Overview:

Purpose:

We are interested in the evolution of altruistic punishment in a population of strategic agents playing a public good game. We examine the effects of population size, and the initial population on the trajectory of punishment.

State variables and scales:

We consider individual strategic agents as well as groups in which they interact. The strategic individuals are characterized by the group they are in, their fitness, the number of others in the group that they expect to contribute to public good, the number of others they expect to punish those that don't contribute, how much it costs them to contribute to the public good, the amount they benefit from the public good, the agents attitude towards doing better than the mean of their group in a given game, and the agents attitude towards doing worse than the mean of their group in a given game. Groups are characterized by the fraction of strategies that involved contribution by its members over a round of games.

Process overview and scheduling:

We consider a single round of games. At the beginning of each round strategic agents rest their fitness to 0 and groups reset their contribution count. Next every group plays the game; for each game the strategic agents calculate an expected fitness for each strategy (Expectation), choose a strategy (Strategy) and the results of the game are awarded (Calculate Fitness). Following the games group fraction of contribution is calculated. Strategic agents update their expectations of contribution and punishment (Learn). Next strategic agents modify their attitudes to that of another agent with probability (Mimic). Groups engage in conflict probabilistically with the winning group replacing the losing group (Conflict). Finally strategic agents mutate introducing change to both their attitudes and expectations (Mutate). This process is iterated until contribution and punishment levels in the population stabilize.

Design Concepts:

Basic principles: **Which general concepts, theories, hypotheses, or modeling approaches are underlying the model's design?**

The model is based on cultural group selection. We consider strategic agents that have expectations and learn based on insights from behavioral economics. We restrict the number of possible decisions but allow agents to make deliberate decisions. We allow for probabilistic error in every decision.

Emergence: **What key results or outputs of the model are modeled as emerging from the adaptive traits, or behaviors, of individuals?**

The emergence of population wide cooperation and punishment from a single cooperative seed group. In addition we observe the evolution of attitudes (norms) related to cooperation in our agents.

Adaptation: **What adaptive traits do the individuals have? What rules do they have for making decisions or changing behavior in response to changes in themselves or their environment?**

Agents update their expectation for the levels of contribution and punishment according to learning algorithm, they copy the attitudes of more successful agents and they take on the attitudes and expectations of members of more successful groups. Finally both agent's attitude and their expectations are subject to mutation.

Objectives: **If adaptive traits explicitly act to increase some measure of the individual's success at meeting some objective, what exactly is that objective and how is it measured?**

Agents act to optimize their attitude adjusted fitness. To do so they compute expectations, and learn from previous interactions.

Learning: **Many individuals or agents (but also organizations and institutions) change their adaptive traits over time as a consequence of their experience? If so, how?**

We implement a learning algorithm for expectations, mimicry for attitudes, and a group replacement that changes both expectations and attitudes.

Prediction: **Prediction is fundamental to successful decision-making; if an agent's adaptive traits or learning procedures are based on estimating future consequences of decisions, how do agents predict the future conditions (either environmental or internal) they will experience?**

Agents have an expected ratio of behaviors from other agents in their group. They use the expected ratio to compute the expected utility for each of their possible strategies where the utility depends on the expected winnings, the expected mean of the winnings of the others in the group, and the attitudes of the agent.

Sensing: **What internal and environmental state variables are individuals assumed to sense and consider in their decisions?**

Agents can sense the number of contributors and the number of punishers in their group, in the previous round.

Interaction: **What kinds of interactions among agents are assumed? Are there direct interactions in which individuals encounter and affect others, or are interactions indirect, e.g., via competition for a mediating resource?**

Agents interact with each other through a public goods game. Groups interact with each other through simulated conflict.

Stochasticity: **What processes are modeled by assuming they are random or partly random?**

The strategy an agent chooses has a random component as do the results of imitation and of group conflicts. The occurrence of imitation and group conflict are fully random as is mutation and assignment of the costs and benefits of public good to an agent.

Collectives: **Do the individuals form or belong to aggregations that affect, and are affected by, the individuals?**

Agents belong in groups that compete with other groups. The group's fitness is determined by the contribution of its members to public good. If a group loses a conflict the agents in it are changed to near copies of the winning group.

Observation: **What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected?**

We look at contribution and punishment levels in the population as well as the attitudes and expectations of agents. The majority of data collection is done once the results have stabilized.

Details:

Initialization: **What is the initial state of the model world, i.e., at time t = 0 of a simulation run?**

At initialization we consider 128 groups of N agents. All but one of these groups are selfish rational agents with no expectation of contribution or punishment. The other group has attitudes such that any payoff is has equal utility to the mean payoff of the others in the group, and expect all other members to both contribute and punish non-contributors.

