Han, X., Cao, S., Bao, J.Z., Wang, W.X., Zhang, B., Gao, Z. Y. y Sánchez, A. (2018). Equal status in Ultimatum Games promotes rational sharing. *Scientific Reports*, 8, 1222.

DOI: https://doi.org/10.1038/s41598-018-19503-x

SCIENTIFIC REPORTS

OPEN

Equal status in Ultimatum Games promotes rational sharing

Xiao Han[1,2], Shinan Cao[3], Jian-Zhang Bao[2], Wen-Xu Wang[2], Boyu Zhang[4], Zi-You Gao[1] & Angel Sánchez[5,6,7,8]

Received: 31 August 2017
Accepted: 18 December 2017
Published online: 19 January 2018

Experiments on the Ultimatum Game (UG) repeatedly show that people's behaviour is far from rational. In UG experiments, a subject proposes how to divide a pot and the other can accept or reject the proposal, in which case both lose everything. While rational people would offer and accept the minimum possible amount, in experiments low offers are often rejected and offers are typically larger than the minimum, and even fair. Several theoretical works have proposed that these results may arise evolutionarily when subjects act in both roles and there is a fixed interaction structure in the population specifying who plays with whom. We report the first experiments on structured UG with subjects playing simultaneously both roles. We observe that acceptance levels of responders approach rationality and proposers accommodate their offers to their environment. More precisely, subjects keep low acceptance levels all the time, but as proposers they follow a best-response-like approach to choose their offers. We thus find that status equality promotes rational sharing while the influence of structure leads to fairer offers compared to well-mixed populations. Our results are far from what is observed in single-role UG experiments and largely different from available predictions based on evolutionary game theory.

The Ultimatum Game (UG) was proposed more than three decades ago[1,2] as a simple and clear way to measure social preferences[3]. In UG experiments, experimenters work with two subjects and give one of them (the "proposer") an amount of money. The proposer makes an offer as to how to split the money to the other player (the "responder"). The responder can only accept the proposal as is or reject it outright, and in case of rejection none of the players receives any money. Clearly, rational people, where the term "rational" is used in the sense of self-interest, will both offer and accept the minimum possible amount, as responders have no incentives to reject any positive amount of money. However, all available experiments provide strong evidence that low offers are often rejected, with low meaning lower than 20–30% of the pot. Correspondingly, it appears that proposers anticipate this behaviour and offer amounts larger than the minimum, with fair splits being frequent. It is worth stressing that in the last three decades literally thousands of experiments have been carried out[4–9] giving the same qualitative results.

A common variant of the standard ultimatum game is that the responder precommits a minimum acceptable offer (MAO) that he/she will accept (and any lower offer will be rejected) rather than simply decides whether to accept a specific offer. This MAO variant is more informative about responders' preferences. Experiments found that the minimum acceptable offers for most subjects are around 30%[6,10–12]. This is consistent with the observations of the standard UG that offers lower than 30% are often rejected[8,9]. In particular, the MAO UG has some relation with the coordination game in the sense that when the proposer's offer equals to the responder's acceptance level, then both of them has no incentive to change their strategies[13].

[1]MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, 100044, P. R. China. [2]School of Systems Science, Beijing Normal University, Beijing, 100875, P. R. China. [3]School of Finance, University of International Business and Economics, Beijing, 100029, P. R. China. [4]Laboratory of Mathematics and Complex Systems, Ministry of Education, School of Mathematical Sciences, Beijing Normal University, Beijing, 100875 P. R. China. [5]Grupo Interdisciplinar de Sistemas Complejos (GISC), Departamento de Matemáticas, Universidad Carlos III de Madrid, 28911 Leganés, Spain. [6]UC3M-BS Institute of Financial Big Data (IFIBID), Universidad Carlos III de Madrid, 28911 Leganés, Spain. [7]Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, Zaragoza, Spain. [8]Unidad Mixta Interdisciplinar de Comportamiento y Complejidad Social (UICCS) UC3M-UV-UZ, Universidad Carlos III de Madrid, 28911 Leganés, Spain. Correspondence and requests for materials should be addressed to B.Z. (email: zhangby@bnu.edu.cn) or Z.-Y.G. (email: zygao@bjtu.edu.cn) or A.S. (email: anxo.sanchez@gmail.com)

These clear, reproducible experimental results have puzzled economists, but also evolutionarily biologists for a long time. Indeed, the fact that human subjects reject positive amounts of money out of anger about what is considered to be unfair is hard to reconcile with the self-interested decisions one would expect from evolutionarily selected species. In this respect, it is interesting to note that this behaviour has been also observed among non-human primates[14]. Therefore, a number of explanations from different disciplines and perspectives have been proposed in order to understand the reasons for this behaviour. Prominent among these are theoretical approaches based on evolutionary game theory[10,15–25]. Most of these studies considered a variant of standard UG, where subjects play both roles and their strategies are given by two parameters p and q. When acting as proposer, the player offers the amount p. When acting as responder, the player rejects any offer smaller than q. Thus, in this variant, subjects play both roles in a MAO UG. Similarly to the standard UG, a rational responder should accept any (non-zero) offer, and therefore, a rational proposer will offer the minimum. It is worth noting that if everyone follows the same strategy in a dual-role UG, then all the subjects will obtain the same payoff, which is a fair outcome. In this paper, we define fairness as fair share in order to keep consistent with previous theoretical and empirical studies.

While these evolutionary game-theoretical models touch upon different aspects of the problem, two factors have been identified in the literature as possibly relevant for the arising of quasi-fair offers and the rejection of unfair offers: First, subjects taking part in the UG in both roles, meaning that they have to be both proposers and responders. Second, subjects interacting with fixed partners during the repeated play (i.e., there is a fixed interaction structure)[16,19,21–25]. In other words, evolutionary game theory on a well-mixed population where everybody plays with everybody else predicts convergence to self-interested behaviour, and only when interactions are restricted to a few, fixed members of the population non-zero offers and acceptance levels arise.

In this paper, we experimentally test the above two factors. Our dual-role repeated UG experiments cover three key issues hitherto unaddressed in the literature. First, it allows for direct comparison to the theoretical predictions. The starting point of most theoretical research based on evolutionary game theory is the dual-role UG: agents act in both roles, and their choices for the acceptance level, q, and the offer, p, evolve with different prescriptions based on their and other subjects' payoffs and/or actions[10,15–25]. We note that some of these studies assumed that players are equally likely to be in either of the two roles and, in this case, the expected payoff of a player is half of that earned when acting in both roles[15–17]. However, in the midst of the abundant experimental literature on the UG, only a few studies consider a setup given by a dual-role UG. These studies typically resort to the strategy method: subjects pick the two values and then they are randomly matched in a specific role in which the game is realized[4,26–30]. Second, to our knowledge, all the available experiments on the dual-role UG are one-shot and have not provided a clear picture of the behaviour in dual-role UGs: Carter and Irons[26] and Weg and Smith[27] found that proposer demands were greater if subjects play both roles. Conversely, Güth and Tietz[4] found that proposers who play both roles make smaller demands than those who do not. Third, aside from an experiment on a bipartite structure of proposers and responders, the effects of population structures on UG experiments are unclear[31]. Although several theoretical studies indicated that a fixed interaction structure can enhance fairness[16,19,21–25], this point has not been tested by laboratory experiments. In addition, a structure setting allows us to get insight into how subjects decide their actions given the information about their neighbours' behaviours in the previous round. Finally, a word is in order about the connection of this work with real contexts. While we do neither claim nor believe that there are actual, specific systems or situations that can be modeled by our experimental setup (or by the corresponding theoretical models for that matter), we do think that the mechanisms we are exploring will indeed be relevant to many such situations in order to provide possible explanations of the observed behaviours.

Results

We conducted a series of repeated dual-role UG experiments to test the effects of playing both roles and population structures on UG. The experiments include 9 treatment groups T1-T9 with structured populations and 2 control groups C1-C2 with well-mixed populations (see Methods, Supplementary Notes 1.1–1.3, Supplementary Table 1 and Supplementary Figures 1–3 for more details). In the 9 treatment groups, every subject plays with four fixed partners. In the 2 control groups, a subject also plays with four subjects, but he/she randomly encounters his/her neighbours in each round. At each round, a subject submits his/her choices for p and q simultaneously ($0 \leq p, q \leq 100$), and plays the standard UG with each of his/her neighbours with the two different roles. Subjects' choices are applied to all their neighbours in a round, i.e., a subject makes the same offer and acceptance level to all of his/her neighbours. Thus, our experimental setting is same as many previous theoretical models[19,21–25].

We begin by analyzing subjects' offer p and acceptance level q. The time evolution, spatio-temporal patterns and distributions of p and q are shown in Fig. 1 and Supplementary Figures 4, 5, respectively, and histograms of p and q at rounds 1, 35 and 70 are shown in Supplementary Figure 6. These figures show that the mean values of p decrease over rounds. To be specific, in both the treatment and control groups, the mean value of p is about 50 at the beginning and it decreases faster in the control groups than the treatment groups (see power law regressions in Supplementary Figure 7). In contrast, the mean value of q keeps constant at about 10 over all the 70 rounds in the treatment groups, and decreases from 10 to 5 in the control groups. The mean values of p and q in the treatment groups are significantly higher than those of the control groups (see Table 1). These results imply that fixed interaction structures have some effects on promoting fairness as compared to well-mixed populations, although p and q in these dual-role experiments are significantly lower than the values found in single-role experiments and in most theoretical models (see a comparison in Supplementary Table 2). Furthermore, the standard deviations of q over all rounds and over the last 35 rounds are similar in both treatment groups (F-test, P-value = 0.5929) and control groups (F-test, P-value = 0.3576), which implies that the diversity of q does not change significantly along rounds. The standard deviation of p over all rounds and over the last 35 rounds in the treatment groups have no significant difference (F-test, P-value = 0.2023), however, in the control groups the

