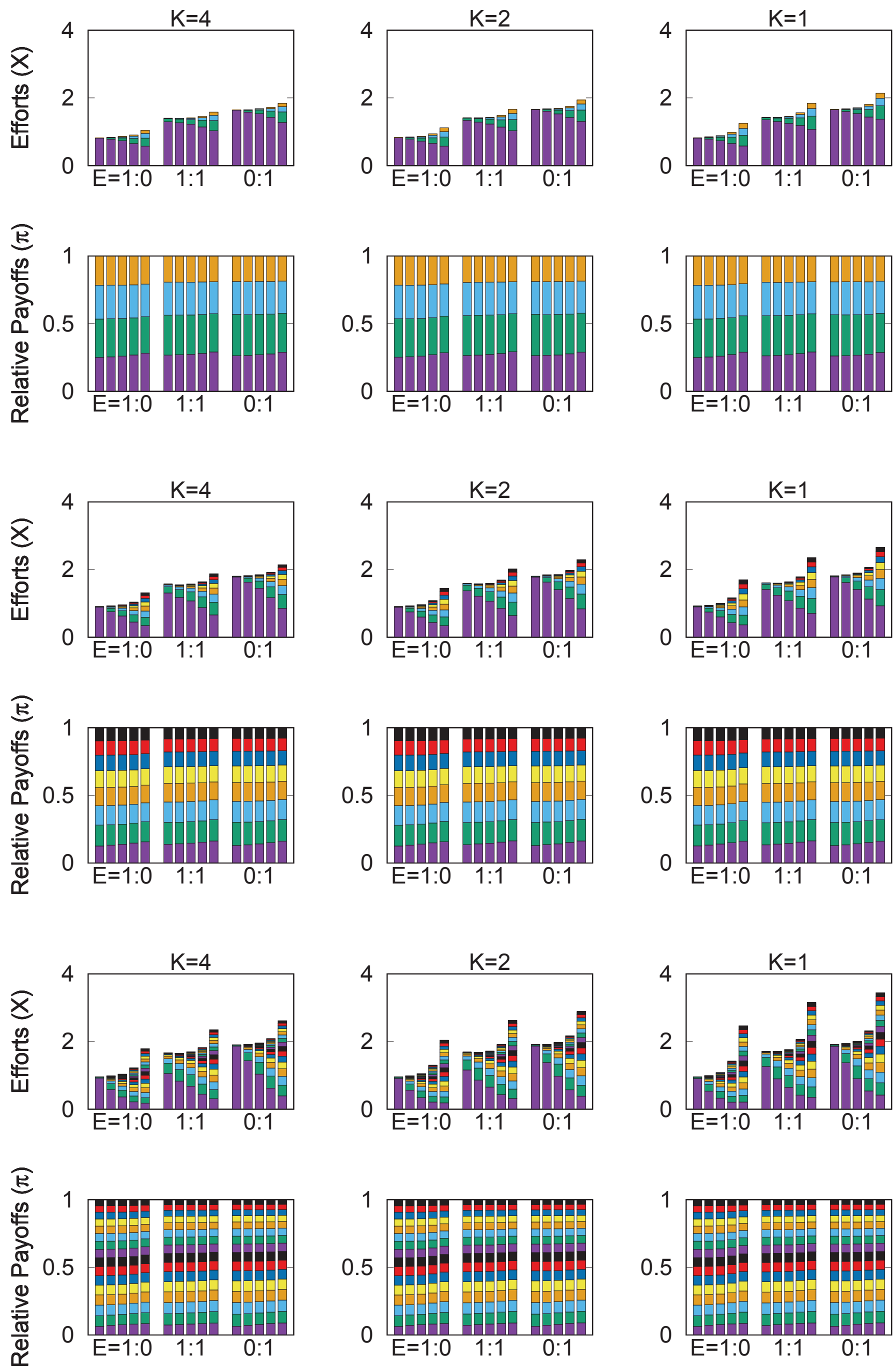

Fig. S4: Effects of the game frequencies ratio E, the number of candidate strategies K, and the precision parameter λ on the group effort X and the relative payoffs π of individuals of different ranks in the basic hierarchical model. The 1st and 2nd rows correspond to $N = 4$, the 3rd and 4th rows to $N = 8$, and the 5th and 6th to $N = 16$. For each value of E, the bars (from left to right) correspond to $\lambda = \infty, 80, 40, 20, 10$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $b = 1.0, c = 0.5, \beta = 1.0$, and $E_3 = 0.1$.

