



OPEN

Quantal response equilibrium for the Prisoner's Dilemma game in Markov strategies

T. S. Kozitsina^{1,2✉}, I. V. Kozitsin^{1,3} & I. S. Menshikov^{1,2,4}

Within the studies of human cooperation, there are gaps that require further investigation. One possible area for growth is developing theoretical concepts which describe high levels of cooperation. In this paper, we present a symmetrical quantal response equilibrium (QRE) in Prisoner's Dilemma game (PD) constructed in Markov strategies (tolerance to defection and mutual cooperation). To prove the adequacy of the resulting equilibrium, we compare it with the previously found Nash equilibrium in PD in Markov strategies: the QRE converges with the Nash equilibrium that corresponds with the theory. Next, we investigate the properties of QRE in PD in Markov strategies by testing it against experimental data. For low levels of rationality, the found equilibrium manages to describe high cooperation. We derive the levels of rationality under which the intersection between Nash and QRE occurs. Lastly, our experimental data suggest that QRE serves as a dividing line between behavior with low and high cooperation.

There are still gaps in the research of human behavior that require investigation. We know that our way of thinking, actions, and beliefs depend on various internal and external factors. Human behavior includes the important social ability of cooperation as “the act of working together with someone or doing what they ask you”¹. Perhaps more importantly, cooperation relates to sharing mutual profit, equality, costs, and skills. Throughout the pandemic, we have realized the importance of cooperation for the benefit of society. In fact, human lives have depended on cooperation: wearing a mask, keeping social distance, and practicing patience and generosity with the larger public. Thus, studying people's ability to cooperate helps us better understand the choices and measures taken during the world pandemic².

Cooperation could be investigated through situations regarded as social dilemmas. In a social dilemma, an individual has a higher profit for defection when others cooperate. However, mutual cooperation leads to better outcomes than mutual defection. For the community, cooperation is preferable to defection; therefore, studying the factors that lead to cooperation is an important step toward understanding people's behavior in social dilemmas.

Researchers make different arguments for which factors increase cooperation in social dilemmas³: using communication^{4,5} or socialization^{6–9}, mobility and dynamics^{10–12}, connectivity^{13–15}, and aspects of an individual's identity^{16,17}. The choice to cooperate is often more of an intuitive act than a meaningful one. It is an emotional, quick, automatic operation that does not involve effort. To support this claim, Rand et al. compared the amount of time that participants in the experiments spent choosing between cooperation and non-cooperation strategies¹⁸. Their results indicated that quick choice could be one predictor of cooperation. Effects of sociality could also lead to an increased likelihood of cooperation¹⁹. For example, Capraro et al. introduce the cooperative equilibrium for explaining deviation from the Nash equilibrium, based on the idea that people have some tendency to cooperate by default^{20,21}. At the moment, this direction of unselfish behavior research is concentrated around the moral preferences hypothesis which postulates that such a sort of behavior appears because of our “internal standards about what is right or wrong in a given situation” (see Ref.²², p. 1).

There are numerous approaches to shifting strategies from individual to social. Which models can explain irrational cooperation in social dilemmas remains a question. Below we list concepts that accept relatively high cooperation levels:

¹Moscow Institute of Physics and Technology, Institutsky lane 9, Dolgoprudny, Moscow region 141700, Russian Federation. ²Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Vavilova street 44/2, Moscow 119333, Russian Federation. ³V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Profsoyuznaya street 65, Moscow 117997, Russian Federation. ⁴I. S. Menshikov is deceased. ✉email: tatyana.babkina@phystech.edu

Payoff	Left	Right
High	5, 5	0, 10
Low	10, 0	1, 1

Table 1. Payoff matrix in PD game.

- Quantal response equilibrium²³.
- Level-k²⁴.
- Cognitive hierarchy²⁵.
- Quantal level-k²⁶.
- Trembling hand perfect equilibrium²⁷.
- Proper equilibrium²⁸.

Previous studies have demonstrated that social interaction significantly increases the level of cooperation in iterated Prisoner's Dilemma (PD) games (the most commonly known example of social dilemma), from a 20% cooperation rate prior to socialization to 53% after socialization^{9,29–31}. We require a specific approach to model such a high level of cooperative strategy choice. Menshikov et al.³⁰ proposed to consider PD in Markov strategies. A symmetric totally mixed Nash equilibrium was found for this game. However, this equilibrium better fits strategies prior to socialization than after. Therefore, we developed a new model that can describe high-cooperation strategies.

Model

Prisoner's Dilemma game (PD). This work is based on the broadly known PD game. In this game, two participants choose between two strategies: Left or Right for the first participant, and High and Low for the second. The choices are simultaneous and independent from each other. Payoffs correspond to the following payoff matrix (see Table 1), which was employed in the model and experiments. Left and High strategies represent Cooperation, Right and Low represent Defection.

Nash equilibrium for PD. PD has one Nash equilibrium: it is a mutual choice of Defection strategy that gives a payoff of 1 for two players. However, laboratory experiments show that people under some conditions avoid Nash equilibrium^{9,30}. For example, individuals under social framing may more frequently choose the Cooperation strategy, behavior that could be considered irrational. For this reason, it would be interesting to discern a theoretical concept underlying this specific behavior.

Nash equilibrium for PD in Markov strategies. Several researchers^{7,9,32,33} argue that for some subjects, social context led to an increase in cooperative choices of up to 100%. Therefore, the behavior under social context is far from a Nash equilibrium. One way to describe this cooperative behavior is to consider the PD game in Markov strategies.

Consider two participants $i \in \{1, 2\}$. Let us denote the probability of cooperation in round t for the first participant as $p_1^c(t)$. We describe participants' behavior by means of the following two quantities: (1) γ – *mutual cooperation* (the probability of cooperative choice as the response to the opponent's cooperative choice on the previous round); (2) α – *tolerance to defection* (the probability of cooperative choice as the response to the opponent's defection choice on the previous round). Then, we assumed that the decision in round t depends only on the results in round $t - 1$. Thus, two variables γ and α imply that individuals' strategies at round $t - 1$ completely determine their behavior in round t . This model will be referred to as PD in Markov strategies^{30,34}. For brevity, we will refer to subsequent γ and α as Markov strategies. The dynamics of participants' actions can be presented as follows:

$$\begin{cases} p_1^c(t) = \gamma_1 p_2^c(t-1) + \alpha_1 (1 - p_2^c(t-1)), \\ p_2^c(t) = \gamma_2 p_1^c(t-1) + \alpha_2 (1 - p_1^c(t-1)). \end{cases} \quad (1)$$

In a stationary state, we have:

$$\begin{cases} p_1^c = \gamma_1 p_2^c + \alpha_1 (1 - p_2^c), \\ p_2^c = \gamma_2 p_1^c + \alpha_2 (1 - p_1^c), \end{cases} \quad (2)$$

where p_1^c and p_2^c are stationary probabilities of cooperation.