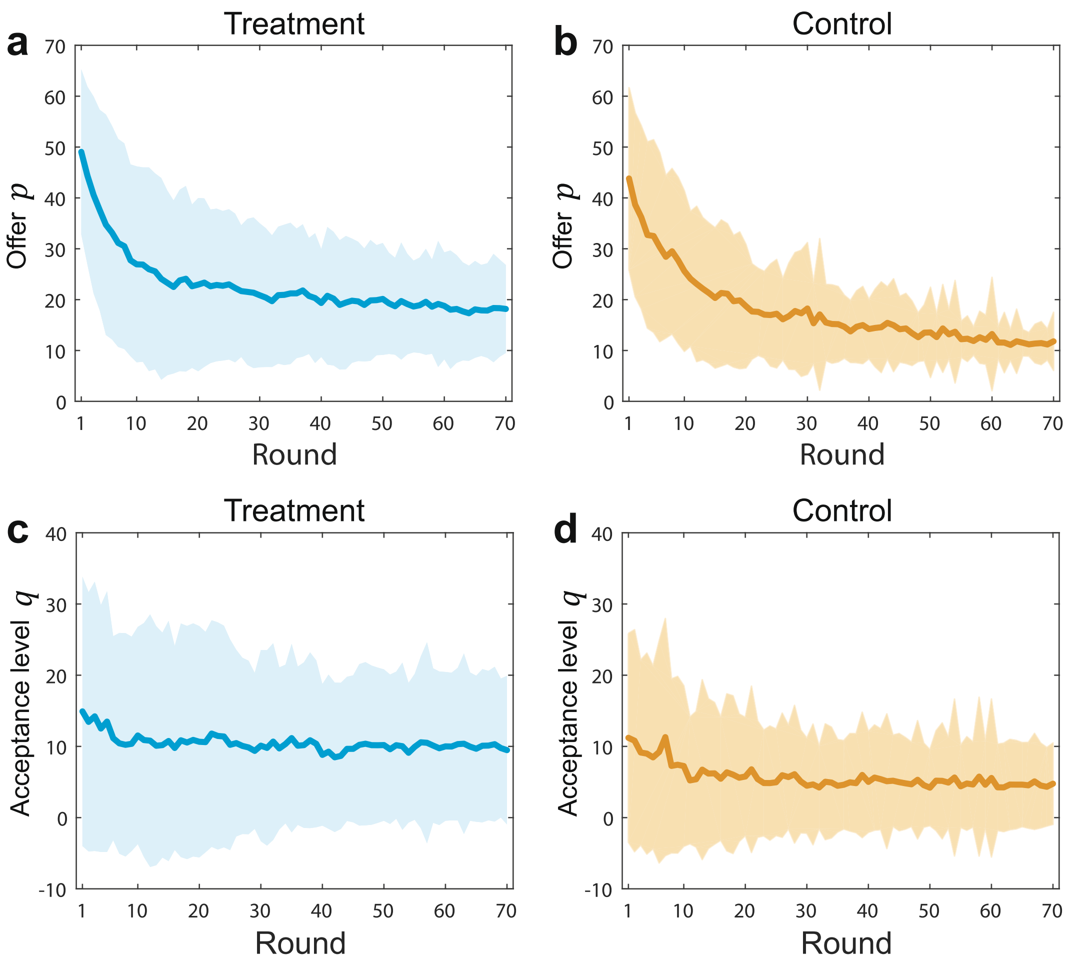


Figure 1. Time evolution of mean values and standard deviations of offers p and acceptance levels q. (**a**,**b**) Mean values and standard deviations of offers p from round 1 to round 70. Fair splits emerge at the beginning of all groups. The mean values of offer p decrease rapidly as the game progresses. The mean values in the treatment groups are larger than the control groups. The result indicates that fixed interaction structures can enhance the fairness compared with well-mixed populations. (**c**,**d**) Mean values and standard deviations of acceptance levels q from round 1 to round 70. Mean values of acceptance levels q are quite low from beginning to end. The mean values of q are stable in the treatment groups while the mean values of q slight decrease in the control groups. The results demonstrate that most of responders are quasi-rational.

	1-70 rounds/1-35 rounds/36-70 rounds	
	Treatment	Control
Mean(p)	22.78/26.37/19.18	17.78/22.52/13.03
Mean(q)	10.46/10.99/9.92	5.70/6/50/4.49
SD(p)	9.23/11.92/8.45	5.40/8.65/3.26
SD(q)	8.19/10.05/8.50	5.82/7.35/5.32

Table 1. The mean values and standard deviations of offers and acceptance levels. We calculate mean values and standard deviations of p and q for all 70 rounds, and separately for rounds 1 to 35 and rounds 36 to 70. Mean(p) and SD(p) represent the mean value and the standard deviation of offers of all proposers, respectively, in which a proposer's offer p is taken as the average of his/her offers p over 1-70 rounds/1-35 rounds/36-70 rounds. Similarly, Mean(q) and SD(q) represent the mean value and the standard deviation of acceptance levels q of all responders, respectively, in which a responder's acceptance level is taken as the average of his/her acceptance levels q over 1-70 rounds/1-35 rounds/36-70 rounds.

standard deviation of p over last 35 rounds is significantly lower than that over all rounds (F-test, P-value < 0.001) (see Table 1). This is consistent with the single-role UG study[31] that a fixed interaction structure promotes the diversity of offers p but not of acceptance levels q.

Furthermore, we find there are positive relationships between mean payoffs and rounds in both treatment groups and control groups, and the average values of payoffs over all rounds are 90.47 and 92.14 in the treatment groups and control groups, respectively. Moreover, there are negative relationship between standard deviations of payoffs and rounds in the treatment groups and control groups (see Supplementary Figure 8). The results of payoffs indicate that coordinating behaviours shown in our experiments can enhance cooperative bonus and reduce payoff differences which lines up with some earlier observations[32,33]. In fact, one could see our setup as a related one to the consensus experiments in[33], where players are incentivized to choose the same colour as their neighbours. However, the fact that there are very many available options to the participants in our setup make it difficult to carry the analogy beyond a general resemblance.

Going from the aggregate to the individual level, let us now look at the behavioural patterns of proposers and responders. In our dual-role UG experiments, the proportion of rational behaviours of responders (i.e., do not reject $q = 0$ or only reject zero offer $q = 1$) is quite high compared with those of single-role experiments. In both the treatment groups and control groups, about half of the responders are rational in the first round (see Supplementary Figures 9–11). Furthermore, we observe that the proportion of rational responders decays over time in the treatment groups, while in the control groups the proportion remains constant (see Supplementary Figure 11). All in all, the proportion and evolution of rational responders is not very different in the treatment and control groups, and therefore the interaction structure does not seem to have large influence on this aspect.

On the other hand, proposers, who in principle should offer the minimum possible amount, seem to behave in a rational, best-response manner, offering the amount that maximizes their payoff given the acceptance levels of their neighbours in the previous round (see Supplementary Note 1.4 for the definition of best-response behaviour). In all groups, the proportions of rational behaviours among proposers are quite low at the beginning, but these proportions increase significantly over rounds and reach about 50 percent at the end (see Supplementary Figures 10,11). Thus, behaviours of proposers show a clear learning trend. As subjects gain experience from repeated observations, they make more precise estimates of the best response offer.

Distributions of individual strategies in the dual-role UG are shown in Supplementary Figure 5. In both the treatment groups and control groups, strategies around (20, 0) and (10, 0) are most popular. Furthermore, in the treatment groups, about 5% of strategies are close to (20, 20), but in the control groups, the proportion is less than 3%. A typical measurement for individual strategies in the dual-role UG is empathy. An empathic individual will offer an amount that is equal to his/her minimum acceptance level[17,20,23,25]. We use this approach to define empathic behaviours ($p = q$) as well as altruistic behaviours ($>q$, i.e., offer more than expected) and selfish behaviours ($p < q$, i.e., offer less than expected). It has been shown in theory that empathy can emerge spontaneously if there is a fixed interaction structure and the role of proposer or responder is randomly changed from round to round[25]. In our experiments, although the proportion of empathic behaviours slightly increases with rounds (which results in the decline of p), the proportion of altruistic behaviours is much higher than the other two types of behaviours (76.37% in the treatment groups and 83.84% in the control groups, see Supplementary Figure 12). Specifically, the difference between p and q keeps stable at about 10 in the second half of the games in all the groups (see Fig. 1 and Table 1). Furthermore, selfish behaviours have the lowest single round average payoff in both treatment groups and control groups. The average payoff of empathic behaviours is slightly higher than that of altruistic behaviours in the treatment groups, but there is no significant difference between the payoffs of altruistic behaviours and empathic behaviours in the control groups (see Supplementary Figure 13).

Although the above analysis is informative as to why p declines but q does not, the reason behind these rational behaviours is still unknown. Previous studies have indicated that individual behaviours in games involving fairness and cooperation are sensitive to decision time, it is then important to investigate the effect of time pressure on decision making[34–37]. Overall, the mean decision time decreases over rounds (see Supplementary Figure 14 and Supplementary Table 3). This means that subjects make decisions faster as they play the game repeatedly. Interestingly, there is a clear positive correlation between actual decision time and proportions of rational behaviours of proposers (Pearson correlation, coefficient $=0.9549$, P-value$=0.0451$, see Fig. 2). Note that a subject's best response choice for p (denoted by BR (p)) depends on his/her neighbours' strategies in the previous round. Therefore, it makes sense that shorter decision times decrease the ratio of rational behaviours since BR (p) may needs to be recalculated at each round. Since the actual decision time decreases over rounds, we further look at the relationship between relative decision time (i.e., actual decision time minus mean decision time in that round) and the proportion of best-response behaviours (Supplementary Figure 15). In T1-T2 and T5-T9, both faster and slower decisions are more likely to be best-response. The reason is simple. If a subject's neighbours' acceptance levels are stable, then his/her optimal offer does not change via rounds so he/she can make a fast decision. In contrast, if the acceptance levels are unstable, then the subject needs more time to calculate the optimal offer so he/she will make a slow decision. Supplementary Figure 14 further reveals that subjects in T3-T4 do not have enough time to make a slow decision. This explains why the proportion of best-response behaviours in T3-T4 is lower than T1-T2 and T5-T9. Moreover, when proposers do not best respond, their offers are often higher than the best-response (i.e., best-response in general leads to smaller p). We find a weak negative correlation between the discrepancy $p - \mathrm{BR}(p)$ and decision time (Pearson correlation, coefficient $= -0.1412$, P-value$=0.0117$). This result agrees well with the observation in Cappelletti *et al.*[35], where proposers are likely to make higher offers under time pressure. In contrast, because a subject's optimal q does not change via rounds (always $q = 0$ or 1), the proportions of rational behaviours of responders are affected little by the decision time.

We now analyze the effects of population size on individual behaviors. Supplementary Tables 4, 5 show that the mean values of q in different treatment groups are similar. In contrast, the mean value of p in T3-T4 is higher than T1-T2 and T5-T9, but the difference between T1-T2 and T5-T9 is not significant. Moreover, we find there is no significant difference in proportion of best-response behaviours between T1-T2 and T5-T9 (see Supplementary Table 6). Overall, subjects in T1-T2 and T5-T9 display very similar behaviours. This implies that the population size does not affect the pattern of individual behaviours. In addition, p in T3-T4 is higher because the average decision time in T3-T4 is shorter, which affects the proportion of best-response behaviours.

