Mutual Cooperation and Tolerance to Defection in the Context of Socialization: The Theoretical Model and Experimental Evidence

T. S. Kozitsina (✉ tatyana.babkina@phystech.edu)
Moscow Institute of Physics and Technology (National Research University)

I. V. Kozitsin
V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences

I. S. Menshikov
Russian Academy of Sciences

Research Article

Keywords: quantal response equilibrium, Nash equilibrium, Prisoner's Dilemma, cooperation, socialization

Posted Date: November 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1003237/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Mutual cooperation and tolerance to defection in the context of socialization: the theoretical model and experimental evidence.

T.S. Kozitsina a,b*, I.V. Kozitsin c,a, I.S. Menshikov a,b

a Moscow Institute of Physics and Technology (National Research University), Moscow, Russian Federation

b Federal Research Center Computer Science and Control, Russian Academy of Sciences, Moscow, Russian Federation

c V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

* Corresponding author
E-mail: tatyana.babkina@phystech.edu

The study of the nature of human cooperation still contains gaps needing investigation. Previous findings reveal that socialization effectively promotes cooperation in the well-known Prisoner’s dilemma (PD) game. However, theoretical concepts fail to describe high levels of cooperation (probability higher than 50%) that were observed empirically. In this paper, we derive a symmetrical quantal response equilibrium (QRE) in PD in Markov strategies and test it against experimental data. Our results indicate that for low levels of rationality, QRE manages to describe high cooperation. In contrast, for high rationality QRE converges to the Nash equilibrium and describes low-cooperation behavior of participants. In the area of middle rationality, QRE matches the curve that represents the set of Nash equilibrium in Markov strategies. Further, we find that QRE serves as a dividing line between behavior before and after socialization, according to the experimental data. Finally, we successfully highlight the theoretically-predicted intersection of the set of Nash equilibrium for PD in Markov strategies and the QRE curve.

Keywords: quantal response equilibrium, Nash equilibrium, Prisoner’s Dilemma, cooperation, socialization

Introduction
Human behavior is still a question and still contains gaps needing investigation. What we know is that our way of thinking, actions, and beliefs depend on many different factors: internal and external. Human behavior includes the important social ability of cooperation, defined by the Cambridge dictionary as “the act of working together with someone or doing what they ask you”1. Perhaps more importantly, cooperation is about sharing mutual profit, equality, costs, and skills. Nowadays, during the pandemic, we realize how important cooperation is for society. It is not about money, but human lives. Examples of cooperation include wearing a mask, keeping social distance, and being patient and generous with public. Thus, studying people’s ability to cooperate helps in making beneficial choices during the world pandemic2.

Despite the evidence of the clear advantages of cooperation, rational to choose defection rather than cooperation when faced with a social dilemma. It is for this reason that these situations are called dilemmas, and studying the factors that lead to cooperation is an important step toward understanding this example of behavioral economics.

Authors make different arguments on which factors may increase cooperation in social dilemmas3; using communication4,5 or socialization6-9, mobility and dynamics10,11, connectivity12, or aspects of an individual’s identity13,14. The choice to cooperate is 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, the authors compared the amount of time that participants in the experiments spent choosing between cooperation and non-cooperation strategies. Their results indicate that quick choice could be a predictor of cooperation. Effects of sociality could also lead to increasing of cooperation. For example, recent articles introduce the cooperative equilibrium for explaining deviation from Nash equilibrium, based on the idea that people have some tendency to cooperate by default.

There are different approaches to shifting strategies from individual to social. The question remains regarding what models can explain irrational cooperation in social dilemmas. Here we list some concepts accepting relatively high cooperation level:

- Quantal response equilibrium
- Level-k
- Cognitive hierarchy
- Quantal level-k
- Trembling hand perfect equilibrium
- Proper equilibrium

Previous studies demonstrate that social interaction significantly increases the cooperation level in iterated Prisoner’s Dilemma games, from a 20% cooperation rate prior to socialization to 53% after socialization. To model such a high level of cooperative strategy choice we require a specific approach. In the paper it was proposed to consider Prisoner’s Dilemma in Markov strategies. For this game, a symmetric totally mixed Nash equilibrium was found. However, this equilibrium better fits strategies prior to socialization than after. Therefore, we developed a new model that is able to describe high-cooperation strategies.