According to the paper³⁰ payoff function for participant 1 takes the following form:

$$U(p_1^c, p_2^c) = 5 * p_1^c * p_2^c + 0 * p_1^c * (1 - p_2^c) + 10 * (1 - p_1^c) * p_2^c + 1 * (1 - p_1^c) * (1 - p_2^c) = -4 p_1^c p_2^c - p_1^c + 9 p_2^c + 1. \quad (3)$$

Menshikov et al.³⁰ found a symmetric (whereby $\gamma_1 = \gamma_2 = \gamma$ and $\alpha_1 = \alpha_2 = \alpha$) totally mixed Nash equilibrium for PD in Markov strategies in explicit form. This equilibrium can be represented as the points (α, γ) that meet the following equation:

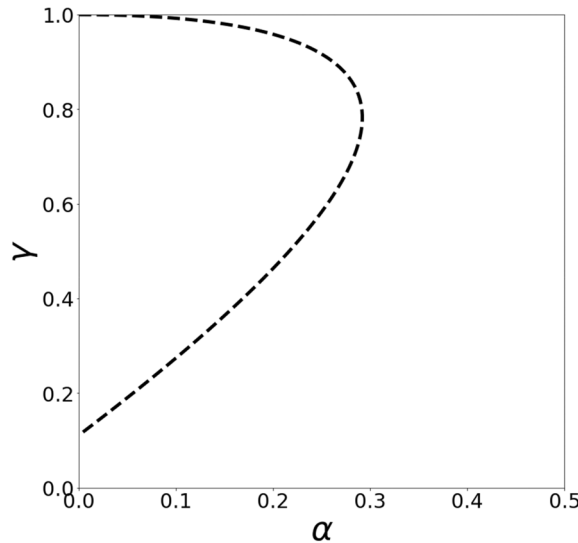


Figure 1. Symmetric totally mixed Nash equilibrium for PD in Markov strategies.

$$5\alpha^2 + 9\gamma^2 - 14\alpha\gamma - 10\gamma + 1 = 0 \tag{4}$$

located in the unit square (see Fig. 1). We will refer to this equilibrium as a Nash equilibrium and this curve as a Nash equilibrium curve.

It is evident from Fig. 1 that curve (4) exists in the area characterized by relatively small values of α (the tolerance to defection does not exceed 0.3). However, experimental results (see “Experimental results” section) reveal that tolerance to defection could exceed 0.5. Also, recent studies indicate that tolerance plays a highly important role in promoting cooperation within a population^{35–37}. To reconcile this problem, we derive quantal response equilibrium for PD in Markov strategies in the next section.

Quantal response equilibrium (QRE). The QRE model was established to explain participants’ observed behavior in laboratory experiments that differs significantly from the Nash equilibrium²³. QRE is an internally consistent equilibrium model in the sense that the quantum response functions are based on the distribution of equilibrium probabilities in the choice of strategies of opponents, not simply on arbitrary beliefs that players may have about these probabilities. One feature of this model is that it allows the possibility of player error. QRE requires that expectations must match an equilibrium choice of probabilities. However, in contrast to the classical Nash equilibrium, the definition of QRE assumes that participants strive for the best answer only in the probabilistic sense: the better the answer, the more likely the participant will choose it^{38,39}. A comparison of the QRE with experimental data indicated that this approach provided a better fit than the Nash equilibrium⁴⁰. In practice, the QRE is dependent upon employing logistic distribution. The answer s_i to the mixed strategy s_{-i} of the remaining players (the probability of choosing strategy s_i) is expressed through the following formula:

$$\sigma_i(s_i|s_{-i}) = \frac{e^{\lambda U_i(s_i,s_{-i})}}{\sum_{s'_i} e^{\lambda U_i(s'_i,s_{-i})}}, \tag{5}$$

where λ – is the parameter of a participant’s rationality, and $U_i(s_i, s_{-i})$ is the expected gain of participant i when the strategies of other players s_{-i} and the strategy s_i of participant i are given. Therefore, when $\lambda \rightarrow 0$ (low rationality), participants chose strategies randomly. When $\lambda \rightarrow \infty$ (high rationality), participants chose strategies with no errors (with the highest expected payoff).

The paper³⁰ demonstrated that the concept of QRE for a simple PD game works only when the probabilities of cooperative choices do not exceed 50%. Here we present a QRE for PD in Markov strategies. Consider $\{\alpha_1, \gamma_1\}$ – Markov strategies of player 1, and $\{\alpha_2, \gamma_2\}$ – Markov strategies of player 2, $\{\lambda_1, \lambda_2\}$ – players’ rationalities.

From the definition of QRE (5):

$$\begin{cases} \alpha_1 = \frac{e^{\lambda_1 U(\alpha_1, \gamma_1, \alpha_2, \gamma_2) |_{\alpha_1=1}}}{e^{\lambda_1 U(\alpha_1, \gamma_1, \alpha_2, \gamma_2) |_{\alpha_1=0}} + e^{\lambda_1 U(\alpha_1, \gamma_1, \alpha_2, \gamma_2) |_{\alpha_1=1}}}, \\ \gamma_1 = \frac{e^{\lambda_1 U(\alpha_1, \gamma_1, \alpha_2, \gamma_2) |_{\gamma_1=1}}}{e^{\lambda_1 U(\alpha_1, \gamma_1, \alpha_2, \gamma_2) |_{\gamma_1=0}} + e^{\lambda_1 U(\alpha_1, \gamma_1, \alpha_2, \gamma_2) |_{\gamma_1=1}}}, \end{cases} \tag{6}$$

where $\{\alpha_1, \gamma_1\} \in [0; 1]$ and $\{\alpha_2, \gamma_2\} \in [0; 1]$ are unknown, $\lambda_1 \in [0; +\infty)$ is fixed, and $U(\alpha_1, \gamma_1, \alpha_2, \gamma_2)$ is the payoff function. Also, we should mention that $\alpha_1 = 0, \alpha_1 = 1, \gamma_1 = 0, \gamma_1 = 1$ are the pure strategies of player 1. To solve system (6), we switched to the symmetrical case by assuming $\alpha_1 = \alpha_2 = \alpha, \gamma_1 = \gamma_2 = \gamma, \lambda_1 = \lambda_2 = \lambda$. Next, we determined the equations of payoff function for different cases.

From system (2), we established the following expressions for stationary state probabilities of cooperation:

$$p_1^c = \frac{\alpha_1 - \alpha_2(\alpha_1 - \gamma_1)}{1 - (\alpha_1 - \gamma_1)(\alpha_2 - \gamma_2)}, \quad p_2^c = \frac{\alpha_2 - \alpha_1(\alpha_2 - \gamma_2)}{1 - (\alpha_1 - \gamma_1)(\alpha_2 - \gamma_2)}. \tag{7}$$

Then, we established the profiles of pure strategies for both players when the others' pure strategies were fixed. To do this, we fixed the profile of strategies $\alpha_2, \gamma_2, \gamma_1$ and calculated the probabilities of cooperative choice for pure strategies $\alpha_1 = 0$ and $\alpha_1 = 1$.