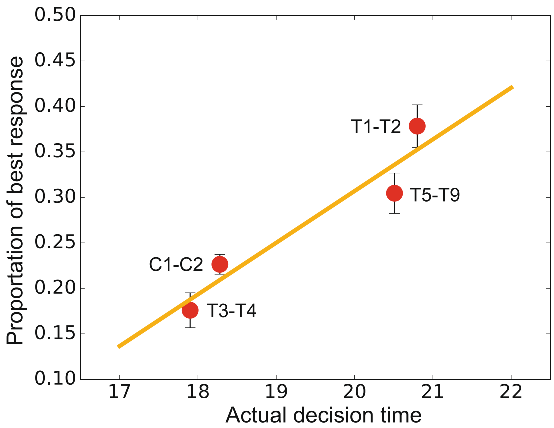


Figure 2. Relationship between best-response behaviours and actual decision time. We analyze the data by classifying the 11 groups into 4 categories depending upon the experimental settings, namely, T1-T2 (large groups with 45 seconds maximum time allowed), T3-T4 (median groups with 30 seconds maximum time allowed), T5-T9 (small groups with 45 seconds maximum time allowed) and C1-C2 (large groups with 45 seconds maximum time allowed). Plotting proportion of best-response behaviours as a function of actual decision time in the 4 categories shows a clear positive correlation. The line is the result of linear regression by using the least squares approach. Error bars denote mean ± s.e.m.

Finally, it is interesting to compare the results of the dual-role UG experiments with our previous single-role UG experiments[31] (two treatment groups with fixed interaction structures and two control groups with well-mixed populations, total $n = 200$, including 100 proposers and 100 responders, 60 rounds, see Supplementary Note 1.5 for the design of the single-role UG experiments). We use the data of the last 35 rounds in the control and treatment groups of the dual-role UG experiments and the last 30 rounds in the control and treatment groups of the single-role UG experiments. We first compare p and q between the dual-role UG experiments treatment groups and the single-role UG experiments treatment groups. As shown in Fig. 3a, both p and q in the dual-role UG experiments treatment groups are much lower than in the single-role UG ones (p: 19.18 vs 43.33, Mann-Whitney U-test, P-value < 0.001; q: 9.92 vs 35.83, Mann-Whitney U-test, P-value < 0.001). The same is true for well-mixed populations, dual-role UG experiments control groups compared to single-role UG experiments control groups (see Fig. 3b, p: 13.03 vs 41.80, Mann-Whitney U-test, P-value < 0.001; q: 4.90 vs 33.65, Mann-Whitney U-test, P-value < 0.001). These observations support our first general conclusion that equal status in the UG is actually detrimental in terms of fairness. This is likely to arise from the fact that single-role UG responders can only increase their payoff by rising their acceptance level, thus forcing proposers to increase their offers; on the contrary, when playing both roles, subjects would mainly focus on the offer in order to increase their payoffs, instead of trying to do it by rising their acceptance levels. This is supported by what we observe in the first round of our dual-role experiments, i.e., proposers choose fair offers but responders agree to accept low offers. In particular, this observation is consistent with those by Oxoby and McLeish[29], where they found that people are more likely to accept low offers in a strategy method UG compared to a sequential decision UG.

In agreement with the explanation above, Fig. 3c shows that the proportion of rational behaviours of responders in the dual-role UG experiments treatment groups is much higher than the proportion in the single-role UG experiments treatment groups (37.53% vs 3.80%). However, the proportion of rational behaviours of proposers (best-response behaviours) in the dual-role UG experiments treatment groups is lower than the proportion in the single-role UG experiments treatment groups (39.78% vs 54.73%, Mann-Whitney U-test, P-value = 0.0027). For well-mixed populations, we find similar but even more dramatic results (see Fig. 3d): The proportion of rational behaviours of responders in the dual-role UG experiments control groups is much higher than the proportion in the single-role UG experiments control groups (55.47% vs 1.53%), but the proportion of rational behaviours of proposers (best-response behaviours) in the dual-role UG experiments control groups is lower than in the single-role UG experiments control groups (30.03% vs 45.93%, Mann-Whitney U-test, P-value < 0.001). This indicates that the acceptance level choices in the populations may be driving the offers, and that single-role UG responders move away from rationality in an attempt to increase their payoffs.

Discussion

In sum, our experiments show that equal status promotes rational splits, and that a fixed interaction structure has positive effects on fairness. Responders keep their acceptance level roughly constant and low all along the experiment, and proposers grow more inclined towards a best-response approach as the experiment proceeds. The best-response behaviour leads to fairer offers in the structured population than in the well-mixed population, although the values are far from what most theoretical models predict[19,21–25] (see a comparison in Supplementary Table 2).

We note that the best-response offer for a proposer must be equal to one of his/her neighbors' acceptance levels, and in many cases, it coincides exactly with the maximum acceptance level. We then calculate the proportion of behaviours that offer the maximum acceptance level, finding that this proportion is slightly lower than that of the best-response behaviours (see Supplementary Notes 1.6). This reveals that some subjects indeed choose their offers based on the best-response consideration rather than simply adopt their neighbors' maximum acceptance

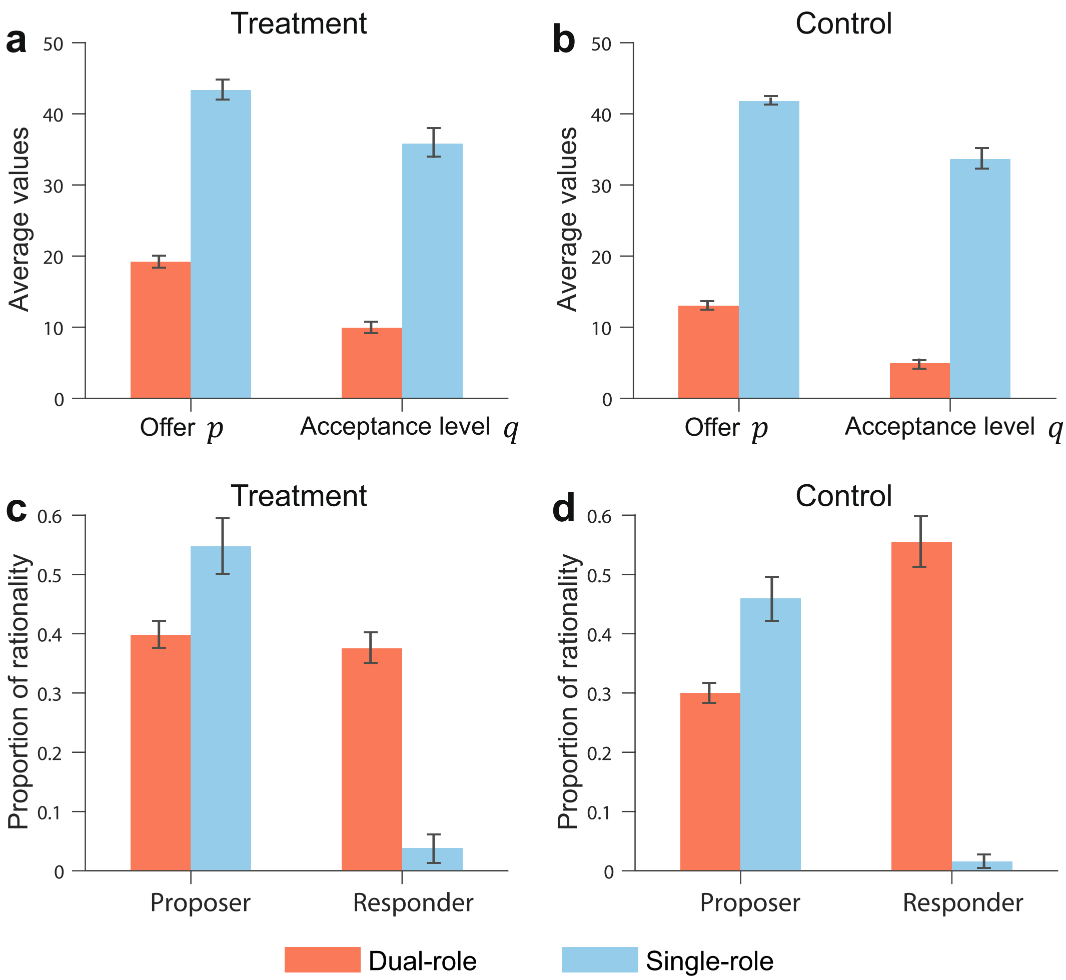


Figure 3. Comparison between dual-role UG experiments and single-role UG experiments. We analyze the data of the last 35 rounds in the treatment and control groups of the dual-role UG experiments and the last 30 rounds in the treatment and control groups of the single-role UG experiments. For each role, there are 318 samples in the dual-role UG experiments and 100 samples in the single-role UG experiments. (**a**) Mean values of p and q in the dual-role UG treatment groups and single-role UG treatment groups which include 50 proposers and 50 responders. (**b**) Mean values of p and q in the dual-role UG control groups and single-role UG control groups which include 50 proposers and 50 responders. (**c**) Proportions of rational behaviours of proposers and responders in the dual-role UG treatment groups and single-role UG treatment groups. (**d**) Proportions of rational behaviours of proposers and responders in the dual-role UG control groups and single-role UG control groups. Error bars indicate ±1 s.e.m.

level. In addition to a large amount of best-response behaviours, there are also many subjects that choose a p value that is 1–5 higher than the best-response offer (see Supplementary Figure 9). This type of behaviour can be explained by a reinforcement learning model, where proposers update their offers based on trial-and-error and responders randomly pick their acceptance level from a database which is built upon the experimental data (see Supplementary Notes 1.7). The simulation results match the experimental results quite well. In both the structured and well-mixed populations, p declines fast in the first 20 rounds, but remains constant or decreases slowly in the last 20 rounds (see Supplementary Figure 16). This is consistent with previous studies showing that reinforcement learning can explain human behaviours in the spatial prisoner's dilemma experiments[38,39].

Models based on evolutionary game theory fail to predict the experimental results because most subjects in the experiments used best-response like strategies or trial-and-error rather than imitation. Rejecting unfair offers can reduce the payoff difference between a subject and his/her neighbors but cannot improve his/her own payoff. Therefore, higher q can spread through (local) imitation, but subjects based on payoff considerations prefer to choose lower q. In particular, the proportions of rational behaviours of responders decrease over time in the treatment groups but remain constant in the control groups (see Supplementary Figure 11). This may be because some rational responders in the treatment groups raise their q to fit the minimum offer of their neighbors (see Supplementary Figure 6). Note that such change is riskless when their neighbors' offers are stable. In contrast, rational responders keep low acceptance level all the time in the control groups because they don't know the history offers of their opponents. This explains why a fixed interaction structure has some effects on promoting fairness.