Model

Prisoner’s Dilemma game (PD). This work is based on the broadly-known Prisoner’s Dilemma game. In this game, two participants choose between two strategies: Left or Right for the first strategy, 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 were employed in the model and experiments. Strategies have the next sense: Left and High for Cooperation, Right and Low for Defection.

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

Nash equilibrium for PD. PD has one Nash equilibrium: it is a mutual choice of Defect strategy which gives the payoff of 1 for two players. However, laboratory experiments show that people in some conditions avoid Nash equilibrium. For example, under social framing, individuals may start to choose more frequently the Cooperate strategy, a sort of 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. The papers argue that for some subjects, social context led to the increase in cooperative choices up to 100%. So, the behavior under social context is far from Nash equilibrium. One of the ways to somehow describe the cooperative behavior is to consider Prisoner’s Dilemma game in Markov strategies.

Consider two participants \( i \in \{1, 2\} \). Let us denote the probability to cooperate in round \( t \) for the first participant as \( p_i^C(t) \). We describe participants’ behavior by means of the following
two quantities: (1) \( \gamma \) – *mutual cooperation* (probability of cooperative choice as the respond to cooperative choice of opponent on the previous round); (2) \( \alpha \) – *tolerance to defection* (probability of cooperative choice as the response to a defective choice of opponent on the previous round). These two variables imply that individuals’ strategies at round \( t-1 \) determine completely their behavior at round \( t \). This model will be referred as PD in Markov strategies\(^{26,30}\), and for brevity we will refer to subsequent \( \gamma \) and \( \alpha \) as Markov strategies. Dynamics of participants’ actions can be presented as follows:

\[
\begin{align*}
\frac{p^1_c(t)}{p^2_c(t-1)} & = \gamma \frac{p^1_c(t-1)}{1-p^2_c(t-1)} + \alpha \left(1 - \frac{p^2_c(t-1)}{1-p^1_c(t-1)}\right), \\
\frac{p^2_c(t)}{p^1_c(t-1)} & = \gamma \frac{p^2_c(t-1)}{1-p^1_c(t-1)} + \alpha \left(1 - \frac{p^1_c(t-1)}{1-p^2_c(t-1)}\right).
\end{align*}
\] (1)

In a stationary state, we have:

\[
\begin{align*}
p_1^c & = \gamma p_2^c + \alpha (1-p_2^c), \\
p_2^c & = \gamma p_1^c + \alpha (1-p_1^c),
\end{align*}
\] (2)

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

Payoff function for participant 1 has the following form:

\[
U(p_1^c, p_2^c) = -4p_1^c p_2^c - p_1^c + 9p_2^c + 1. \] (3)

The paper\(^{26}\) found a symmetric (whereby \( \gamma_1 = \gamma_2 = \gamma \) and \( \alpha_1 = \alpha_2 = \alpha \)) totally mixed Nash equilibrium for Prisoner’s Dilemma in Markov strategies in explicit form. This equilibrium can be represented as the points \( (\alpha, \gamma) \) that meet equation

\[
5\alpha^2 + 9\gamma^2 - 14\alpha\gamma - 10\gamma + 1 = 0 \] (4)

and located in the unit square (see Figure 1). Further we will refer to equilibrium as Nash equilibrium and this curve as Nash equilibrium curve.

![Figure 1](image.png)

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

It is evident from Figure 1 that curve (4) exists in the area which is characterized by relatively small values of \( \alpha \) (the tolerance to defection does not exceed 0.3). However, experimental results (Section 4) show that tolerance to defection could be even more than 0.5. Also, recent studies show that tolerance plays a highly important role in promoting cooperation.
inside a population\textsuperscript{31–33}. Attempting to reconcile this problem, we derive QRE for PD in Markov strategies in the next Section.