For $\alpha_1 = 0$ we expressed:

$$p_1^c |_{\alpha_1=0} = \frac{\alpha_2 \gamma_1}{1 + \gamma_1(\alpha_2 - \gamma_2)}, \quad p_2^c |_{\alpha_1=0} = \frac{\alpha_2}{1 + \gamma_1(\alpha_2 - \gamma_2)}.$$

After substituting $\alpha_1 = 1$ into (6), we obtained:

$$p_1^c |_{\alpha_1=1} = \frac{1 - \alpha_2 + \alpha_2 \gamma_1}{1 + (1 - \gamma_1)(\alpha_2 - \gamma_2)}, \quad p_2^c |_{\alpha_1=1} = \frac{\gamma_2}{1 - (\alpha_2 - \gamma_2)(1 - \gamma_1)}.$$

Secondly, we fixed the profile of strategies $\alpha_2, \gamma_2, \alpha_1$ and obtained the probabilities of cooperative choice for pure strategies $\gamma_1 = 0$ and $\gamma_1 = 1$.

For $\gamma_1 = 0$:

$$p_1^c |_{\gamma_1=0} = \frac{\alpha_1 - \alpha_2 \alpha_1}{1 - \alpha_1(\alpha_2 - \gamma_2)}, \quad p_2^c |_{\gamma_1=0} = \frac{\alpha_2 - \alpha_2 \alpha_1 + \alpha_1 \gamma_2}{1 - \alpha_1(\alpha_2 - \gamma_2)}.$$

When $\gamma_1 = 1$:

$$p_1^c |_{\gamma_1=1} = \frac{\alpha_1 + \alpha_2 - \alpha_2 \alpha_1}{1 - (\alpha_2 - \gamma_2)(\alpha_1 - 1)}, \quad p_2^c |_{\gamma_1=1} = \frac{\alpha_2 - \alpha_2 \alpha_1 + \alpha_1 \gamma_2}{1 - (\alpha_2 - \gamma_2)(\alpha_1 - 1)}.$$

In the symmetrical case, we have $\alpha_1 = \alpha_2 = \alpha, \gamma_1 = \gamma_2 = \gamma$, and $\lambda_1 = \lambda_2 = \lambda$. The probabilities of cooperative choice then took the following forms:

$$p_1^c |_{\alpha=0} = \frac{\alpha \gamma}{1 + \gamma(\alpha - \gamma)}, \quad p_2^c |_{\alpha=0} = \frac{\alpha}{1 + \gamma(\alpha - \gamma)}, \tag{8}$$

$$p_1^c |_{\alpha=1} = \frac{1 - \alpha + \alpha \gamma}{1 - (1 - \gamma)(\alpha - \gamma)}, \quad p_2^c |_{\alpha=1} = \frac{\gamma}{1 - (1 - \gamma)(\alpha - \gamma)}, \tag{9}$$

$$p_1^c |_{\gamma=0} = \frac{\alpha - \alpha^2}{1 - \alpha(\alpha - \gamma)}, \quad p_2^c |_{\gamma=0} = \frac{\alpha - \alpha^2 + \alpha \gamma}{1 - \alpha(\alpha - \gamma)}, \tag{10}$$

$$p_1^c |_{\gamma=1} = \frac{2\alpha - \alpha^2}{1 - (\alpha - 1)(\alpha - \gamma)}, \quad p_2^c |_{\gamma=1} = \frac{\alpha - \alpha^2 + \alpha \gamma}{1 - (\alpha - 1)(\alpha - \gamma)}. \tag{11}$$

Note that expressions like $(\dots) |_{\alpha=0}$ in formulas (8)–(11) are more formal notations than traditional substitution. From this perspective, it is not problematic that expression for $p_1^c |_{\alpha=0}$ includes α . Additionally, despite the symmetrical assumption, formulas (8)–(11) are naturally asymmetric as expressions for p_1^c and p_2^c are different. If the symmetrical assumption was made prior to substituting pure strategies into resulted expressions, a significantly poorer set of (symmetric) stationary state cooperation probabilities would result. Namely, the symmetrical assumption would transform (7) into (for clarity, we utilize capital letters):

$$P_1^c = P_2^c = P^c = \frac{\alpha - \alpha(\alpha - \gamma)}{1 - (\alpha - \gamma)(\alpha - \gamma)}.$$

From there, we derived the following expressions:

$$P^c |_{\alpha=0} = 0, \quad P^c |_{\alpha=1} = \frac{1}{2 - \gamma}, \quad P^c |_{\gamma=0} = \frac{\alpha}{1 + \alpha}, \quad P^c |_{\gamma=1} = 1. \tag{12}$$

It is evident that formula (12) occur only in the extreme cases of strategies (8)–(11). For example, it is true that $P^c |_{\alpha=0} = p_1^c |_{\alpha=0} = p_2^c |_{\alpha=0}$ if we use $\alpha = 0$ in (8).

Next, we established the payoff functions using (3) for specific strategies by substituting the found expressions (8)–(11) of probabilities:

$$U |_{\alpha=0} = -4p_1^c |_{\alpha=0} p_2^c |_{\alpha=0} - p_1^c |_{\alpha=0} + 9p_2^c |_{\alpha=0} + 1 = \frac{-4\alpha^2 \gamma}{(\gamma(\alpha - \gamma) + 1)^2} - \frac{\alpha \gamma}{(\gamma(\alpha - \gamma) + 1)} + \frac{9\alpha}{(\gamma(\alpha - \gamma) + 1)} + 1, \tag{13}$$

$$U |_{\alpha=1} = \frac{-4\gamma(-\alpha(1 - \gamma) + 1)}{(-(\alpha - \gamma)(1 - \gamma) + 1)^2} + \frac{9\gamma}{(-(\alpha - \gamma)(1 - \gamma) + 1)} - \frac{(-\alpha(1 - \gamma) + 1)}{(-(\alpha - \gamma)(1 - \gamma) + 1)} + 1, \tag{14}$$

$$U|_{\gamma=0} = \frac{-(\alpha^2 + \alpha)}{(-\alpha(\alpha - \gamma) + 1)} - \frac{4(-\alpha^2 + \alpha)(-\alpha(\alpha - \gamma) + \alpha)}{(-\alpha(\alpha - \gamma) + 1)^2} + \frac{9(-\alpha(\alpha - \gamma) + \alpha)}{(-\alpha(\alpha - \gamma) + 1)} + 1, \quad (15)$$

$$U|_{\gamma=1} = \frac{-\alpha(\alpha - 1) + \alpha}{(-\alpha - 1)(\alpha - \gamma) + 1} - \frac{4(-\alpha(\alpha - 1) + \alpha)(-\alpha(\alpha - \gamma) + \alpha)}{(-\alpha - 1)(\alpha - \gamma) + 1)^2} + \frac{9(-\alpha(\alpha - \gamma) + \alpha)}{(-\alpha - 1)(\alpha - \gamma) + 1} + 1. \quad (16)$$