Interestingly, the result of the dual-role UG is very different from the single-role UG experiment[31] where, when subjects play only as responders, they depart considerably from rationality, in contrast with what we observe here. This in turn raises an interesting question. In the single-role UG, the statuses of subjects are unequal, but they still seem to acknowledge that the responders should obtain about 40% of the pot, which is close to the fair split. By contrast, although subjects have equal status in the dual-role UG, subjects forgo getting large payoffs as responders. It thus seems that when there is no fundamental inequity in initial allocation, offers are more salient to subjects. This is probably because they are perceived as an 'active' choice, while the acceptance level may be regarded as more 'passive' and then less useful to improve the subjects' payoffs. In contrast, in the single-role UG, responders' payoffs depend entirely on proposers making substantial offers. Thus, responders can only increase their payoff by rising their acceptance level. Another possible explanation would be that the unequal status in the single-role UG activates inequity aversion or 'negative' emotions of responders[40,41], while equal status in the dual-role UG promotes rational thinking. Whatever the ultimate reasons behind these anomalous behaviours, the fact that subjects adjust their offers to their environments while keeping their acceptance levels low can provide a sound basis on which new, more accurate theoretical approaches to understand the evolution of fairness can be designed.

Methods

We conducted a series of repeated dual-role UG experiments, including 9 treatment groups T1-T9 with structured populations and 2 control groups C1-C2 with well-mixed populations, in computer labs at Beijing Normal University. All 321 subjects were freshmen and sophomores recruited from Beijing Normal University that had not taken courses on game theory or economy. The interactions were anonymous, and via computers. In the 9 treatment groups, every subject plays with four fixed partners. To be precise, each participant occupies a location of a static 4-degree ring structure and plays the dual-role UG with his/her four immediate neighbours (see Supplementary Figures 3). In addition, a sketch map of the ring structure is showed to subjects. Thus, subjects knew that any two of them don't have exactly the same neighbours, but they are not completely independent. In the 2 control groups, a subject also plays with four subjects, but he/she randomly encounters his/her neighbours in each round. At each round, a subject submits his/her choices for p and q simultaneously ($0 \leq p, q \leq 100$), and plays the standard UG with each of his/her neighbours with the two different roles. To be specific, when a subject plays proposer, all his/her neighbours play responders, and when he/she plays responder, all his/her neighbours play proposers.

In the dual-role UG, a subject's total points interacting with one of his/her neighbours are the sum of his/her points obtained as a proposer and as a responder. For instance, denote the strategy of subject i by (p_i, q_i). Then the points of subject i interacting with neighbor j can be denoted as follows

$$U_{ij} = \begin{cases} p_j + 100 - p_i & p_i \geq q_j \text{ and } p_j \geq q_i \\ 100 - p_i & p_i \geq q_j \text{ and } p_j < q_i \\ p_j & p_i < q_j \text{ and } p_j \geq q_i \\ 0 & p_i < q_j \text{ and } p_j < q_i, \end{cases} \tag{1}$$

where $p_i, p_j, q_i, q_j \in [0, 100]$[15–17]. In the experiments, the payoff of a subject (in a round) is taken as the average points of his/her four interactions. That is, subject i's payoff can be calculated as $U_i = \sum_{j \in \Gamma_i} U_{ij}/4$, where Γ_i is the set of his/her neighbours. At the end of each round, subjects are informed of their neighbours' choices and payoffs (see Supplementary Figures 2 for more details). Note that this differs from setups in which there is competition among responders or proposers, as all deals that verify that the offer is larger than the acceptance level are actually realized.

Ethics. All participants provided written informed consent. All experimental methods were carried out in accordance with the approved guidelines. All experimental protocols were approved by the Ethics Review Committee of Beijing Normal University.

References

1. Güth, W., Schmittberger, R. & Schwarze, B. An experimental analysis of Ultimatum bargaining. *J. Econ. Behav. Org.* **3**, 367–388 (1982).
2. van Damme, E. *et al.* How Werner Güth's ultimatum game shaped our understanding of social behavior. *J. Econ. Behav. Org.* **108**, 292–318 (2014).
3. Fehr, E. & Fischbacher, U. Why social preferences matter–The impact of non-selfish motives on competition, cooperation and incentives. *Econ. J.* **112**, C1–C33 (2002).
4. Güth, W. & Tietz, R. Ultimatum bargaining behavior: A survey and comparison of experimental results. *J. Econ. Psychol.* **11**, 417–449 (1990).
5. Kagel, J. H. & Roth, A. E. *The Handbook of Experimental Economics*. (Princeton University, Princeton, NJ, 1995).
6. Camerer, C. *Behavioral Game Theory*. (Princeton University, Princeton, NJ, 2003).
7. Henrich, J. *et al. Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. (Oxford University Press, Oxford, 2004).
8. Oosterbeek, H., Sloof, R. & Kuilen, G. V. D. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Exp. Econ.* **7**, 171–188 (2004).
9. Henrich, J. *et al.* Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).
10. Rand, D. G., Tarnita, C. E., Ohtsuki, H. & Nowak, M. A. Evolution of fairness in the oneshot anonymous ultimatum game. *Proc. Natl. Acad. Sci. USA* **108**, 19193–19198 (2013).
11. Rodríguez-Lara, I. Equity and bargaining power in ultimatum games. *J. Econ. Behav. Org.* **130**, 144–165 (2016).

12. Lee, K. & Shahriar, Q. Fairness, one's source of income, and others'decisions: An ultimatum game experiment. *Managerial. Decis. Econ.* **38**, 423–431 (2017).
13. Weber, R. A., Camerer, C. F. & Knez, M. Timing and virtual observability in ultimatum bargaining and "weak link" coordination games. *Exp. Econ.* **7**, 25–48 (2004).
14. Brosnan, S. F. & de Waal, F. B. M. Monkeys reject unequal pay. *Nature* **425**, 297–299 (2003).
15. Nowak, M. A., Page, K. M. & Sigmund, K. Fairness versus reason in the ultimatum game. *Science* **289**, 1773–1775 (2000).
16. Nowak, M. A., Page, K. M. & Sigmund, K. The spatial ultimatum game. *Proc. R. Soc. Lond. B* **267**, 2177–2182 (2000).
17. Page, K. & Nowak, M. A. A generalized adaptive dynamics framework can describe the evolutionary ultimatum game. *J. Theor. Biol.* **209**, 173–179 (2000).
18. Sigmund, K., Hauert, C. & Nowak, M. A. Reward and punishment. *Proc. Natl. Acad. Sci. USA* **98**, 10757–10762 (2001).
19. Killingback, T. & Studer, E. Spatial ultimatum games, collaborations and the evolution of fairness. *Proc. R. Soc. Lond. B* **268**, 1797–1801 (2001).
20. Page, K. & Nowak, M. A. Empathy leads to fairness. *Bull. Math. Biol.* **64**, 1101–1116 (2002).
21. Kuperman, M. & Risau-Gusman, S. The effect of topology on the spatial ultimatum game. *Eur. Phys. J. B* **62**, 233–238 (2008).
22. Sinatra, R. *et al.* The ultimatum game in complex networks. *J. Stat. Mech. Theory Exp.* **2009**, P09012 (2009).
23. Szolnoki, A., Perc, M. & Szabó, G. Defense mechanisms of empathetic players in the spatial ultimatum game. *Phys. Rev. Lett.* **109**, 078701 (2012).
24. Iranzo, J., Román, J. & Sánchez, A. The spatial Ultimatum game revisited. *J. Theor. Biol.* **278**, 1–10 (2011).
25. Iranzo, J., Floría, L. M., Moreno, Y. & Sánchez, A. Empathy emerges spontaneously in the ultimatum game: small groups and networks. *PLoS ONE* **7**, e43781 (2012).
26. Carter, J. R. & Irons, M. D. Are economists different, and if so, why? *J. Econ. Perspec.* **5**, 171–177 (1991).
27. Weg, E. & Smith, V. On the failure to induce meager offers in ultimatum game. *J. Econ. Psychol.* **14**, 17–32 (1993).
28. Oosterbeek, H., Sloof, R. & Van De Kuilen, G. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Exp. Econ.* **7**, 171–188 (2004).
29. Oxoby, R. J. & McLeish, K. N. Sequential decision and strategy vector methods in ultimatum bargaining: evidence on the strength of other-regarding behavior. *Econ. Lett.* **84**, 399–405 (2004).
30. Brandts, J. & Charness, G. The strategy versus the direct-response method: a first survey of experimental comparisons. *Exp. Econ.* **14**, 375–398 (2011).
31. Han, X. *et al.* Emergence of Communities and Diversity in Social Networks. *Proc. Natl. Acad. Sci. USA* **114**, 2887–2891 (2017).
32. Fehr, E. & Fischbacher, U. Social norms and human cooperation. *Trends Cogn. Sci.* **8**, 185–190 (2004).
33. Kearns, M. Experiments in social computation. *Commun. ACM* **55**, 56–67 (2012).
34. Sutter, M., Kocher, M. & Strauß, S. Bargaining under time pressure in an experimental ultimatum game. *Econ. Lett.* **81**, 341–347 (2003).
35. Cappelletti, D., Güth, W. & Ploner, M. Being of two minds: Ultimatum offers under cognitive constraints. *J. Econ. Psychol.* **32**, 940–950 (2011).
36. Rand, D. G., Greene, J. D. & Nowak, M. A. Spontaneous giving and calculated greed. *Nature* **489**, 427–430 (2012).
37. Bear, A. & Rand, D. G. Intuition, deliberation, and the evolution of cooperation. *Proc. Natl. Acad. Sci. USA* **113**, 936–941 (2016).
38. Cimini, G. & Sánchez, A. Learning dynamics explains human behaviour in Prisoner's Dilemma on networks. *J. R. Soc. Interface* **11**, 20131186 (2014).
39. Ezaki, T., Horita, Y., Takezawa, M. & Masuda, N. Reinforcement learning explains conditional cooperation and its moody cousin. *PLoS Comput. Biol.* **12**, e1005034 (2016).
40. Fehr, E. & Schmidt, K. M. A Theory of Fairness, Competition, and Cooperation. *Q. J. Econ.* **114**, 817–868 (1999).
41. Sanfey, A. G. *et al.* The neural basis of economic decision-making in the ultimatum game. *Science* **300**, 1755–1758 (2003).

Acknowledgements

We thank Long Ma, Xiao-Yan Sun, Zhao-Long Hu for assistance with carrying out the experiments and Xin-Di Wang, Ke-Qiang Li for helpful discussions. This work was partially supported by the National Science Foundation of China (Grant No. 71771026, BZ; Grant No. 71401037, CS; Grant No. 71631002, WW, BZ), by EU through FET-Proactive Project DOLFINS (contract no. 640772, AS) and FET-Open Project IBSEN (contract no. 662725, AS), and by grant FIS2015-64349-P (MINECO/FEDER, UE, AS).

Author Contributions

X.H., W.W., B.Z., Z.G. and A.S. conceived the original idea for the experiment, X.H., S.C., J.B., B.Z. and W.W. contributed to the final experimental setup, X.H., S.C., J.B., B.Z. and W.W. carried out the experiments, X.H., B.Z. and W.W. analyzed the data, X.H., W.W., B.Z., Z.G. and A.S. discussed the analysis results and X.H., W.W., B.Z., Z.G. and A.S. wrote the paper. All authors gave final approval for publication.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-19503-x.