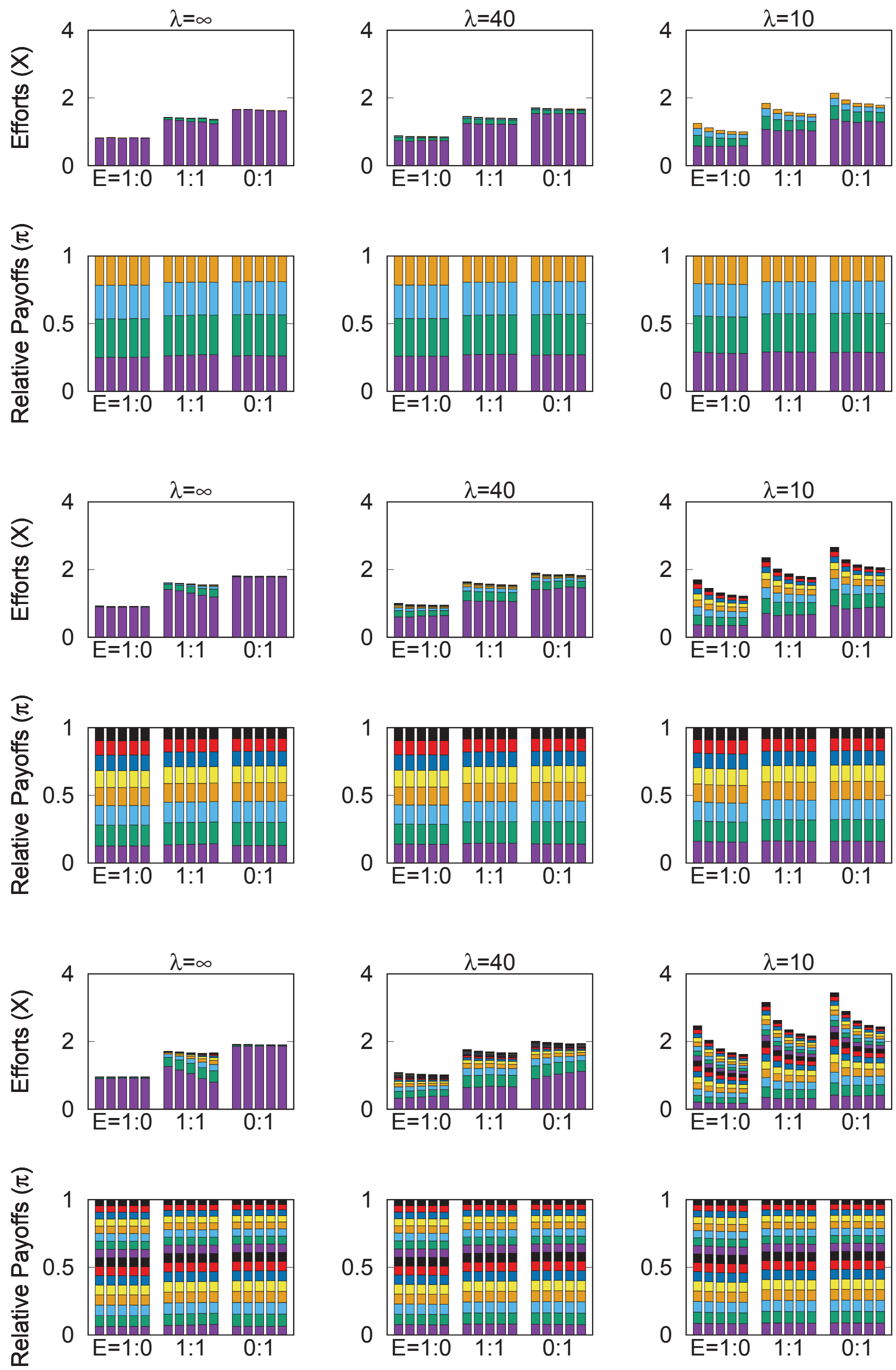

Fig. S5: Effect of the game frequencies ratio E, the precision parameter λ, and the number of candidate strategies K on the group effort X and the relative payoffs π of individuals of different ranks in the basic hierarchical model. The 1st and 2nd rows correspond to $N = 4$, the 3rd and 4th rows to $N = 8$, and the 5th and 6th to $N = 16$. For each value of E, the bars (from left to right) correspond to $K = 1, 2, 4, 8, 16$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $N = 8, b = 1.0, c = 0.50, \beta = 1.0$, and $E_3 = 0.10$.