Finally, we used the found expressions (13)–(16) of the payoff functions in system (6) that gave:

$$\begin{cases} \alpha = \frac{e^{\lambda U}|_{\alpha=1}}{e^{\lambda U}|_{\alpha=0} + e^{\lambda U}|_{\alpha=1}}, \\ \gamma = \frac{e^{\lambda U}|_{\gamma=1}}{e^{\lambda U}|_{\gamma=0} + e^{\lambda U}|_{\gamma=1}}, \end{cases} \quad (17)$$

where $\alpha \in [0; 1]$ and $\gamma \in [0; 1]$ are unknown, $\lambda \in [0; +\infty)$ is fixed, and U is the payoff function. Note that different values of the rationality could lead to different profiles of strategies.

We propose to solve system (17) numerically, reducing it to finding (as feasibly as possible) the optimal solution for the following optimization problem:

$$\begin{cases} \min_{\alpha, \gamma} \left(\frac{e^{\lambda U}|_{\alpha=1}}{e^{\lambda U}|_{\alpha=0} + e^{\lambda U}|_{\alpha=1}} - \alpha \right)^2 + \left(\frac{e^{\lambda U}|_{\gamma=1}}{e^{\lambda U}|_{\gamma=0} + e^{\lambda U}|_{\gamma=1}} - \gamma \right)^2, \\ \alpha \in [0; 1], \\ \gamma \in [0; 1]. \end{cases} \quad (18)$$

To solve optimization problem (18), we used the Python package *minimize* from *scipy.optimize*⁴¹. We solved the problem (18) for different values of λ shifting from $\lambda = 0$ to $\lambda = 200$ with the step of 0.01. As the result, the solution of optimization problem (18) gave us the set of points (α, γ) for values of $\lambda \in [0; 200]$ – sought-for QRE for PD in Markov strategies. To prove the adequacy of the optimization method, we investigated the behavior of the objective function:

$$f = \left(\frac{e^{\lambda U}|_{\alpha=1}}{e^{\lambda U}|_{\alpha=0} + e^{\lambda U}|_{\alpha=1}} - \alpha \right)^2 + \left(\frac{e^{\lambda U}|_{\gamma=1}}{e^{\lambda U}|_{\gamma=0} + e^{\lambda U}|_{\gamma=1}} - \gamma \right)^2. \quad (19)$$

In Fig. 2, we illustrate the contour lines of objective function (19) for different values of rationality λ with the solutions of optimization problem (18) for the same value of rationality (yellow points) and with the Nash equilibrium curve. We noticed that for small values of $\lambda < \sim 4$, objective function (19) has a unique local minimum. Additionally, the optimization method found the solution within these areas of the local minimum. However, for $\lambda > \sim 4$, objective function (19) contained more than one local minimum and most of these local minimums were located within the Nash equilibrium curve. For $\lambda \in (\sim 4, \sim 7)$, the optimization method found the global minimum inside one of the local minimums. For $\lambda > \sim 7$, the results of the optimization method started to shift to the area within the standard Nash equilibrium $\alpha = 0, \gamma = 0$ (defection/defection) and finally reached a point outside the Nash equilibrium curve in local minimums of objective function (19). This result corresponds to the theory, as the Nash equilibrium describes the behavior of fully rational participants²³. Therefore, we can conclude that the chosen optimization method worked sufficiently.

In Fig. 3, we plot the arrangement of the obtained QRE for PD in Markov strategies (the solution of (18)) and the Nash equilibrium curve. The QRE for PD in Markov strategies forms an almost near smooth curve in the range of small λ (approximately less than 5). For these values of rationality, objective function (19) consistently showed a unique global minimum that was perfectly caught by the solver. In the middle range of λ (approximately located at the interval [5, 7.08]), the solution of optimization problem (18) approaches the Nash equilibrium curve that corresponds to the theory²³. Nonetheless, at these levels of rationality, the QRE for PD in Markov strategies curve (further, QRE curve) loses its smoothness and solutions “leapfrog” on the Nash equilibrium curve (see the blue triangles on Fig. 3). This could be the result of the optimization method weakness (in Fig. 2, we demonstrate that when $\lambda > \sim 5$ objective function (19) has many local minimums). However, the exact “first” intersection between the QRE curve and the Nash equilibrium curve (which is located near $\alpha \approx 0.2, \gamma \approx 0.5$ and is derived under $\lambda \approx 5$) fits the experimental data best when compared to other intersections (see “Experimental results” section). For large values of the rationality ($\lambda > 7.08$), solutions of (18) converge to the point $\alpha = 0, \gamma = 0$ which does not belong to the Nash equilibrium curve. Instead, it marks the strategies’ profiles of the standard Nash equilibrium (defection/defection).

Experimental results

In this section, we evaluate the equilibrium found against the data from laboratory experiments which were presented in several publications^{9,29,30}. The primary goal of these experiments was to identify the effect of socialization on the level of cooperation choice in the PD game.

The full description of the experiments (N = 14) can be found in Supplementary 1. The following is a schematic representation of the experimental design:

1. The sample for one experiment included 12 recruited participants (all strangers).
2. Participants played iterated PD (Table 1) in a mixed-gender group of 12 participants for 11–22 rounds.