**Quantal response equilibrium (QRE).** The QRE model was invented to explain the observed behavior of participants in laboratory experiments when it differs significantly from the Nash equilibrium\textsuperscript{19}. 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 of the features of the model is that it allows consideration of “players making mistakes”. QRE imposes a requirement 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\textsuperscript{34,35}. The QRE was compared with experimental data, and determined that this approach provided better fit than the Nash equilibrium\textsuperscript{36}. In practice, the QRE is built upon employing logistic distribution. 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:

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

(5)

where $\lambda$ – is the parameter of participant’s rationality, and $U_i(s_i, s_{-i})$ is the expected gain of participant $i$ when strategies of other players $s_{-i}$ and strategy $s_i$ of participant $i$ are given. Therefore, when $\lambda \to 0$ (low rationality) choices are equally random, and when $\lambda \to \infty$ (high rationality) participants chose the strategies with the highest expected payoff.

The paper\textsuperscript{26} demonstrated that the concept of QRE for simple PD game works only when the probabilities of cooperative choices do not exceed 50\%. Studying high cooperation, we found 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. We found QRE for this game by fixing both parameters of one of the players (for example, player 2’s parameters $\alpha_2$ and $\gamma_2$) and only one parameter of another player (for example, $\gamma_1$). We observed the remaining parameter under two conditions (for example, $\alpha_1 = 0$ and $\alpha_1 = 1$). It provides the combinations of specific strategies of players 1 and 2. Following this, we switched to the symmetrical case and used $\alpha_1 = \alpha_2$, $\gamma_1 = \gamma_2$, $\lambda_1 = \lambda_2$. Finally, the following system provided us QRE for PD in Markov strategies (the detailed description of QRE finding can be found in Appendix 1):

$$
\begin{align*}
\alpha &= \frac{e^{\lambda U|_{\alpha=1}}}{e^{\lambda U|_{\alpha=0}} + e^{\lambda U|_{\alpha=1}}}, \\
\gamma &= \frac{e^{\lambda U|_{\gamma=1}}}{e^{\lambda U|_{\gamma=0}} + e^{\lambda U|_{\gamma=1}}},
\end{align*}
$$

(6)

where $\alpha \in [0; 1], \gamma \in [0; 1], U$ is the payoff function (expressions for $U|_{\alpha=0}, U|_{\alpha=1}, U|_{\gamma=0}$, and $U|_{\gamma=1}$ are given in Appendix 1), and $\lambda \in [0; +\infty)$ is fixed. We propose to solve system (6) numerically, reducing it to finding (as far as feasible) the optimal solution of the following optimization problem:

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

(7)

Let us first investigate the behavior of the objective function
\[ f = \left( \frac{e^{\lambda U|a=1}}{e^{\lambda U|a=0} + e^{\lambda U|a=1}} - \alpha \right)^2 + \left( \frac{e^{\lambda U|y=1}}{e^{\lambda U|y=0} + e^{\lambda U|y=1}} - \gamma \right)^2. \] (8)

In Fig. 2, we demonstrate how solutions obtained correspond to the contour lines of the objective function for different values of rationality. We observe that for values of \( \lambda \) close to null, the (unique) minimum of the objective function is reached near the point \([0.5, 0.5]\) that corresponds to the sense of \( \lambda \) (under assumption of low rationality, individuals should act at random). With the increasing of \( \lambda \), we notice several local minima gradually shifting to the Nash equilibrium curve. When rationality is high, local minima converge to the Nash equilibrium curve on one hand and the strategies’ profile corresponding to the standard Nash equilibrium \( \alpha = 0, \gamma = 0 \) (defect/defect) on other hand. This result corresponds to the theory, as the Nash equilibrium describes behavior of fully rational actors\(^{19}\).