Competing Interests: The authors declare that they have no competing interests.

Supplementary Information for

Equal status in Ultimatum Games promotes rationality

Xiao Han[1,2], Shinan Cao[3], Jian-Zhang Bao[2], Wen-Xu Wang[2], Boyu Zhang[4], Zi-You Gao[1], and Angel Sánchez[5,6,7]

[1]*MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, 100044, P. R. China*

[2]*School of Systems Science, Beijing Normal University, Beijing, 100875, P. R. China*

[3]*School of Finance, University of International Business and Economics, Beijing, 100029, P. R. China*

[4]*School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, P. R. China*

[5]*Grupo Interdisciplinar de Sistemas Complejos (GISC), Departamento de Matemáticas, Universidad Carlos III de Madrid, Spain*

[6]*UC3M-BS Institute of Financial Big Data (IFIBID), Universidad Carlos III de Madrid, Spain*

[7]*Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, Spain*

Contents

1 Supplementary Notes

1.1 Experimental settings for dual-role ultimatum game experiments

We conducted a series of repeated dual-role UG experiments, including 9 treatment groups with structured populations and 2 control groups with well-mixed populations in computer labs of Beijing Normal University. Detail settings of the groups are shown in Supplementary Table 2. All 321 subjects were freshmen and sophomores recruited from Beijing Normal University without taking classes of game theory and economy. The interactions were anonymous, and via computers. Frosted glass dividers ensured that the students could not see each other (see Supplementary Figure 1). We built the experimental platform by using PHP, mySQL and javascript, and ran the platform programs on the server. Schematic diagrams of our experimental platform are shown in Supplementary Figure 2.

Before starting the experiment, we explained the game to all subjects, including the rules of the game, the purpose of the game, and the feedback information in the computer in 20 minutes. All subjects in each session were given the same instructions (in Chinese). To ensure that all subjects fully understand the game, we implemented 2 exercises and 5 practice rounds before the formal experiment (last about 10 minutes). During the period, all subjects can raise their hands, and our experimenters would answer their questions. The formal experiment lasted about 60 minutes, and subjects were not told the number of rounds, so as to avoid end round effects. Each round is time limited (except for the first round). In T3-T4, subjects have 30 seconds to submit their decisions at each round, and in other groups, subjects have 45 seconds to submit their decisions. Subjects knew that if they did not decide within the given time, they would be allocated their own decisions in the previous round. Since the subjects had familiarized themselves with the game during the practice rounds, this happened only 412 times in 24672 decisions (1.67 %). After the experiment, the total payoffs of each subject obtained in the formal experiment was converted to Chinese Yuan at a ratio of 100 : 1. This pay plus 30 Chinese Yuan is his/her final income (see Supplementary Table 2 for details).

In all experiments, the data of three subjects (two in C1 and one in T3) are excluded because we notice that they do not really understand the game. To keep the comparison unbiased, all results were calculated using data in 1-70 rounds.

1.2 Experimental instructions for T1-T9

Welcome and thanks for participating in this game. Please read the game instruction carefully. If you have any questions please raise your hand. One experimenters will then come to you and answer your questions. From now on, communication with other participants is not allowed. Please switch off your mobile phone and keep quiet in the whole game.

You will play a decision making game. In the game, you would not know other persons' true identity. Your scores depend on your and your partners' decisions. Your final income = fixed income 30 Chinese Yuan + 0.01 × total scores.

Game instruction

1. In this game, you play two roles, proposer and responder, and submit your offer and demand simultaneously. You play the game with four fixed partners, who also play the two roles.
2. At each round, you play the game twice with different roles at the same time. When you play proposer, all your partners play responders; when you play responder, all your partners play proposers.
3. A proposer and a responder share 100 points. If proposer' offer greater than or equal to responder' demand, the proposer receives (100$-$**proposer's offer**), and the responder receives **(proposer's offer)**, otherwise, both receive 0.
4. Your total points are the sum of your four interactions. Your score = (your total points)/(number of partners). After all the participants submit their offer and demand, the system will calculate your points obtained as a proposer, your points obtained as a responder, your total points, and your scores.

Example (A sketch map of the ring structure is showed to subjects.)

1. Suppose you have four partners, A, B, C and D.
2. At each round, you play the game with all your four partners.
3. You submit your offer p and demand q.
4. Suppose the four partners' offers are p_A, p_B, p_C and p_D, and demands are q_A, q_B, q_C, q_D.
5. Suppose your offer p satisfies $q_C, q_D > p \geq q_A, q_B$. Then, as a proposer, your offer is accepted by A and B, and your points obtained as a proposer are $(100 - p) + (100 - p)$.
6. Suppose your demand q satisfies $p_A < q \leq p_B, p_C, p_D$. Then, as a responder, you accept offers from B, C and D, and your points obtained as a responder are $p_B + p_C + p_D$.
7. In this round, your total points are $(200 - 2p) + p_B + p_C + p_D$.
8. Your scores are $[(200 - 2p) + p_B + p_C + p_D]/4$, where 4 is the number of your partners.

Exercise 1

Now we generate your and your partners' offers and demands randomly. For simplicity, we only generate multiples of 10 for offers and demands. You need calculate your points obtained as a proposer, your points obtained as a responder, your total points and your scores. Different subjects may have different partners, so you cannot calculate your partners' scores.

Exercise 2

Same as Exercise 1.

1.3 Experimental instructions for C1-C2

Game instruction

1. In this game, you play two roles, proposer and responder, and submit your offer and demand simultaneously. You play the game with four partners, who also play the two roles.
2. At each round, you play the game twice with different roles at the same time. When you play proposer,

all your partners play responders; when you play responder, all your partners play proposers.
3. A proposer and a responder share 100 points. If proposer' offer greater than or equal to responder' demand, the proposer receives (100−**proposer's offer)**, and the responder receives **(proposer's offer)**, otherwise, both receive 0.
4. Your total points are the sum of your four interactions. Your score = (your total points)/(number of partners). After all the participants submit their offer and demand, the system will calculate your points obtained as a proposer, your points obtained as a responder, your total points, and your scores.
5. At the beginning of each round, you will randomly encounter four new partners.

[The rest parts are same as *Instructions for T1-T2*.]

1.4 Calculating best-response behaviour

We used a rigorous definition of best-response behaviour to identify whether proposers were rational. The best strategy for rational proposers in each round was to offer the amount that maximizes payoff, keeping in mind the acceptance levels of indicated by neighbouring responders in the previous round [1, 2]. For a proposer with k neighbouring responders whose acceptance levels in the previous round were respectively $q_1, ..., q_k$ (with $q_1 < \cdots < q_k$), the best strategy was $p = \text{argmax}_{q_i}\{i \times (100 - q_i)/k\}$, where $i \times (100 - q_i)/k$ was the payoff if the proposers offered q_i. We found that the proportion of rational proposers gradually increased and eventually nearly about half of all proposers take best-response behaviours in all groups. Our definition of best-response behaviours of proposers was extremely rigorous which indicate that the proportion of rational proposers was quite high in the last several rounds.

1.5 Experimental settings for single-role ultimatum game experiments

In our previous single-role UG experiments, we totally conducted 2 treatment groups and 2 control groups [3]. In each group, we recruited 50 subjects, half of whom were randomly assigned proposers and the rest randomly assigned responders. The interactions were executed via computer and were anonymous. We built the experimental platform by using z-Tree [4]. In the single-role UG experiments, each subject only enacted one role, proposer or responder, and his/her role didn't change during the experiment. In the treatment groups, each subject was assigned in a location within a static bipartite network, and played UG with their neighbours who enacted the other role. All of the proposers' neighbours are responders and vice versa. We use two static bipartite networks including a regular bipartite network in which each node has four neighbors and a random bipartite network in which the number of neighbors ranges from 2 to 6 (with an average degree of 4). In the control groups, the population structure changes and players randomly encounter their neighbours in each round.

In the single-role UG, all subjects must use one decision behavior as they interact with their neighbors; that is, a proposer must make the same offer $p(0 \leq p \leq 100)$ to all of his or her neighboring responders, and a responder must indicate the same minimum acceptance level $q(0 \leq q \leq 100)$ to all of his or her neighboring proposers. The payoff of a subject (proposer or responder) is taken as the average points of his/her all interactions. That is, a subject i's payoff can be calculated as $U_i = \sum_{j \in \Gamma_i} U_{ij}/k_i$, where Γ_i is the set of his/her neighbours, k_i is the number of his/her neighbours, and U_{ij} is the points of subject i interacting with neighbor j.

1.6 Calculating 'offering the maximum acceptance level' behaviour

We note that the best-response offer for a proposer must be equal to one of his/her neighbors' acceptance levels. In particular, the best-response offer is exactly the maximum acceptance level if $\max\{q_i(t)\} \leq 25$. We then calculate the proportion of behaviours that offer the maximum acceptance level, i.e., $p(t+1) = \max\{q_i(t)\}$. The proportions of 'offering the maximum acceptance level' behaviours are 0.3043 and 0.2250 for the treatment groups and the control groups, respectively. Note that the proportions of best-response behaviours are 0.3047 and 0.2264 for the treatment groups and the control groups, respectively, i.e., the proportion of behaviours that offer the maximum acceptance level is slightly lower than that of the best-response behaviours. This reveals that some subjects indeed choose their offers based on the best-response consideration rather than simply adopt their neighbors' maximum acceptance level.

1.7 Reinforcement learning model

In order to get a deeper insight into this theoretical significance of our experimental results, we have run simulations based on a type of reinforcement learning model [5]. Firstly, we build two databases of all responder acceptance levels obtained from all treatment groups and all control groups, respectively. Then we randomly pick responder acceptance level sequences from the two databases and use reinforcement learning model to reproduce proposers' offer p for treatment groups and control groups, respectively. We use a static 4-degree ring structure and well-mixed population with 50 subjects in treatment simulations and control simulations, respectively. The simulation process of reinforcement learning model is as follows.

1. *Initial propensities*: We reduce the offer set of proposers into $\{0, 5, 10, ..., 100\}$ in our simulations and assume that all proposers have the same initial propensities for all offers p in the simplified strategy set, which are set equal to fair split 50.

2. *Update propensities* : Suppose a proposer i has chosen offer p_k in round t, the propensity in round $t+1$ is updated by

$$Q_i^k(t+1) = \theta u_i^k(t) + (1-\theta) Q_i^k(t), \quad (1)$$

where $Q_i^k(t+1)$ and $Q_i^k(t)$ denote the propensities of proposer i chosen offer p_k at round $t+1$ and round t, respectively, $u_i^k(t)$ is the payoff of proposer i chosen offer p_k at round t, θ is the learning rate which is set to 0.2 in our simulations.