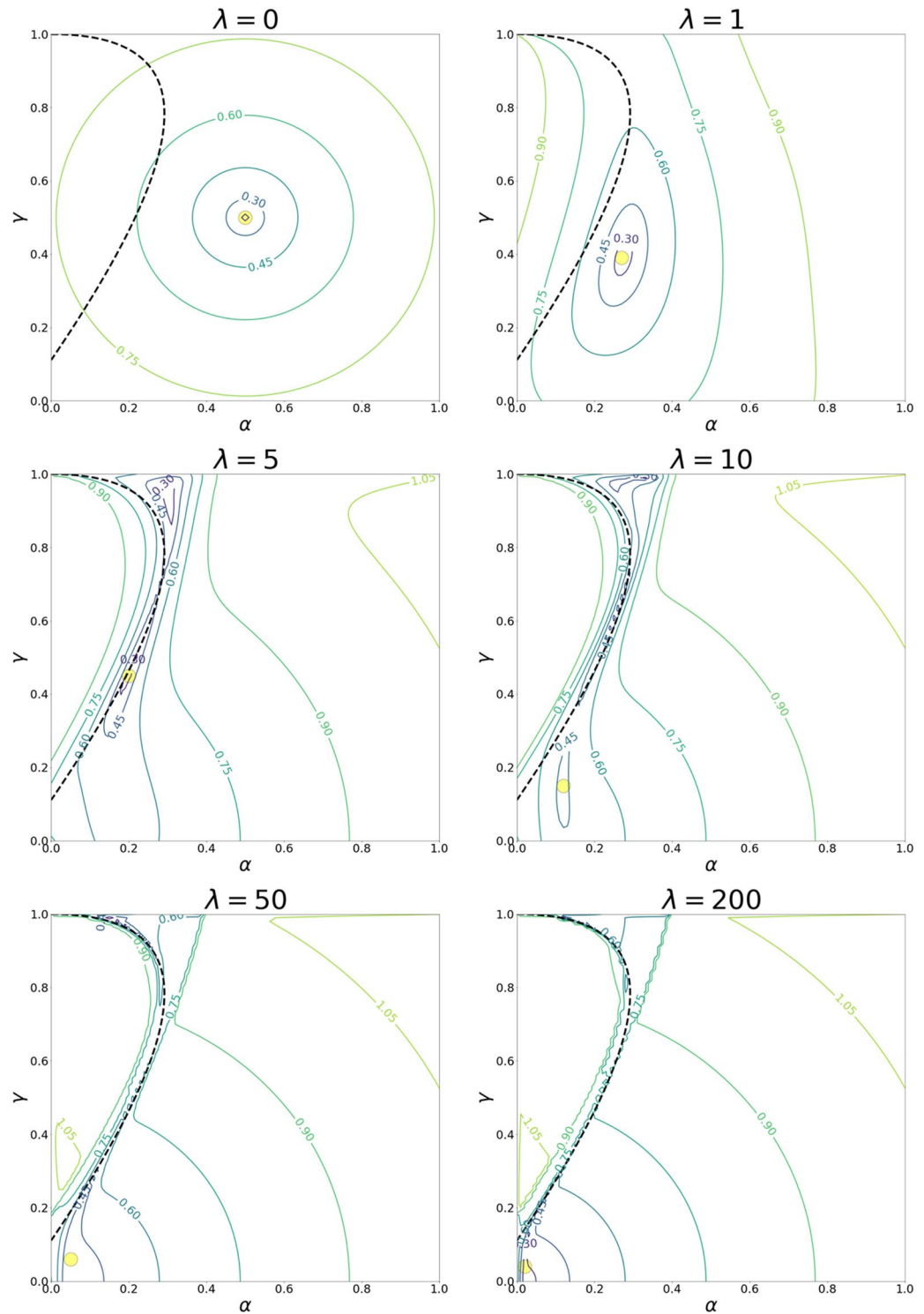


Figure 2. The panels plot contour lines of objective function (19) for different values of rationality (the objective function is shown to the power of 1/5 for demonstrative purposes). The dashed line represents the Nash equilibrium for PD in Markov strategies. The yellow circles symbolize solutions to corresponding optimization problem (18).

3. Socialization occurred across unacquainted group members, who were divided into two groups of 6 participants.
4. Participants played iterated PD (Table 1) in the newly formed groups for 15–20 rounds.

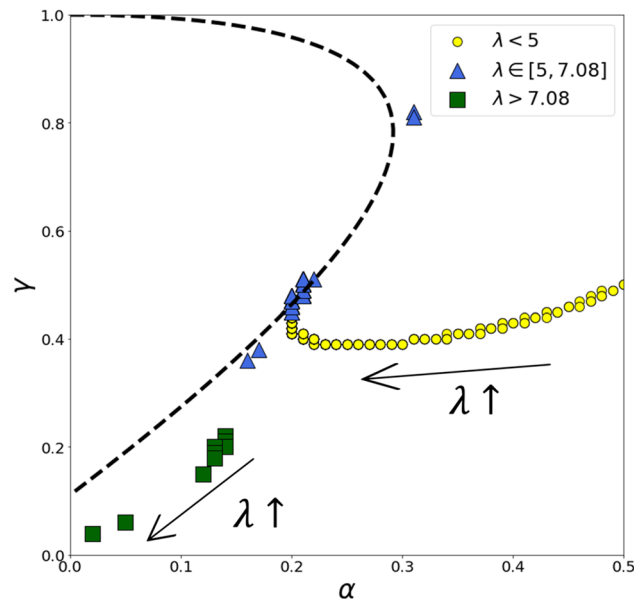


Figure 3. QRE for PD in Markov using solving optimization problem (18) for different values of rationality. The arrows indicate the directional growth of rationality. The resulting QRE curve contains three branches that correspond to various ranges of rationality. The dashed line represents the Nash equilibrium.

5. Participants were compensated for the experiment.

The study procedures involving human participants were approved by the Skolkovo Institute of Science and Technology (Skoltech) Human Subjects Committee. Written informed consents were obtained from participants. All methods were performed in accordance with the relevant guidelines and regulations.

In Supplementary Table S1 (Supplementary 2), we present aggregated results of the 14 experiments (168 participants). We find that the choice of cooperation is higher after socialization (58%) rather than before (22%). We assume that socialization compensates for the irrationality of these choices. This implies that despite the expectation that the payoff of defection is higher than that of cooperation, the utility of sociality is higher than the probable losses of the cooperation choice. In comparing theoretical results with the experimental data, we found for every part of the experiments probabilities of mutual cooperation (γ) and tolerance to defection (α) in all parts of the experiments. Thus, we analyzed the aggregating strategies by participants that represent the most popular strategies in the observed part of the experiment. Due to the experimental design, we aggregated strategies by all 12 participants for every experiment before socialization. We aggregated strategies by two socialized groups of six participants for every experiment after socialization (see Supplementary Table S1, Supplementary 2).

We first analyzed how experimental points corresponded to values of objective function (19) under different levels of rationality (see Fig. 4). We observed that most participants' strategies could be approximated by the minima of the objective function after selecting the appropriate level of rationality. More precisely, we recognized that the behavior of individuals with a high level of cooperation (more than 50%) could be modeled by selecting low rationality rates (which was one of our objectives), whereas low-cooperative participants were well-approximated by high values of rationality. Unfortunately, participants located in the upper right zone of the phase plane are still unexplained. Notably, among all the Nash equilibrium, the best equilibrium that fits the experimental data was found on the initial intersection between the QRE curve and the Nash equilibrium curve (at the rationality level $\lambda \approx 5$). The QRE for PD in Markov strategies for $\lambda < \sim 5$ divided most of the strategies before and after socialization (see Fig. 5), resembling phase boundary.