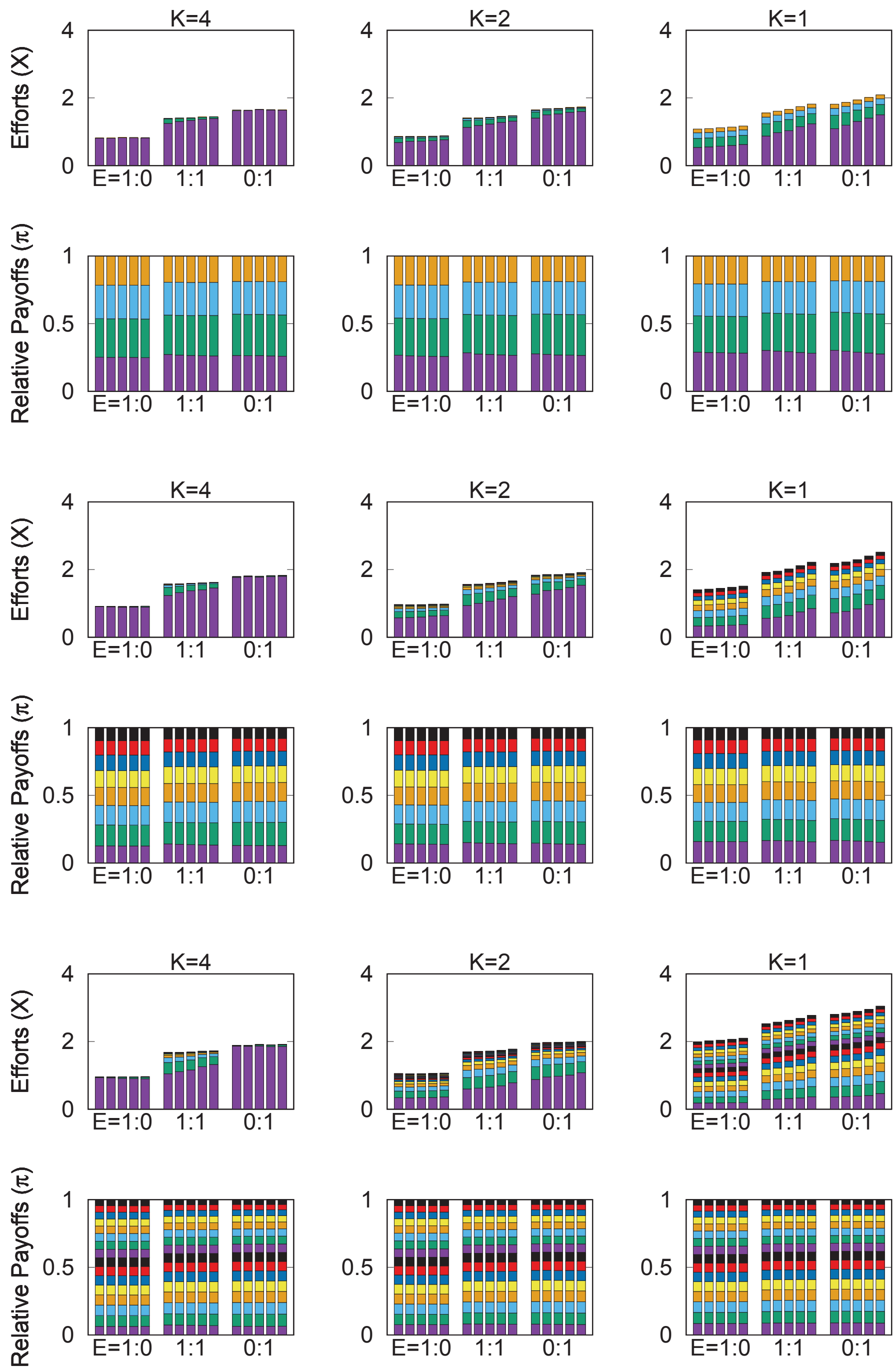

Fig. S6: Effect of the game frequencies ratio E, the precision parameter λ, and the frequency of cultural group selection events E_3 on the group effort X and the relative payoffs π of individuals of different ranks in the basic hierarchical model. The 1st and 2nd rows correspond to $N = 4$, the 3rd and 4th rows to $N = 8$, and the 5th and 6th to $N = 16$. For each value of E, the bars (from left to right) correspond to $E_3 = 0.00, 0.05, 0.10, 0.15, 0.20$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $b = 1.0, c = 0.50, \beta = 1.0$, and $K = 2$.

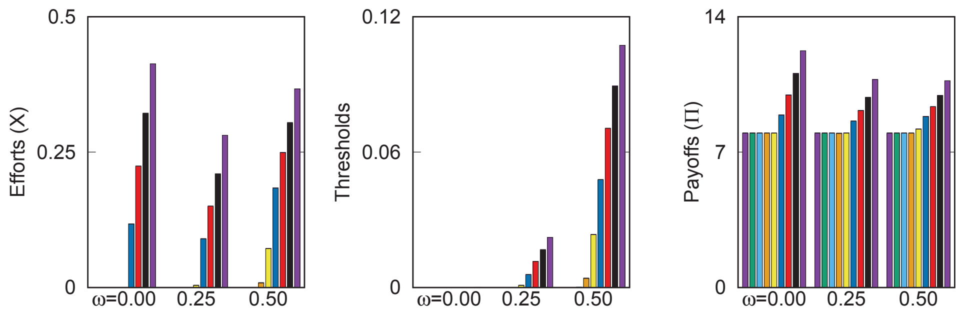

Fig. S7: Results summarizing the effects of the emphasis on future payoffs ω, and the average benefit b for the case of egalitarian groups with perfect precision ,i.e. $\lambda = \infty$, when foresight is present. The graphs (from left to right) summarize group efforts X, group punishment thresholds, and group payoffs Π. Each set of 9 bars corresponds to a specific value of omega, while each bar within a set corresponds to a specific average benefit ranging from the smallest on the left ($b = 0.00$) to the largest on the right ($b = 2.00$) by increments of 0.25. Results are the averages of 20 simulations using: $c = 1, K = 2, X_0 = 1, E = 1 : 0$, and $E_3 = 0$.