To solve optimization problem (7), we use the Python package \texttt{minimize} from \texttt{scipy.optimize}\(^{37}\). Fig. 3 plots the arrangement of obtained a symmetrical quantal response equilibrium for PD in Markov strategies which form a near smooth curve in the range of small \( \lambda \) (approximately less than 5). For these values of rationality, the objective function always has a unique global minimum, which is perfectly caught by the solver. In the middle range of \( \lambda \) (approximately in the interval \([5, 7.08]\)), the solution of the optimization problem approaches the Nash equilibrium curve. Nonetheless, at these levels of rationality, the QRE curve loses its smoothness and solutions “leapfrog” on the Nash equilibrium curve (see blue triangles on Fig. 3). For large values of the rationality \( \lambda > 7.08 \), solutions of (7) converge to the point \( \alpha = 0, \gamma = 0 \) which does not belong to Nash equilibrium curve in Markov strategies. Instead, it marks the strategies’ profile of standard Nash equilibrium (defect/defect). According to the theory, intersections of the Nash equilibrium curve and the QRE curve should exist and mark the branch of Nash equilibrium in Markov strategies that appear to have a special significance in experimental data description\(^{19}\). We demonstrate that there are few intersections of these curves (blue triangles, Fig. 3) that could be the result of the optimization method weakness. However, exactly the “first” intersection (which is located near \( \alpha \approx 0.2, \gamma \approx 0.5 \) and is derived under \( \lambda \approx 4 \)) fits the experimental data best when compared to other intersections (Section 4).
Figure 2. The panels plot contour lines of the objective function (for demonstrative purposes we show the objective function to the power of $1/5$) for different values of rationality. The
dashed line represents Nash equilibrium for PD in Markov strategies. Yellow circles represent solutions of the corresponding optimization problem.

Figure 3. QRE for PD in Markov strategies derived using solving optimization problem (6) for different values of rationality. The arrows indicate the direction in which rationality grows. The resulting QRE curve consists of three branches that correspond to different ranges of rationality. The dashed line represents Nash equilibrium for PD in Markov strategies.

Experimental results
In this section, we compare the equilibrium found against the data from laboratory experiments which were presented in the following publications<sup>9,25,26</sup>. The general 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 can be found in Appendix 2. The following is a schematic representation of the experimental design:

1. 12 recruited participants (all strangers).
2. Participants play iterated PD (Table 1) in a mixed-gender group of 12 people for 11-22 rounds.
3. Socialization of unacquainted members of groups. Division of participants into two groups of 6 people.
4. Participants play iterated PD (Table 1) in the newly formed groups for 15-20 rounds.
5. Participants are 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 Table 1 (Appendix 3) we present aggregated results of the experiments. 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 payoff of defection is higher than of cooperation, the utility of sociality is higher than 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 (gamma) and tolerance to defection (alpha) (see Table 1, Appendix 3).

We first analyzed how experimental points correspond to values of objective function (8) under different levels of rationality (see Fig. 4). We observed that most participants’ strategies can be approximated by the minima of the objective function after selecting the appropriate level of rationality. More precisely, we recognized that behavior of individuals with 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. From this perspective, we conclude that socialization reduces the level of rationality. Unfortunately, participants located in the upper right zone of the phase plane are still unexplained. The fact that the QRE curve have intersections with the Nash equilibrium curve means that our results are consistent with the theory. Notably, among all Nash equilibrium in Markov strategies, the most “successful” equilibrium (in fitting experimental data) is one found on the “first” intersection between the QRE curve and Nash equilibrium curve (at the rationality level $\lambda \approx 4$). Interestingly, the QRE equilibrium for $\lambda \sim < 4$ divides strategies before and after socialization (see Fig. 5), resembles phase boundary.
Figure 4. The panels plot contour lines of the objective function (for demonstrative purposes we show the objection function to the power of $1/5$) for different values of rationality. The dashed line represents Nash equilibrium for PD in Markov strategies. Yellow circles
represent solutions of the corresponding optimization problems (QRE). Orange triangles indicate strategies of experiments’ participants before socialization while violet circles represent strategies after socialization.

Figure 5. The dashed line represents Nash equilibrium for PD in Markov strategies. Pink circles represent the QRE curve. Orange triangles signify strategies of experiments’ participants before socialization, while violet circles signify strategies after socialization. QRE points for $\lambda \sim < 4$ serve as a natural border between points before and after socialization.