Discussion and conclusion

In an era that highlights the importance of every individual's choice while promoting living for oneself, it is crucial to consider cooperation as an effective mechanism to promote the overall well-being of society. In this paper, we present a theoretical concept that sufficiently works well enough with high cooperation levels that were previously obtained in the laboratory experiments^{9,30}, that examining the influence of socialization on strategy choices in the PD game. In this paper, we present a symmetrical QRE for PD in Markov strategies (further, QRE-PD-M) and compared it with the symmetrical Nash for PD in Markov strategies³⁰ and experimental results⁹. Under the PD game in Markov strategies, we specify the following: a) instead of pure PD strategies (cooperation or defection), we employ mutual cooperation (the probability of a cooperative choice in response to an opponent's choice in the previous round), and tolerance to defection (the probability of a cooperative choice as the response to an opponent's defection choice in the previous round); and b) the choice of strategy in the current period depends solely on the strategies from the previous period.

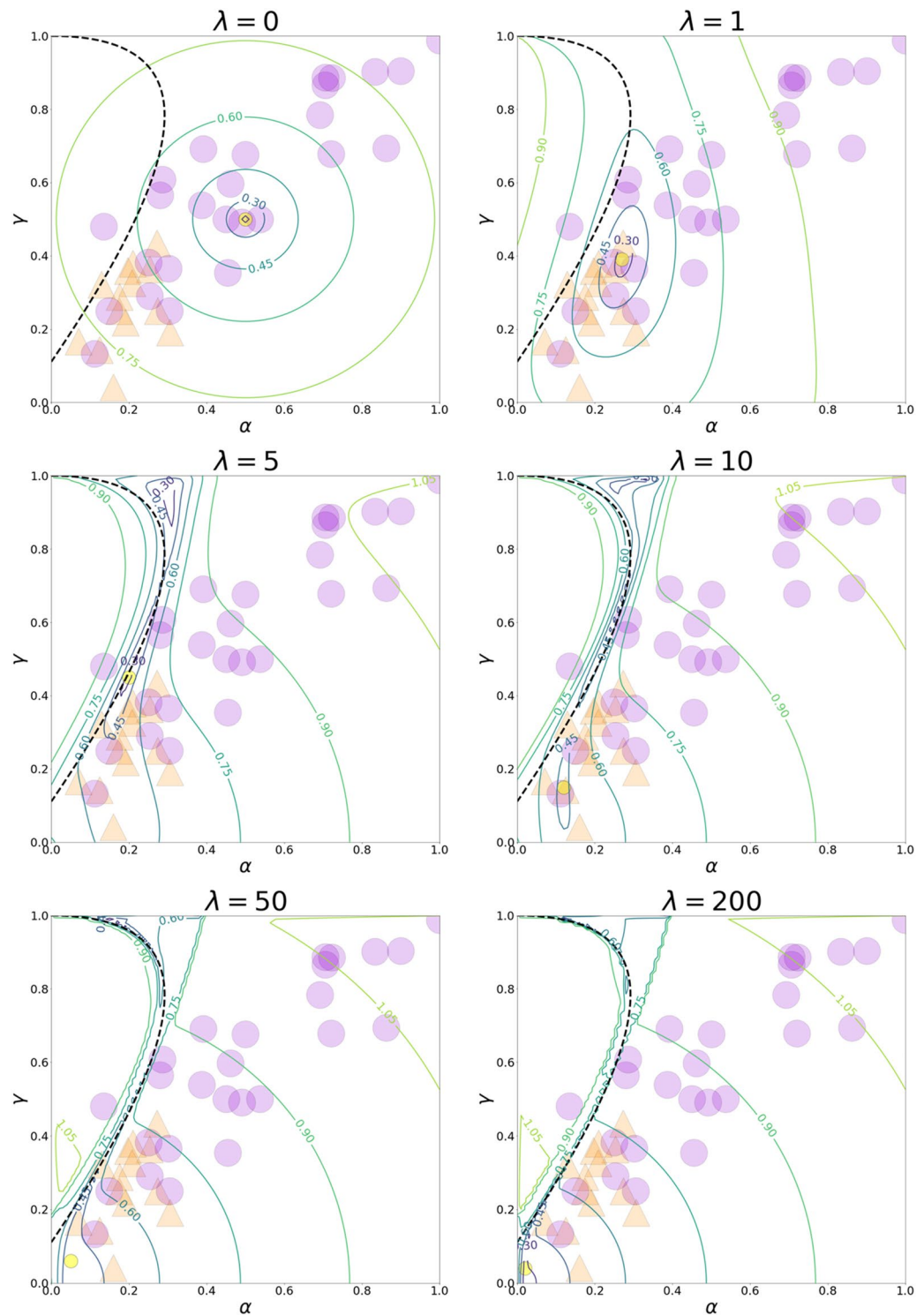


Figure 4. The panels plot contour lines of objective function (19) (the objective function is shown to the power of $1/5$ for demonstrative purposes) for different values of rationality. The dashed line represents the Nash equilibrium. The yellow circles symbolize solutions to the corresponding optimization problems (QRE for PD in Markov strategies). The orange triangles indicate participants' strategies before socialization, while violet circles represent participants' strategies after socialization.

When we matched the QRE-PD-M and the experimental data, we observed that the QRE-PD-M with the low rationality rates described high levels of cooperation after socialization. Conversely, the QRE-PD-M with the high

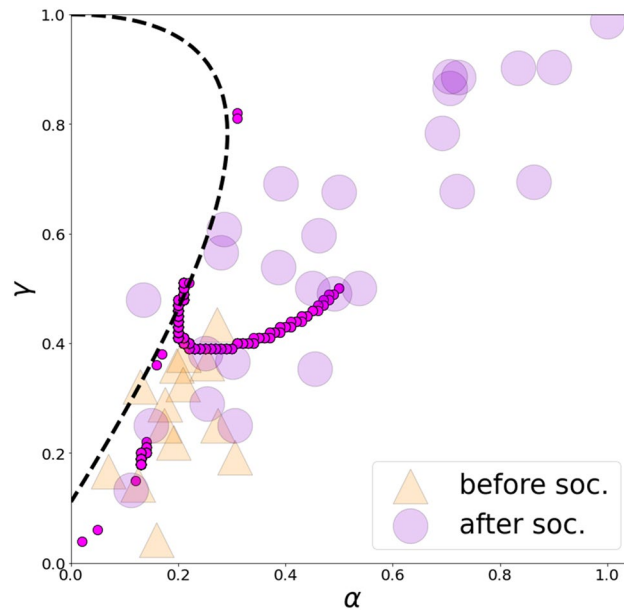


Figure 5. The dashed line represents the Nash equilibrium. The pink circles represent the QRE curve. The orange triangles signify participants' strategies before socialization, while the violet circles signify participants' strategies after socialization. The QRE for PD in Markov strategies points for $\lambda < \sim 5$ serve as a natural border between points before and after socialization.

values of rationality approximates low-cooperative results before socialization. We also found how the QRE-PD-M completes the Nash equilibrium. The intersection between the equilibrium curves under the low parameter of rationality ($\lambda \approx 5$) gives the unique selection of Nash equilibrium, most closely fitting the experimental data the closest (compared to other Nash equilibria for PD in Markov strategies). Additionally, the QRE-PD-M curve (for the parameter of rationality $\lambda < \sim 5$) serves somewhat as a phase boundary for the experimental data before and after socialization; most of the points for before socialization lie below the QRE-PD-M while most of the points for after socialization lie above. However, this result may be coincidental.