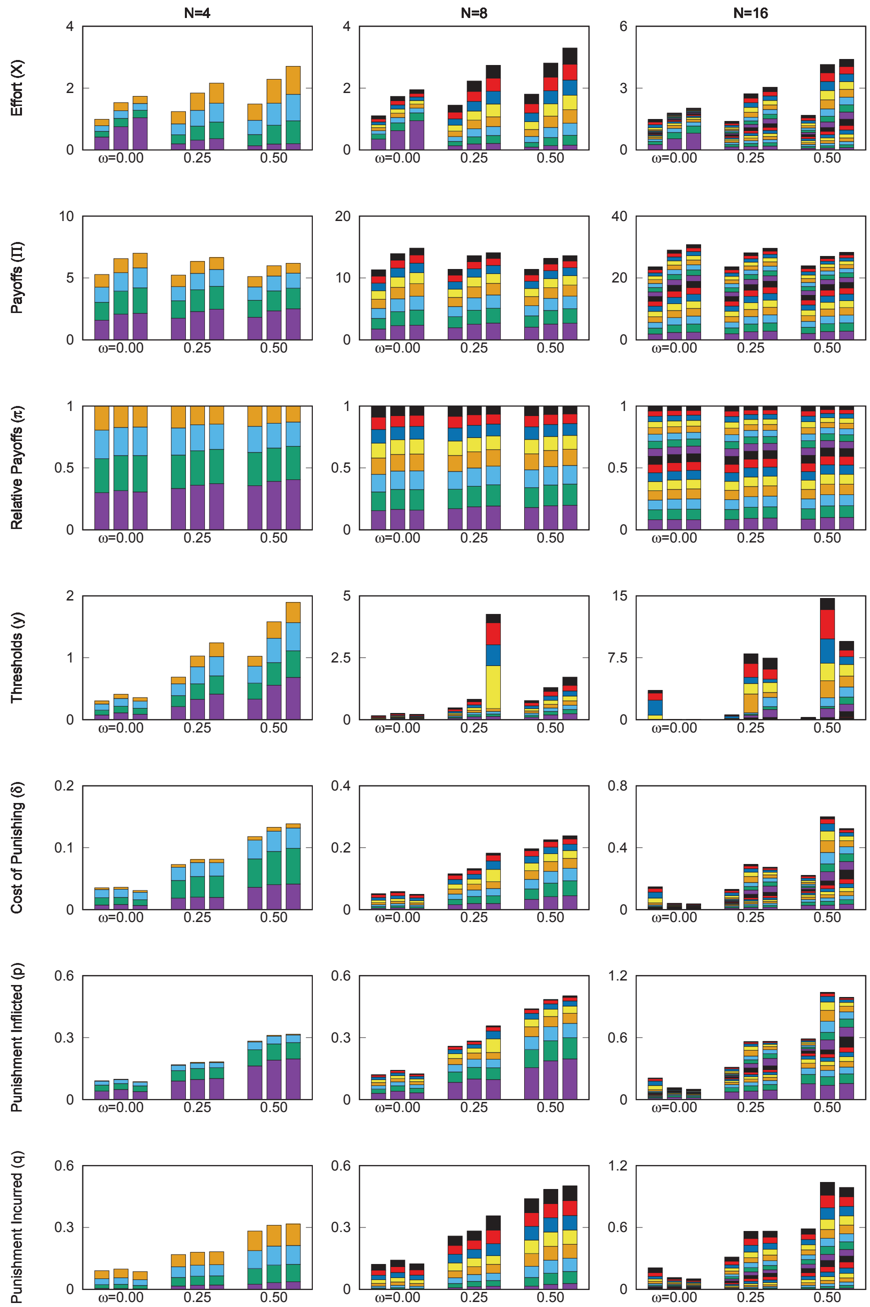

Fig. S8: Effects of the foresight parameter ω, the group size N, and the game frequencies ratio E in the full model for precision $\lambda = 40$. The first column contains results for $N = 4$, the second for $N = 8$, and the third for $N = 16$. The rows from top to bottom summarize the effort levels X, the total payoffs Π, the relative payoffs π, the thresholds y, the cost of punishing δ, the punishment inflicted p, and the punishment incurred q. For each value of ω, the bars (from left to right) correspond to $E = 1:0, 1:1, 0:1$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $b = 1.0, c = 0.5, \beta = 1, K = 2$, and $E_3 = 0.1$.