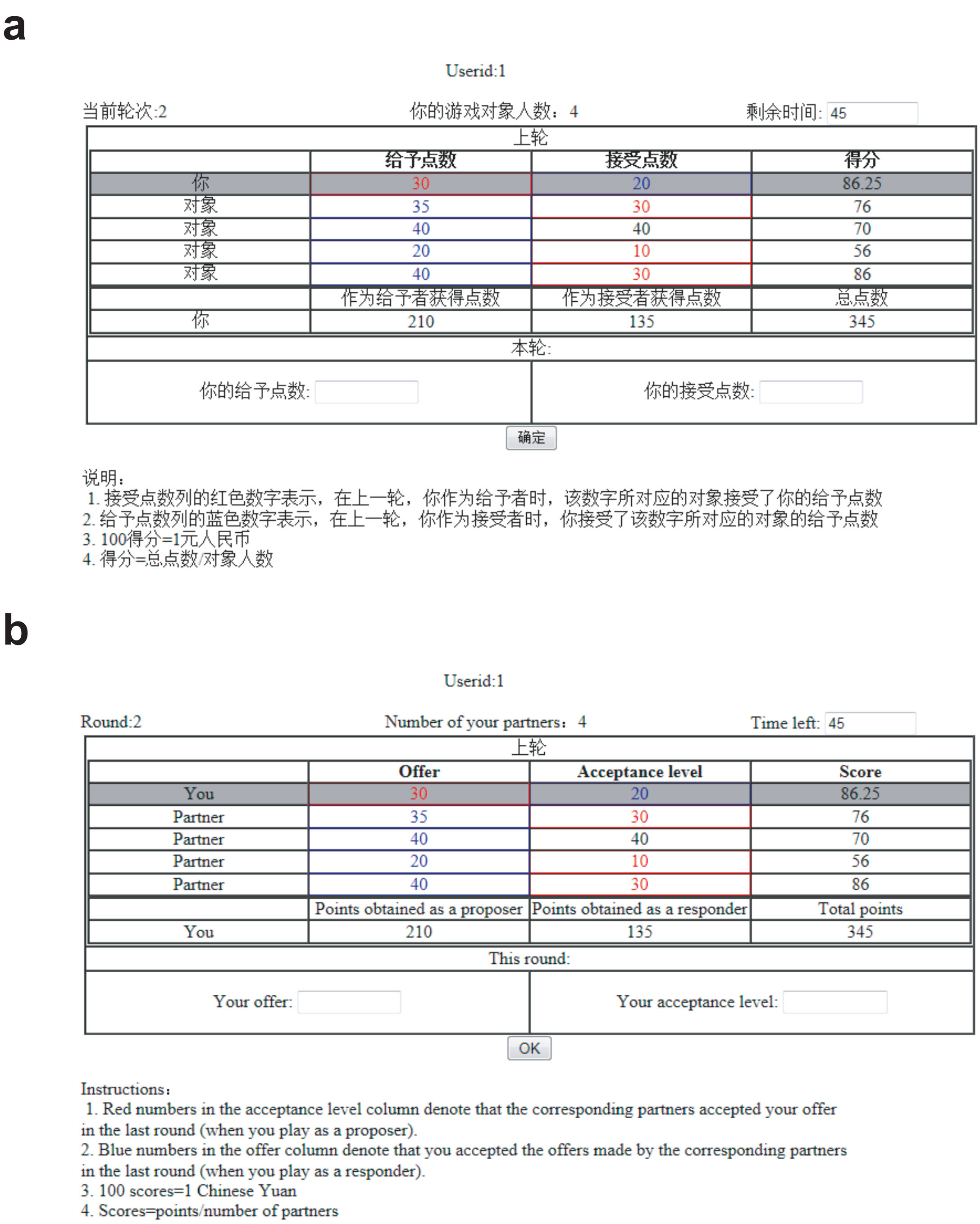
3. *Update probabilities*: The probability of choosing offer p_k in round $t+1$ is determined by

$$P_i^k(t+1) = \frac{e^{\lambda Q_i^k(t+1)}}{\sum_{k=1}^{21} e^{\lambda Q_i^k(t+1)}}, \tag{2}$$

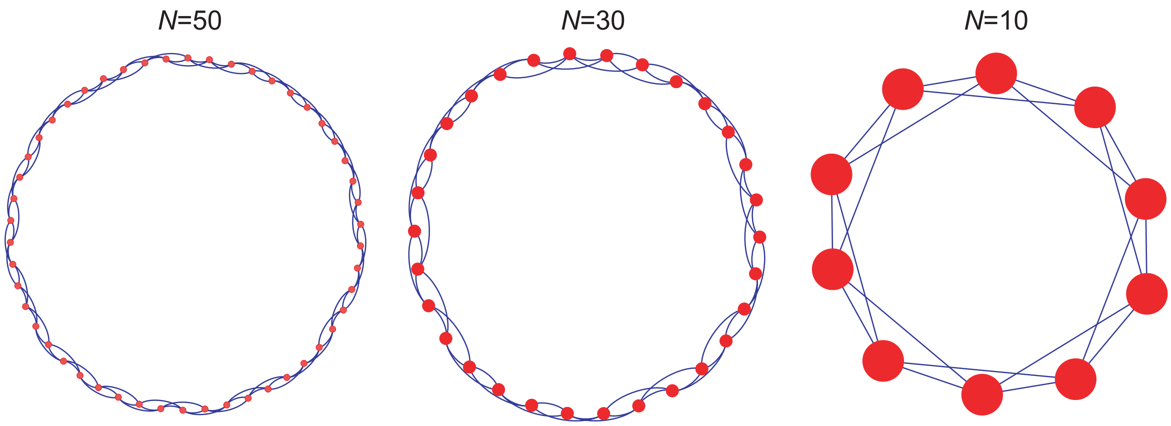
where $k = 1, 2, \cdots, 21$ and λ is a parameter that determines reinforcement sensitivity which is set to 0.2 in our simulations. We repeat step 2 and step 3 until the simulation reaches a predetermined round.

2 Supplementary Figures

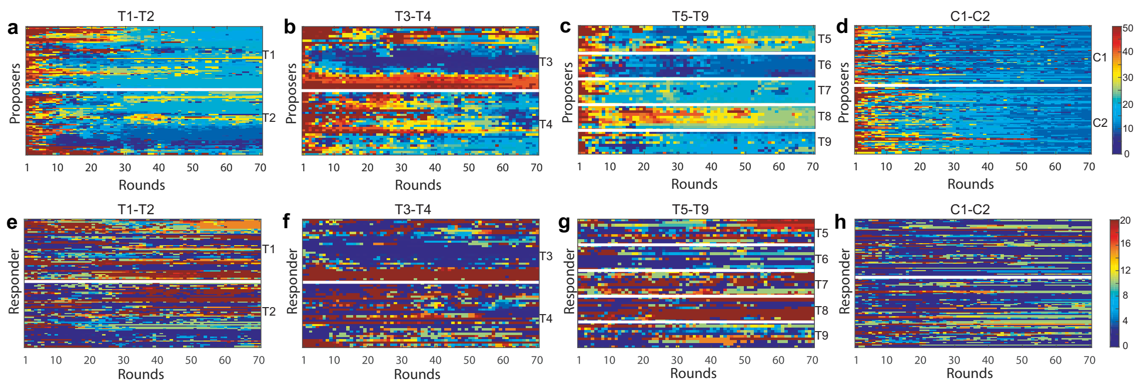
Supplementary Figure 1: Photos of the computer lab. Frosted glass dividers are used to avoid subjects glancing others screens. All subjects can play the game in a quiet environment.



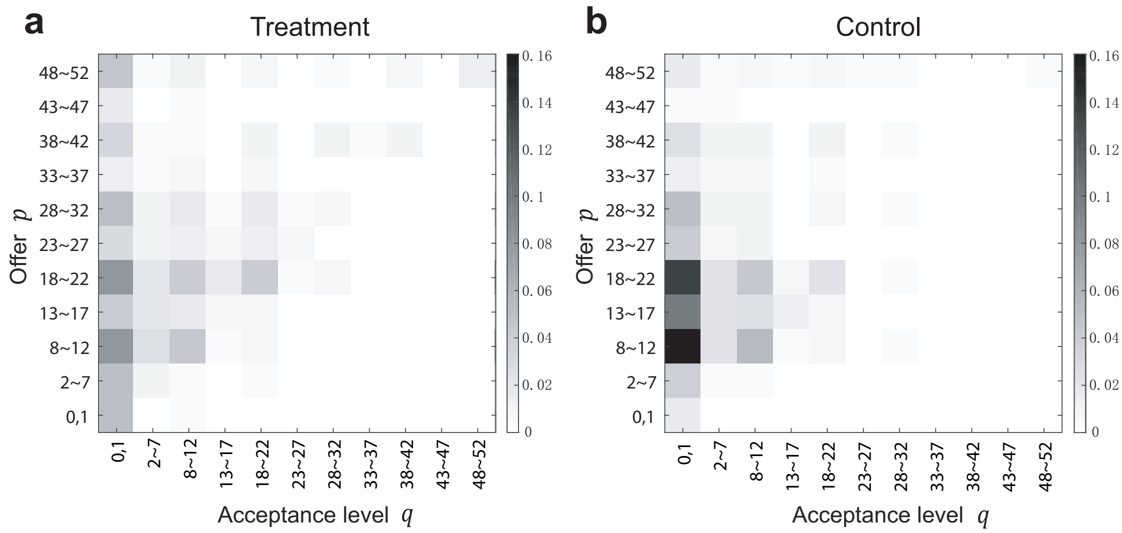
Supplementary Figure 2: Screenshots of the experimental platform. **a**, Chinese version. **b**, English version, translated from the Chinese version. We used the interface in Chinese in the experiments.