Our study is not without limitations. They are as follows:

1. We found and described only the symmetrical case of the QRE-PD-M. Symmetricity allows us to avoid some mathematical difficulties. However, by doing so, we likely lose some equilibria.
2. The optimization problem was solved using numerical methods that come with possible noise in the solutions.
3. We compared the theoretical results against the experiments that involved only one socialization strategy. Further, these experiments were performed within the same socio-demographic and cultural context.

In this paper, our initial motivation was to explain the high levels of cooperation observed in empirics via the concept of QRE. In a nutshell, we obtained that the found QRE-PD-M curve fits the data well. On the one hand, it means that we succeed in our purpose. On the other hand, one could wonder that our approach is rather meaningless because QRE is drawn upon the idea that individuals may “make errors,” and thus, our explanation of high cooperation via the QRE approach could imply that cooperation is simply “an error”.

To reconcile this problem, we put forward the following hypothesis (which should be carefully tested in future studies!) whereby we attempt to bring together our findings and the modern notion of what is a rational sort of behavior. First, we propose that the segment of the QRE-PD-M curve that is characterized by small values of λ ($\lambda < \sim 5$; hereafter, the small-rationality segment) serves as a natural indicator of a participant's state. This state could be individual (leading to selfish behavior), or social (in the social state, individuals should act unselfishly). Further, we suggest that the difference between these two states lies in an individual's utility function. According to the moral preferences hypothesis, the utility function includes the “moral component” that describes our “internal standards about what is right or wrong in a given situation”²². We suggest that in the individual state, the moral component is not well pronounced and doesn't compensate for the risk of cooperative behavior according to our internal moral standards (e.g., behavior in a group of unfamiliar people). Thus, the rational (from the classical view) sort of behavior should happen (corresponding strategies are located beneath the small-rationality segment). In the social state, this component comes with more effect to the utility function because of our need to belong to the social group⁴² (and behave in a more empathetic and sensitive manner towards the group members). Then, people in the social state prefer to cooperate rather than defect (strategies above the small-rationality segment).

In other words, if individuals' strategies fall beneath the small-rationality segment, then we conclude that the people are in the individual state. If strategies are above the small-rationality segment, then we observe the

social state. In the case the strategies are located near the small-rationality segment, then we hypothesize that the corresponding individuals are in “an intermediate regime.” In our empirical context, the changes in participant behavior were caused by the socialization procedure that made the inter-group relations more solid and thus increased the effect of the moral component. Because of this issue, participants’ strategies had a tendency to move from the individual to the social state.

However, questions remain: (1) why does the small-rationality segment play this “dividing” role and (2) how should we interpret strategies that are located near the small-rationality segment? We propose the following answer: one could think about the small-rationality segment as a zone where individuals experience a phase transition between the individual and social states. In this intermediate regime, the participants’ behavior becomes rather “unpredictable” (people face considerable uncertainties about choosing between selfish and unselfish behavior) and thus may be described meaningfully via the concept of QRE. Depending on individuals’ personal characteristics and the socio-demographic and cultural contexts, and, importantly, on the effectiveness of socialization, individuals may find themselves into individual, intermediate, or social states both before and after socialization. Figure 5 may be considered as evidence to support this statement.

A promising avenue for future studies would be to test our hypothesis. We believe that it could be done in two directions: (1) theoretical—by deriving corresponding models that consider a utility function from which the small-rationality segment and the corresponding unpredictable regime should come as a special case that marks the phase transition between individual and social states and (2) empirical—by conducting more experiments involving different socio-demographic and cultural contexts as well as various socialization strategies. Apart from socialization^{9,43}, the phase shift could be triggered by nudging personal norms, turning on morality, and other methods presented by Capraro and Perc²². These experiments may reveal whether the small-rationality segment serves as a boundary between the individual and social states.

If the theory is found to be true, then this segment can be considered as a universal (and, importantly, theoretically motivated) indicator of whether the social relationships in a group of individuals are solid enough. This question is extremely important because solid relationships ensure that the group members are open to cooperative actions and the group can solve complex tasks that require high levels of collaboration. From this perspective, our results could be implemented in team-building activities. For example, one could use the PD game to analyze how participants act before and after implementing team-building tasks. If the tasks were implemented successfully, then it is reasonable to expect that the participants chose strategies above the small-rationality segment afterwards. Further, this methodology could also be used to compare different team-building strategies.