Input data: **Does the model use input from external sources such as data files or other models to represent processes that change over time?**

Global

Variable	Description	Value
$nGroups$	Number of groups	128
N	Population of each group	2,4,8,16,32,64,128
$nGames$	Number of games per round	1
Punishment	Will punishment be allowed	On, Off
m	Probability that an agent mimics another /round	0.01, 0.002, 0.05
μ	Mutation (standard deviation of normal distribution)	0.05
ex	Probability of conflict	0.015, 0.0075, 0.03
λ	Learning rate	0.1
η	Error rate	0.02

Agent Level

Variable	Description	Value
p	Cost of being punished	0.8
k	Cost of punishing	0.2
b	Benefit if every member of group contributes	0.5
c	Cost of contributing	0.2
$E[C]$	Expected portion of others to contribute	$0 \leq E[C] \leq 1$
$E[P]$	Expected portion of contributors to punish	$0 \leq E[P] \leq 1$
α	Utility reduction for doing better than average	$-1 \leq \alpha \leq 1$
β	Utility increase for doing worse than average	$-1 \leq \beta \leq \alpha$

Submodels: **What, in detail, are the submodels that represent the processes listed in 'Process overview and scheduling'? What are the model parameters, their dimensions, and reference values? How were submodels designed or chosen, and how were they parameterized and then tested?**

<u>DECISION MAKING & PAYOFFS</u>
Agents may choose to contribute or free-ride in any game by weighing the following utility function (U):

$$U = x - \alpha * \max(x - \bar{x}, 0) + \beta * \max(\bar{x} - x, 0)$$

Here, x defines the earning of the focal agent and $\bar{x}$ the average earning of all other agents in the group. The parameter α defines the strength of aversion to exploiting others (i.e. the level of guilt an agent feels when it earns more than the average). Conversely, the parameter β defines the strength of altruistic tendencies (i.e. the pride an agent feels when it earns less than the average). Together, these parameters, along with expected contribution, determine whether an agent will cooperate.

In line with Charness and Rabin (2002), we can define the following cases for $\beta \leq \alpha \leq 1$

Case 1: The players like to have their payoffs higher than those of the other players.
If $\beta \leq \alpha \leq 0$, players are highly competitive.

Case 2: Players prefer the payoffs among all players to be equal. This "Inequity Aversion" holds when $\beta < 0 < \alpha \leq 1$ (see Fehr and Schmidt 1999).

Case 3: The third model approximates a "Social Welfare Consideration" which holds when $0 < \beta \leq \alpha \leq 1$. The parameter α captures the extent to which a player weighs the average payoffs of the other *n*-1 agents compared to his own payoff, when his own payoff is higher than the average payoff of the others.

Case 4: If $\alpha = \beta = 0$, then players only care about their welfare

We note that agents compute the expected utility of each course of action:

A. Don't contribute (Defection)
B. Contribute (Cooperation)
C. Contribute & Sanction (Altruistic punishment)

Agents make these calculations based on an initial expectation of cooperation in the first round among group-mates. This "trust" variable (T) is adaptive, subsequently updated based on observed levels of cooperation after each round of play.

<u>DECSISON MAKING WITHOUT THREAT OF PUNISHMENT</u>
We calculate the utility of agent i based on the contribution of her n-1 partners in the groups. For a cooperative agent the expected fraction of cooperators among her (n-1) group members is T_c (and 1 for the focal agent). Therefore we define the expected earnings for a cooperative agent as:

$$x_C^* = 1 + b \cdot \frac{\{T_c \cdot (n-1) + 1\}}{n} - c$$

In a similar way we can define the expected earning of a defecting agent as:

$$x_D^* = 1 + b \cdot \frac{\{T_c \cdot (n-1)\}}{n}$$

When agent i contributes with the expectation that fraction T_c of her group members contribute too, the expected earnings of other agents is equal to:

$$\overline{x_C^*} = T_c \cdot x_C^* + (1 - T_c) \cdot \left(x_D^* + \frac{b}{n}\right)$$

Conversely, if agent i is not contributing the benefit of b/n will not be enjoyed by her partners. Thus, the expected earnings of other agents is defined as:

$$\overline{x_D^*} = T_c \cdot \left(x_C^* - \frac{b}{n}\right) + (1 - T_c) \cdot x_D^*$$

We can define the expected utility of cooperation and defection in the following way:

$$E[U_C] = x_C^* + \beta_i * \left(\overline{x_C^*} - x_C^*\right) - \alpha_i * (x_C^* - \overline{x_C^*})$$
$$E[U_D] = x_D^* + \beta_i * \left(\overline{x_D^*} - x_D^*\right) - \alpha_i * (x_D^* - \overline{x_D^*})$$

DECISION MAKING UNDER THE THREAT OF PUNISHMENT

Suppose contributors may sanction defectors at a cost (k) to themselves (i.e. altruistic punishment). Now we must define expected earnings based on the perceived and actual values of public good.