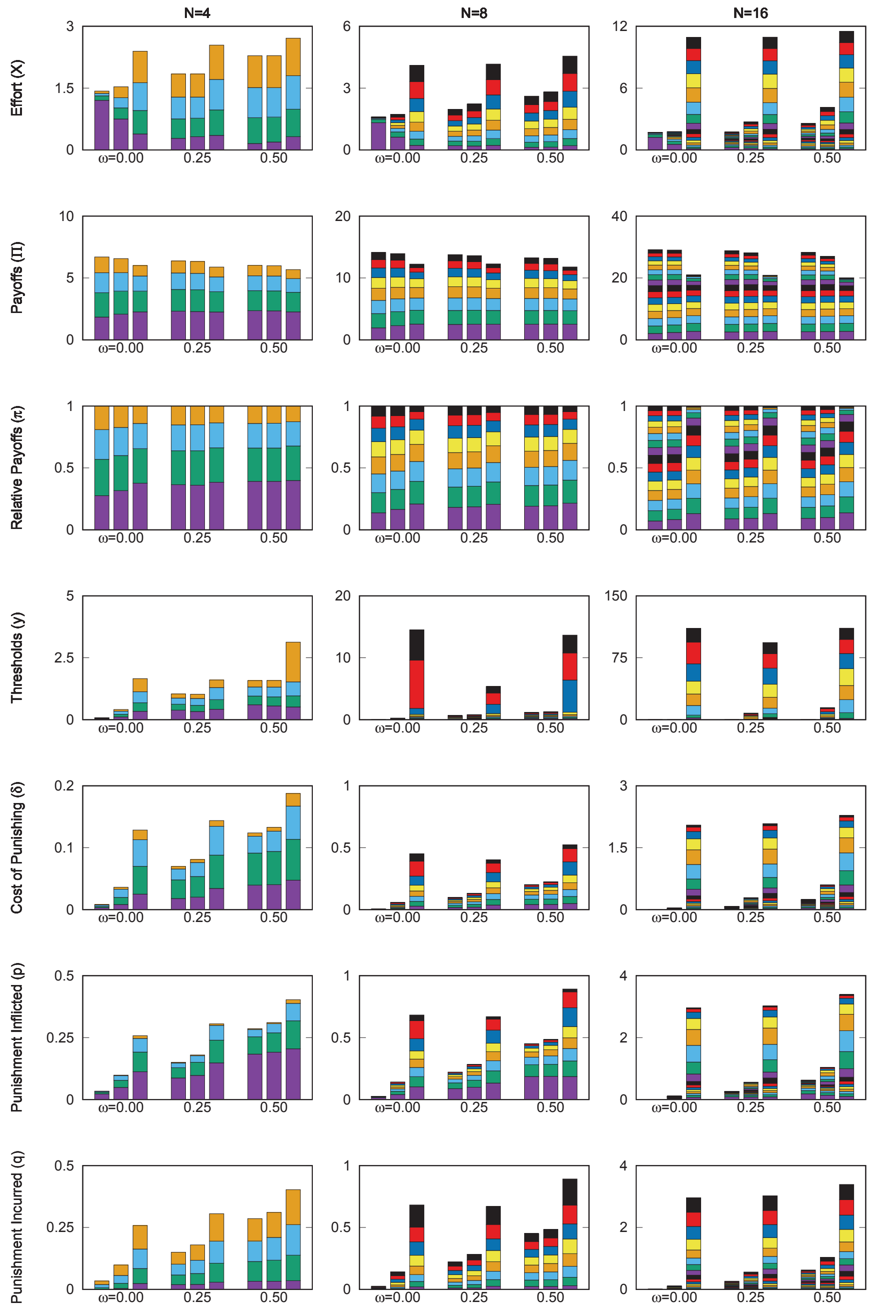

Fig. S9: Effects of the foresight parameter ω, the group size N, and the precision λ in the full model. The first column contains results for $N = 4$, the second for $N = 8$, and the third for $N = 16$. The rows from top to bottom summarize the effort levels X, the total payoffs Π, the relative payoffs π, the thresholds y, the cost of punishing δ, the punishment inflicted p, and the punishment incurred q. For each value of ω, the bars (from left to right) correspond to $\lambda = \infty, 40, 10$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $b = 1.0, c = 0.5, \beta = 1, K = 2, E = 1 : 1$, and $E_3 = 0.1$.