Received: 21 October 2021; Accepted: 7 March 2022

Published online: 16 March 2022

References

1. COOPERATION | meaning in the Cambridge English Dictionary. <https://dictionary.cambridge.org/dictionary/english/cooperation>. Accessed October 8th 2021.
2. Van Bavel, J. J. *et al.* Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **4**(5), 460–471 (2020).
3. Wu, J. Z., Lu, Z. Y. & Zheng, Y. P. A new game theoretical model of social dilemma. *IFAC Proc.* **23**, 393–396 (1990).
4. Jensen, C., Farnham, S. D., Drucker, S. M. & Kollock, P. The effect of communication modality on cooperation in online environments. In *Proceedings of Conference on Human Factors in Computing Systems*. 470–477. <https://doi.org/10.1145/332040.332478> (2000)
5. Kerr, N. L. & Kaufman-Gilliland, C. M. Communication, commitment, and cooperation in social dilemmas. *J. Pers. Soc. Psychol.* **66**, 513–529 (1994).
6. Van Doesum, N. J., Karremans, J. C., Fikke, R. C., de Lange, M. A. & Van Lange, P. A. M. Social mindfulness in the real world: The physical presence of others induces other-regarding motivation. *Soc. Influ.* **13**, 209–222. <https://doi.org/10.1080/15534510.2018.1544589> (2018).
7. Peshkovskaya, A., Babkina, T. & Myagkov, M. Gender effects and cooperation in collective action: A laboratory experiment. *Ration. Soc.* **31**, 337–353 (2019).
8. Peshkovskaya, A., Babkina, T. & Myagkov, M. Social context reveals gender differences in cooperative behavior. *J. Bioecon.* <https://doi.org/10.1007/s10818-018-9271-5> (2018).
9. Babkina, T. *et al.* Choice of the group increases intra-cooperation. In *CEUR Workshop Proceedings*, Vol. 1627, (2016).
10. Joshi, J., Couzin, I. D., Levin, S. A. & Guttal, V. Mobility can promote the evolution of cooperation via emergent self-assortment dynamics. *PLoS Comput. Biol.* **13**, e1005732 (2017).
11. Li, W. J., Jiang, L. L. & Perc, M. A limited mobility of minorities facilitates cooperation in social dilemmas. *Appl. Math. Comput.* **391**, 125705 (2021).
12. Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H.-J. & Milinski, M. Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci.* **107**, 2962–2966 (2010).
13. Rand, D. G., Arbesman, S. & Christakis, N. A. Dynamic social networks promote cooperation in experiments with humans. *Proc. Natl. Acad. Sci.* **108**, 19193–19198 (2011).
14. Allen, B. *et al.* Evolutionary dynamics on any population structure. *Nature* **544**, 227–230 (2017).
15. Su, Q., Li, A., Wang, L. & Eugene Stanley, H. Spatial reciprocity in the evolution of cooperation. *Proc. R. Soc. B* **286**, 20190041 (2019).
16. Realpe-Gómez, J., Andrighetto, G., Nardin, L. G. & Montoya, J. A. Balancing selfishness and norm conformity can explain human behavior in large-scale prisoner’s dilemma games and can poise human groups near criticality. *Phys. Rev. E* **97**, 042321 (2018).
17. Eshghi, S. *et al.* Mathematical models for social group behavior. In *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)* 1–6 (IEEE, 2017). <https://doi.org/10.1109/UIC-ATC.2017.8397423>.
18. Rand, D. G. *et al.* Social heuristics shape intuitive cooperation. *Nat. Commun.* **5**, 3677 (2014).
19. Nishi, A., Christakis, N. A., Evans, A. M., O’Malley, A. J. & Rand, D. G. Social environment shapes the speed of cooperation. *Sci. Rep.* **6**, 29622 (2016).

20. Capraro, V., Polukarov, M., Venanzi, M. & Jennings, N. R. Cooperative equilibrium beyond social dilemmas: Pareto solvable games. (2015).
21. Capraro, V. A solution concept for games with altruism and cooperation. (2013).
22. Capraro, V. & Perc, M. Mathematical foundations of moral preferences. *J. R. Soc. Interface* <https://doi.org/10.1098/rsif.2020.0880> (2021).
23. McKelvey, R. D. & Palfrey, T. R. Quantal response equilibria for normal form games. *Games Econ. Behav.* **10**, 6–38 (1995).
24. Costa-Gomes, M., Crawford, V. P. & Broseta, B. Cognition and behavior in normal-form games: An experimental study. *Econometrica* **69**, 1193–1235 (2001).
25. Camerer, C. F., Ho, T.-H. & Chong, J.-K. A cognitive hierarchy model of games. *Q. J. Econ.* **119**, 861–898 (2004).
26. Stahl, D. O. & Wilson, P. W. Experimental evidence on players' models of other players. *J. Econ. Behav. Organ.* **25**, 309–327 (1994).
27. Bielefeld, R. S. Reexamination of the perfectness concept for equilibrium points in extensive games. In 1–31 (Springer, 1988). https://doi.org/10.1007/978-94-015-7774-8_1.
28. Myerson, R. B. Refinements of the Nash equilibrium concept. *Int. J. Game Theory* **7**, 73–80 (1978).
29. Peshkovskaya, A., Babkina, T. & Myagkov, M. Social context reveals gender differences in cooperative behavior. *J. Bioecon.* <https://doi.org/10.1007/s10818-018-9271-5> (2018).
30. Menshikov, I. S., Shklover, A. V., Babkina, T. S. & Myagkov, M. G. From rationality to cooperativeness: The totally mixed Nash equilibrium in Markov strategies in the iterated Prisoner's Dilemma. *PLoS ONE* **12**, e0180754 (2017).
31. Peshkovskaya, A. G. *et al.* The socialization effect on decision making in the Prisoner's Dilemma game: An eye-tracking study. *PLoS ONE* **12**, e0175492 (2017).
32. Barreda-Tarrazona, I., Jaramillo-Gutiérrez, A., Pavan, M. & Sabater-Grande, G. Individual characteristics vs. experience: An experimental study on cooperation in prisoner's dilemma. *Front. Psychol.* **8**, 596 (2017).
33. Jones, G. Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959–2003. *J. Econ. Behav. Organ.* **68**, 489–497 (2008).
34. Akin, E. The iterated Prisoner's Dilemma: Good strategies and their dynamics. *Ergod. Theory Adv. Dyn. Syst.* <https://doi.org/10.1515/9783110461510-004/HTML> (2016).
35. Fu, F., Wang, L. & Chen, X. Social tolerance allows cooperation to prevail in an adaptive environment Co-diffusion of social contagions View project Evolutionary games on interdependent networks View project Social tolerance allows cooperation to prevail in an adaptive environment. *Artic. Phys. Rev. E* <https://doi.org/10.1103/PhysRevE.80.051104> (2009).
36. Szolnoki, A. & Chen, X. Benefits of tolerance in public goods games. (2015).
37. Szolnoki, A. & Perc, M. Competition of tolerant strategies in the spatial public goods game. *New J. Phys.* **18**, 083021 (2016).
38. Choi, S., Gale, D. & Kariv, S. Social learning in networks: A Quantal Response Equilibrium analysis of experimental data. *Rev. Econ. Des.* **16**, 135–157 (2012).
39. Goeree, J. K., Holt, C. A. & Palfrey, T. R. Regular quantal response equilibrium. *Exp. Econ.* **8**, 347–367 (2005).
40. Zhuang, Q., Di, Z. & Wu, J. Stability of mixed-strategy-based iterative logit quantal response dynamics in game theory. *PLoS ONE* **9**, e105391 (2014).
41. Optimization (scipy.optimize) — SciPy v1.7.1 Manual. <https://docs.scipy.org/doc/scipy/reference/tutorial/optimize.html>. Accessed October 8th 2021.
42. Pickett, C. L., Gardner, W. L. & Knowles, M. Getting a cue: The need to belong and enhanced sensitivity to social cues. *Person. Soc. Psychol. Bull.* <https://doi.org/10.1177/0146167203262085> (2004).
43. Peshkovskaya, A., Myagkov, M., Babkina, T. & Lukinova, E. Do women socialize better? Evidence from a study on sociality effects on gender differences in cooperative behavior. In *CEUR Workshop Proceedings*, Vol. 1968 (2017).

Acknowledgements

The authors are grateful to reviewers for their invaluable comments. This paper is dedicated to the memory of I.S. Menshikov.

Author contributions

I.M. and T.K. conceived and designed the research; T.K. and I.K. analyzed the results; T.K. and I.K. wrote and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-08426-3>.

Correspondence and requests for materials should be addressed to T.S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022