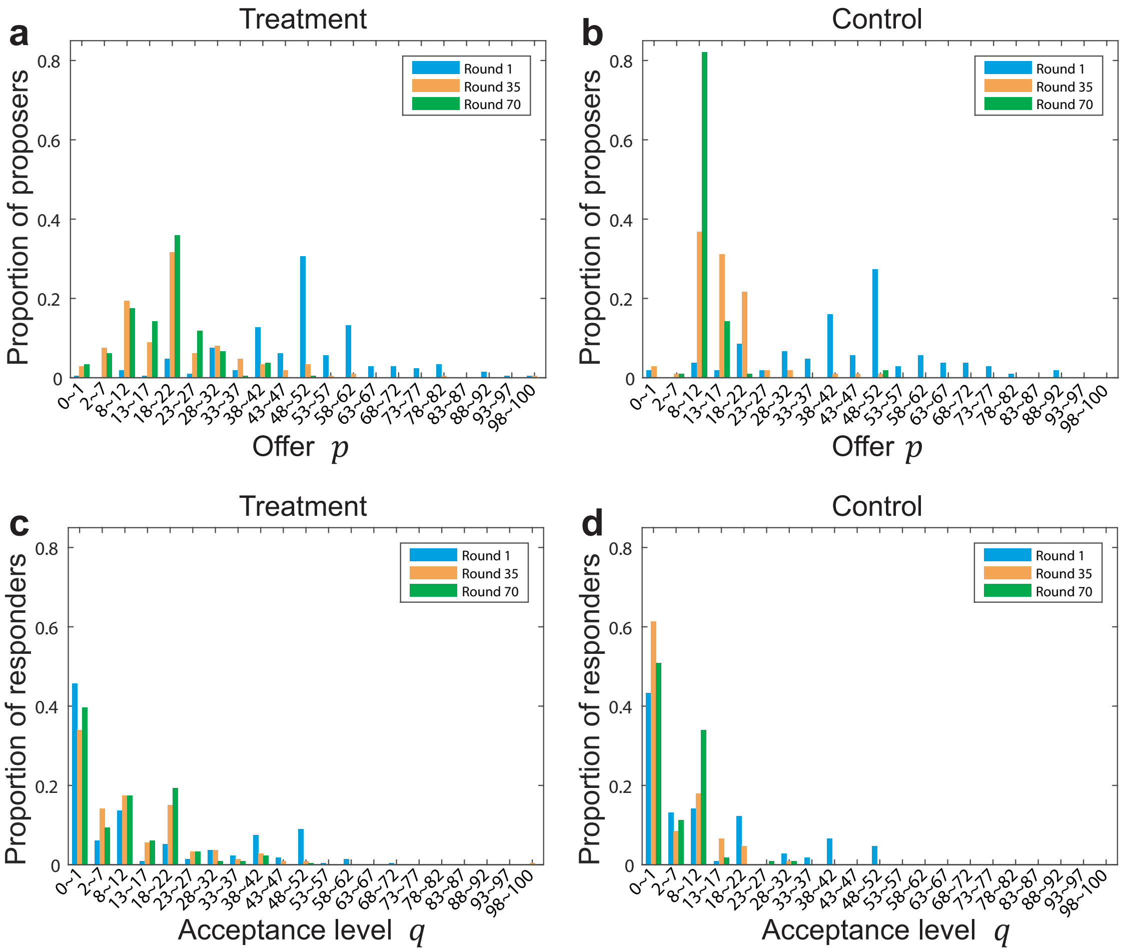
Supplementary Figure 3: Illustrations of the ring structures in the treatments. In the treatment groups, subjects played dual-role UG on the 4-degree ring structures (these figures were showed to subjects). These ring structures have 50 nodes (treatment groups T1-T2, see left), 30 nodes(treatment groups T3-T4, see middle), or 10 nodes (treatment groups T5-T9, see right). We note that two neighbouring subjects have and only have two joint neighbours. Thus, any two subjects don't have exactly the same neighbours, but on the other hand they are not completely independent either.



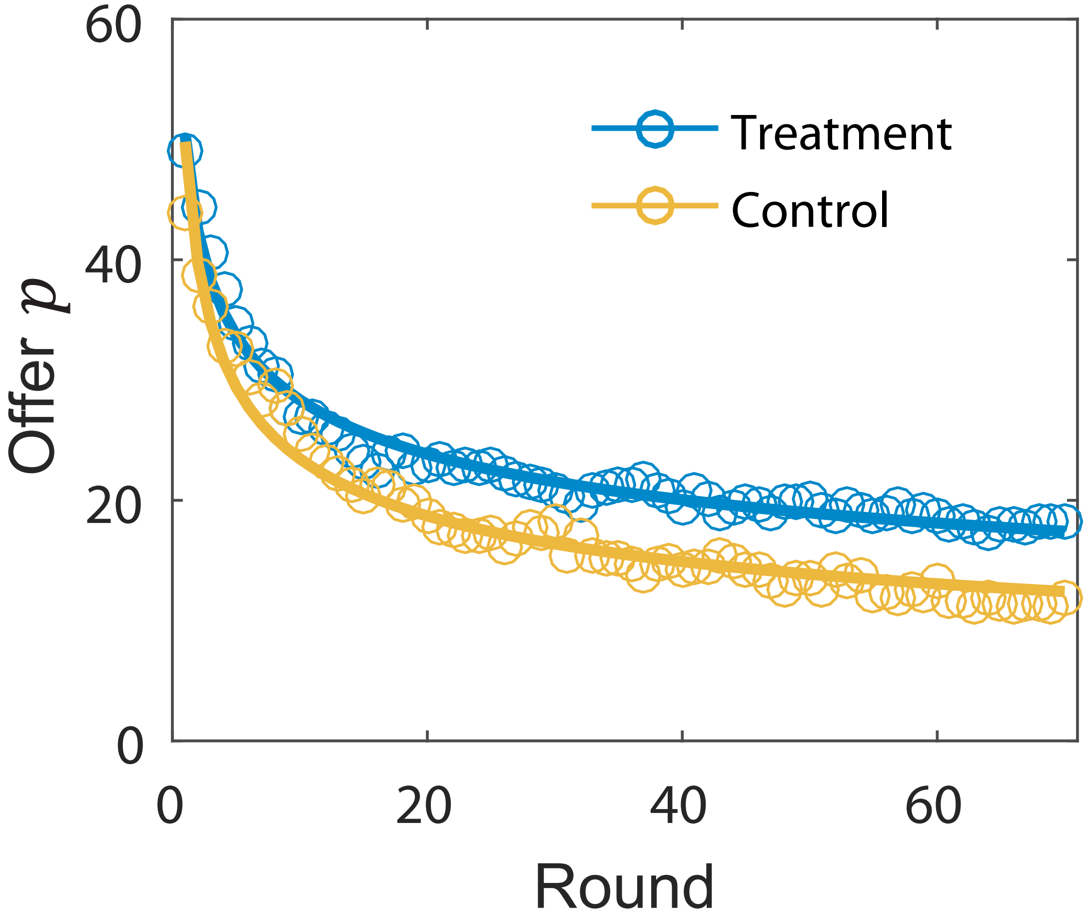
Supplementary Figure 4: Spatio-temporal patterns of proposers' offers and responders' acceptance levels. a-d, Spatio-temporal patterns of the proposers' offer p in T1-T2 (a), T3-T4 (b), T5-T9 (c) and C1-C2 (d). The ordinate represents the spatial orders of proposers. Two proposers with most common neighbours will be adjacent to each other. The color bar represents the value of offer p. **e-f**, Spatio-temporal patterns of the responders' acceptance levels q in T1-T2 (e), T3-T4 (f), T5-T9 (g) and C1-C2 (h). The ordinate represents the spatial orders of responders. Two responders with most common neighbours will be adjacent to each other. The color bar represents the value of acceptance level q.



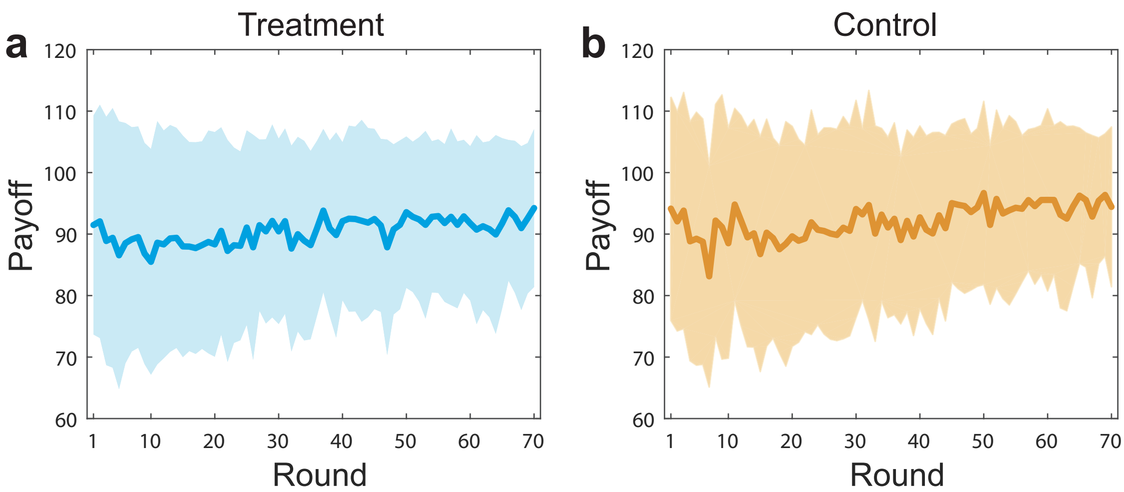
Supplementary Figure 5: Distributions of individual strategies in the dual-role UG. a, The treatment groups, **b**, the control groups.



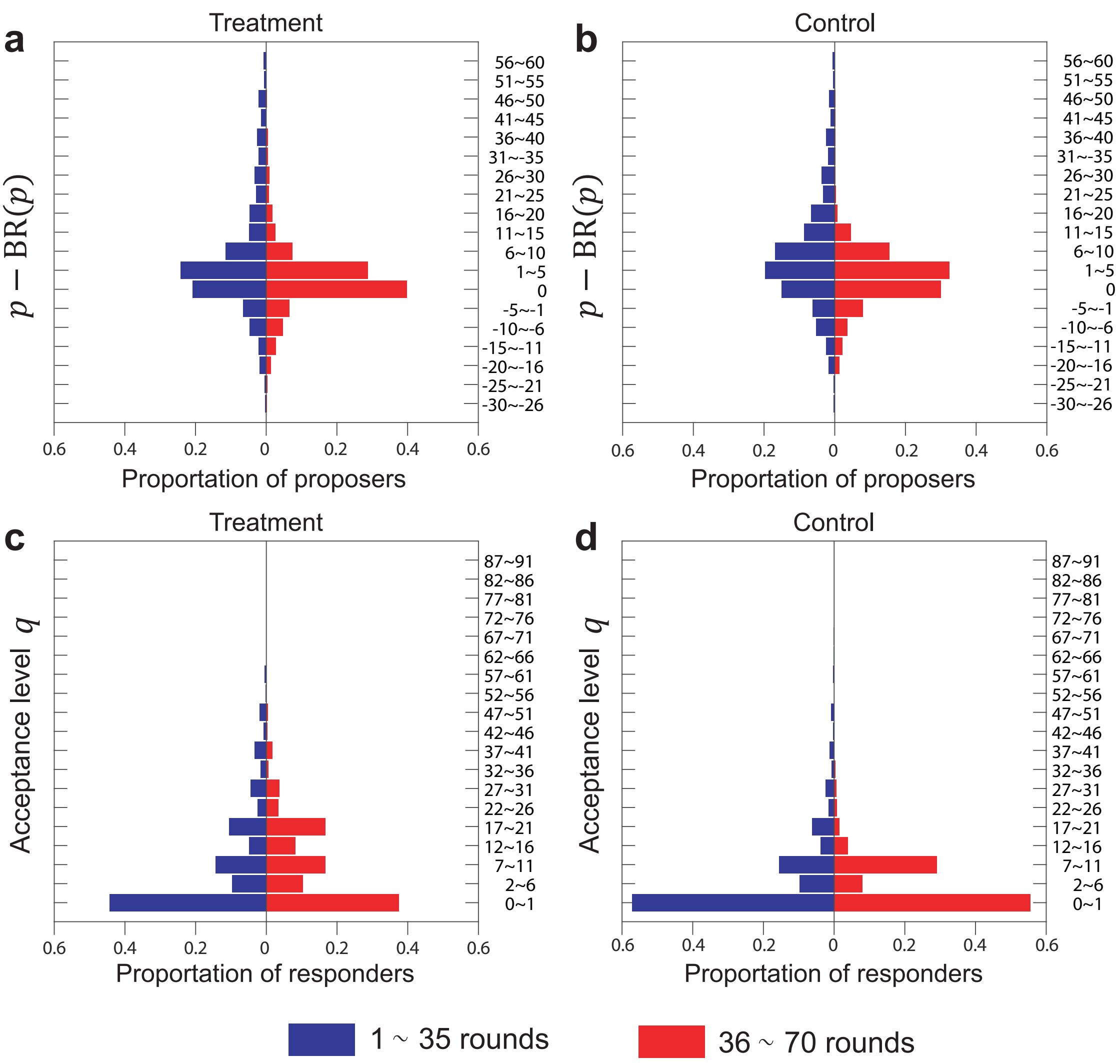
Supplementary Figure 6: Distribution of subjects offers p and acceptance levels q. a-b, Histograms of the proposers' offers p at round 1, round 35 and round 70 in the treatment and control groups, respectively. **c-d**, Histograms of the responders' acceptance levels q at round 1, round 35 and round 70 in the treatment and control groups, respectively.