The expected earning of a cooperative agent is given by:

$$x_C^* = 1 + b \cdot \frac{\{T_c \cdot (n - 1) + 1\}}{n} - c$$

The expected earning of an altruistic punisher is given by:

$$x_{CP}^* = 1 + b \cdot \frac{\{T_c \cdot (n - 1) + 1\}}{n} - c - k(1 - T_c) \cdot (n - 1)$$

The expected earning of a defecting agent is now given by:

$$x_D^* = 1 + b \cdot \frac{\{T_c \cdot (n - 1)\}}{n} - p \cdot T_p \cdot T_c \cdot (n - 1)$$

where T_p denotes the expected fraction of punishers.

The expected earning of other agents when agent i is cooperating is given by:

$$\overline{x_C^*} = T_c \cdot (1 - T_p) \cdot (x_C^*) + (1 - T_c) \cdot \left(x_D^* + \frac{b}{n}\right) + T_c \cdot T_p \cdot (x_{CP}^*)$$

The expected earning of other agents when agent i is an altruistic punisher is:

$$\overline{x_{CP}^*} = T_c \cdot (1 - T_p) \cdot (x_C^*) + (1 - T_c) \cdot \left(x_D^* + \frac{b}{n} - p\right) + T_c \cdot T_p \cdot (x_{CP}^*)$$

The expected earning of other agents when agent j is defecting is given by:

$$\overline{x_D^*} = T_c \cdot (1 - T_p) \cdot \left(x_C^* - \frac{b}{n}\right) + (1 - T_c) \cdot (x_D^*) + T_c \cdot T_p \cdot \left(x_{CP}^* - \frac{b}{n} - k\right)$$

Hence, the expected utility of cooperation, defection, and punishment are:

$$E[U_C] = x_C^* + \beta * \left(\overline{x_C^*} - x_C^*\right) - \alpha * (x_C^* - \overline{x_C^*})$$
$$E[U_D] = x_D^* + \beta * \left(\overline{x_D^*} - x_D^*\right) - \alpha * (x_D^* - \overline{x_D^*})$$
$$E[U_P] = x_{CP}^* + \beta * (\overline{x_{CP}^*} - x_{CP}^*) - \alpha * (x_{CP}^* - \overline{x_{CP}^*})$$

Once these utilities are known, the optimal decision is made with probability $(1 - \eta)$; with probability η the agents will switch from defecting to one of the cooperative strategies or vice versa.

LEARNING

Agents update their "trust" expectations at the end of each round according to the following equations:

$$T_c = (1 - \lambda) * T_c + \lambda * \frac{N_c}{n - 1}$$
$$T_P = (1 - \lambda) * T_P + \lambda * \frac{N_P}{n - 1}$$

where λ is the learning rate, N_c and N_P are the number of other agents that cooperated and punished in the current round respectively.

IMITATION

Every generation prior to group conflicts, agents may imitate the trust expectations, and welfare phenotypes (α & β) of a more successful agent in the group. This models the individual-level selection forces that promote payoff maximizing strategies. This submodel is run with probability m for each agent. Replacement of agent *i* by agent *j* has the probability:

$$\frac{\left(1 + (x_j - x_i)\right)}{2}$$

Where x_i denotes the payoff of agent *i.*

GROUP SELECTION

We assume like Boyd et al. (2003) that group selection occurs through intergroup conflicts. At the end of each generation, groups are randomly paired, and with probability ε the interaction results in one group defeating and replacing the other group. The probability that group *i* defeats group *j* is:

$$\frac{\left(1 + \left(N_{c,j} - N_{c,i}\right)\right)}{2}$$

Where $N_{c,l}$ is fraction of contributors in group l. This assumes agents who are more who are more successful in generating the public good for their group, are more likely to be imitated. As a consequence, cooperation is the sole target of the resulting group selection process.

MUTATION

Mutation occurs every generation after group selection. We assume that agent welfare phenotypes undergo mutations as follows:

Every agent replaces their current alpha value (α_t) and beta value (β_t) with a random number drawn from the normal distribution with standard deviation μ:

$$\alpha_{t+1} \sim N(\alpha_t, \mu)$$
$$\beta_{t+1} \sim N(\beta_t, \mu)$$

We conclude by imposing: $-1 \leq \beta_{t+1} \leq \alpha_{t+1} \leq 1$

We also assume offspring inherit "trust" expectations from their parents with some error:

$$T_{c_{t+1}} \sim N(T_{c_t}, \mu)$$
$$T_{p_{t+1}} \sim N(T_{c_t}, \mu)$$

Where:

$$0 \leq T_{c_{t+1}} \leq 1$$
$$0 \leq T_{p_{t+1}} \leq 1$$

Model implementation

The model is implemented in Netlogo 5.0.3

References

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson. 2003. The evolution of altruistic punishment. Proceedings of the National Academy of Sciences of the United States of America 100: 3531–3535.

Charness, G., and M. Rabin. 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117: 817-869.

Fehr, E., and K. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114: 817-868.