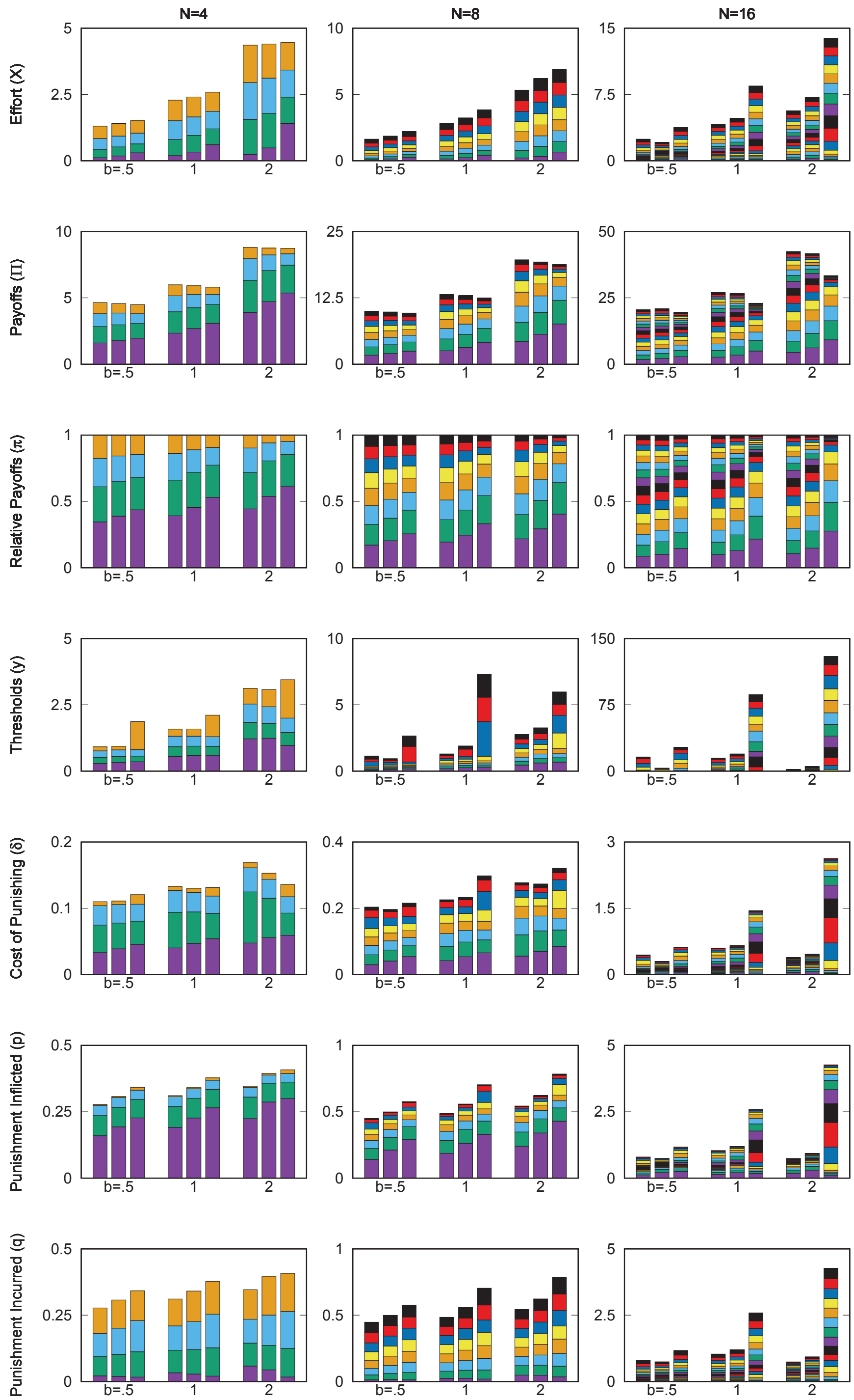

Fig. S10: Effects of the expected benefit b, the group size N, and hierarchy steepness β in the full model. The first column contains results for $N = 4$, the second for $N = 8$, and the third for $N = 16$. The rows from top to bottom summarize the effort levels X, the total payoffs Π, the relative payoffs π, the thresholds y, the cost of punishing δ, the punishment inflicted p, and the punishment incurred q. For each value of b, the bars (from left to right) correspond to $\beta = 1, 2, 4$. The segments of each bar correspond to a particular individual with the dominant at the bottom (purple) and the weakest at the top. Results are the averages of 20 simulations using: $c = 0.5, K = 2, \lambda = 40, E = 1 : 1, E_3 = 0.1$ and $\omega = 0.50$.