Conclusion

In the era of highlighting the importance of every individual’s choice while mass media promotes living for oneself, it is crucial to remember cooperation as an effective mechanism to promote well-being of the whole society. Our paper proposes a new theoretical concept to explain the high levels of cooperation (more than 50%) which was previously obtained in the laboratory experiments$^{9,26}$ (examining the influence of socialization on strategies choices in the Prisoner’s Dilemma game). We found a symmetrical quantal response equilibrium for Prisoner’s Dilemma game in Markov strategies (QRE) and compared it with the symmetrical Nash equilibrium for this game$^{26}$ and experimental results$^{9}$. Under the Prisoner’s Dilemma game in Markov strategies, we specify the following: a) instead of pure PD strategies (cooperate or defect), we employ mutual cooperation (the probability of cooperative choice in response to an opponent’s in the previous round) and tolerance to defection (the probability of cooperative choice as the response to an opponent’s defective choice in the previous round); b)
the choice of strategy in the current period depends solely on the strategies from the previous period.

When we matched QRE and experimental data, we observed that high level of cooperation after socialization can be explained by QRE with the low rationality rates. Conversely, low-cooperative results before socialization are well-approximated by high values of rationality. We also found how QRE completes the existing Nash equilibrium for this game. The intersection between the equilibrium curves under the low parameter of rationality ($\lambda \approx 4$) gives the unique selection of Nash equilibrium, fitting the experimental data the closest (comparing to other Nash equilibrium in Markov strategies). Additionally, QRE curve (for parameter of rationality $\lambda < 4$) serves somewhat as a phase boundary for the experimental data before and after socialization: before socialization points lie below QRE while after socialization points lie above. However, we understand the chance this result was found by coincidence. Therefore, one possible continuation of the study is to investigate the possible property of QRE as a phase boundary in two directions: (1) theoretical (by deriving corresponding models) and (2) empirical (by conducting experiments). Finally, we also observed experimental strategies which deviate far from QRE or Nash equilibrium. This indicates that the investigation of theoretical concepts explaining the behavior of high cooperation is still in progress.

**Acknowledgements**
The authors are grateful to reviewers for their invaluable comments.
This research was supported by the grant in RFBR 19-01-00296A.
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.

**References**
1. COOPERATION | meaning in the Cambridge English Dictionary. Available at: https://dictionary.cambridge.org/dictionary/english/cooperation. (Accessed: 8th October 2021)
2. Bavel, J. J. Van et al. Using social and behavioural science to support COVID-19 pandemic response. Nat. Hum. Behav. 2020 45 4, 460–471 (2020).
3. Wu, J. Z., Lu, Z. Y. & Zheng, Y. P. A New Game Theoretical Model of Social Dilemma. IFAC Proc. Vol. 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. Conf. Hum. Factors Comput. Syst. - Proc. 470–477 (2000). doi:10.1145/332040.332478
5. Kerr, N. L. & Kaufman-Gilliland, C. M. Communication, Commitment, and Cooperation in Social Dilemmas. J. Pers. Soc. Psychol. 66, 513–529 (1994).
6. Doesum, N. J. Van, Karremans, J. C., Fikke, R. C., Lange, M. A. de & Lange, P. A. M. Van. Social mindfulness in the real world: the physical presence of others induces other-regarding motivation. https://doi.org/10.1080/15534510.2018.1544589 13, 209–222 (2018).
7. Peshkovskaya, A., Babkina, T. & Myagkov, M. Gender effects and cooperation in collective action: A laboratory experiment. Ration. Soc. 31, (2019).
8. Anastasia Peshkovskaya, Tatiana Babkina, M. M. Social context reveals gender differences in cooperative behavior. J. Bioeconomics 1–13 (2018).
9. Babkina, T. et al. Choice of the group increases intra-cooperation. in CEUR Workshop Proceedings 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. 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).
13. 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).
14. Eshghi, S. et al. Mathematical models for social group behavior. in 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, 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). doi:10.1109/UIC-ATC.2017.8397423
15. Rand, D. G. et al. Social heuristics shape intuitive cooperation. Nat. Commun. 5, 3677 (2014).
16. 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).
17. Capraro, V., Polukarov, M., Venanzi, M. & Jennings, N. R. Cooperative Equilibrium beyond Social Dilemmas: Pareto Solvable Games. (2015).
18. Capraro, V. A solution concept for games with altruism and cooperation. (2013).
19. McKelvey, R. D. & Palfrey, T. R. Quantal response equilibria for normal form games. Games Econ. Behav. 10, 6–38 (1995).
20. Costa-Gomes, M., Crawford, V. P. & Broseta, B. Cognition and behavior in normal-form games: An experimental study. Econometrica 69, 1193–1235 (2001).
21. Camerer, C. F., Ho, T.-H. & Chong, J.-K. A Cognitive Hierarchy Model of Games. Q. J. Econ. 119, 861–898 (2004).
22. Stahl, D. O. & Wilson, P. W. Experimental evidence on players’ models of other players. J. Econ. Behav. Organ. 25, 309–327 (1994).
23. Bielefeld, R. S. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. in 1–31 (Springer, Dordrecht, 1988). doi:10.1007/978-94-015-7774-8_1
24. Myerson, R. B. Refinements of the Nash equilibrium concept. Int. J. Game Theory 7, 73–80 (1978).
25. Peshkovskaya, A., Babkina, T. & Myagkov, M. Social context reveals gender differences in cooperative behavior. J. Bioeconomics (2018). doi:10.1007/s10818-018-9271-5
26. 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, (2017).
27. Peshkovskaya, A. G. et al. The socialization effect on decision making in the Prisoner’s Dilemma game: An eye-tracking study. PLoS One 12, (2017).
28. 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.* **0**, 596 (2017).

29. Jones, G. Are smarter groups more cooperative? Evidence from prisoner’s dilemma experiments, 1959–2003. *J. Econ. Behav. Organ.* **68**, 489–497 (2008).

30. Akin, E. The iterated Prisoner’s Dilemma: Good strategies and their dynamics. *Ergod. Theory Adv. Dyn. Syst.* **77**–107 (2016). doi:10.1515/9783110461510-004/HTML

31. 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* (2009). doi:10.1103/PhysRevE.80.051104

32. Szolnoki, A. & Chen, X. Benefits of tolerance in public goods games. (2015).

33. Szolnoki, A. & Perc, M. Competition of tolerant strategies in the spatial public goods game. *New J. Phys.* **18**, 083021 (2016).

34. 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).

35. Goeree, J. K., Holt, C. A. & Palfrey, T. R. Regular Quantal Response Equilibrium. *Exp. Econ.* **8**, 347–367 (2005).

36. Zhuang, Q., Di, Z. & Wu, J. Stability of Mixed-Strategy-Based Iterative Logit Quantal Response Dynamics in Game Theory. *PLoS One* **9**, e105391 (2014).

37. Optimization (scipy.optimize) — SciPy v1.7.1 Manual. Available at: https://docs.scipy.org/doc/scipy/reference/tutorial/optimize.html. (Accessed: 8th October 2021)

38. Kozitsina, T. *et al.* Ethnicity and gender influence the decision making in a multinational state: The case of Russia.

39. Fischbacher, U. *z-Tree*: {Zurich} toolbox for ready-made economic experiments. *Exp. Econ.* **10**, 171–178 (2007).
Legends

Table 1
Payoff matrix in Prisoner’s Dilemma game.

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

Figure 2
The panels plot contour lines of the objective function (for demonstrative purposes we show the objection function to the power of 1/5) for different values of rationality. The dashed line represents Nash equilibrium for PD in Markov strategies. Yellow circles represent solutions of the corresponding optimization problem.

Figure 3
QRE for PD in Markov strategies derived using solving optimization problem (6) for different values of rationality. The arrows indicate the direction in which rationality grows. The resulting QRE curve consists of three branches that correspond to different ranges of rationality. The dashed line represents Nash equilibrium for PD in Markov strategies.

Figure 4
The panels plot contour lines of the objective function (for demonstrative purposes we show the objection function to the power of 1/5) for different values of rationality. The dashed line represents Nash equilibrium for PD in Markov strategies. Yellow circles represent solutions of the corresponding optimization problems (QRE). Orange triangles indicate strategies of experiments’ participants before socialization while violet circles represent strategies after socialization.

Figure 5
The dashed line represents Nash equilibrium for PD in Markov strategies. Pink circles represent the QRE curve. Orange triangles signify strategies of experiments’ participants before socialization, while violet circles signify strategies after socialization. QRE points for $\lambda \sim < 4$ serve as a natural border between points before and after socialization.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.docx