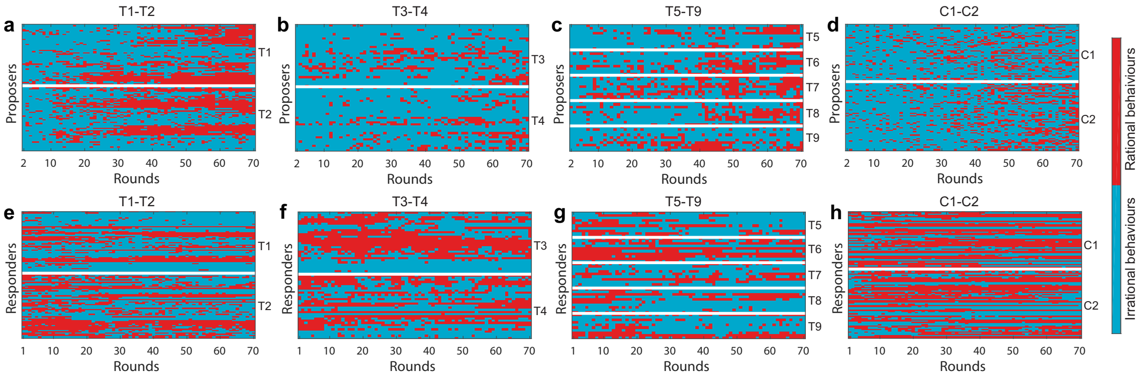
Supplementary Figure 7: Power law regressions for the time evolution of mean values of offers p. In the control groups, $p(t) = 49.81t^{-0.327}$ (t is the round number) with coefficient of determination $R^2 = 0.9576$. In the treatment groups, $p(t) = 50.41t^{-0.250}$ with coefficient of determination $R^2 = 0.9760$. The regression result shows clearly that the mean value of p decreases faster in the control groups than the treatment groups.



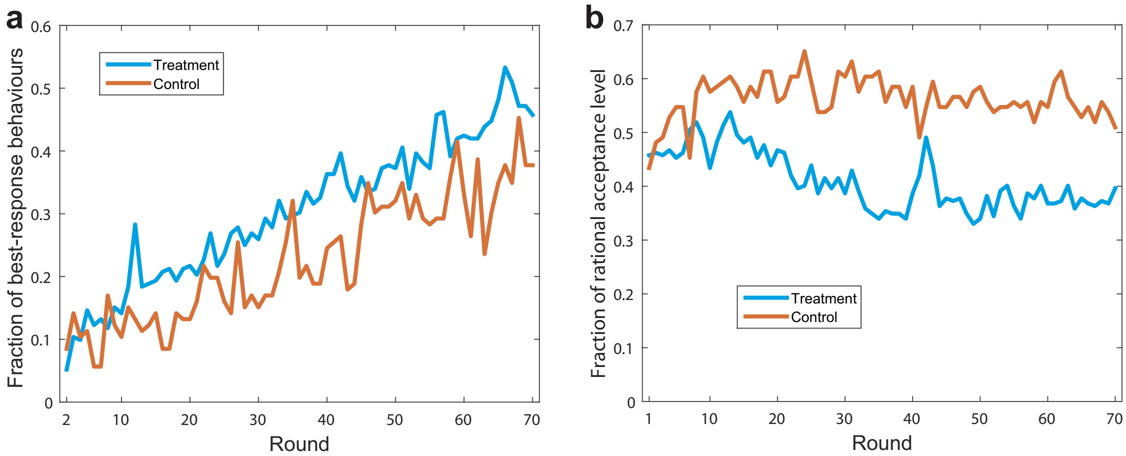
Supplementary Figure 8: Time evolution of mean values and standard deviations of payoffs. (a)Treatment groups, (b) Control groups. The mean payoffs of the treatment groups and control groups are 90.47 and 92.14, respectively. There are positive correlation between the payoffs and rounds in the treatment groups (Pearson correlation coefficient$= 0.6657$, P-value< 0.001) and control groups (Pearson correlation coefficient$= 0.6619$, P-value< 0.001). Moreover, the standard deviations of payoffs and rounds have negative correlations in the treatment groups (Pearson correlation coefficient$= -0.8514$, P-value< 0.001) and control groups (Pearson correlation coefficient$= -0.8633$, P-value< 0.001).



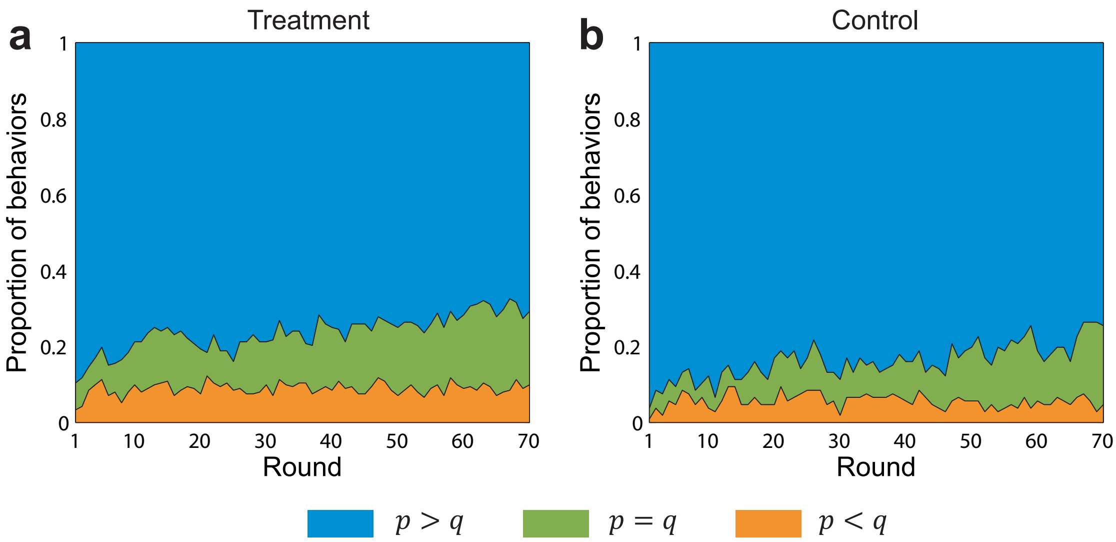
Supplementary Figure 9: Distributions of $p - \mathrm{BR}(p)$ and acceptance level q. We show the distributions of $p - \mathrm{BR}(p)$ and q separately for rounds 1 to 35 and rounds 36 to 70. **a-b**, The distributions of $p - \mathrm{BR}(p)$ in the treatment and control groups, respectively. In all the groups, the proportions of best-response behaviours in the last 35 rounds are higher than that in the previous 35 rounds. **c-d**, The distributions of q in the treatment and control groups, respectively. In all the groups, the distributions in the previous 35 rounds and the last 35 rounds are similar. The result is consistent with Supplementary Figure 7.



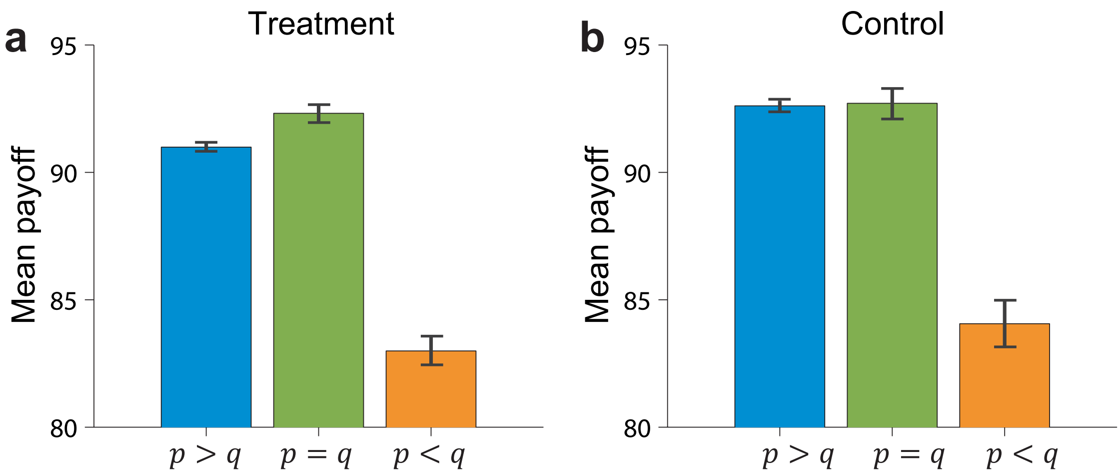
Supplementary Figure 10: Spatio-temporal patterns of rational and irrational behaviours of proposers and responders. a-d, Spatio-temporal patterns of rational behaviours (rigorous best-response behaviours) and irrational behaviours of proposers in T1-T2 (a), T3-T4 (b), T5-T9 (c) and C1-C2 (d). The ordinate represents the spatial orders of proposers. Two proposers with most common neighbours will be adjacent to each other. **e-h**, Spatio-temporal patterns of rational behaviours ($q = 0$ or 1) and irrational behaviours of responders in T1-T2 (e), T3-T4 (f), T5-T9 (g) and C1-C2 (h). The ordinate represents the spatial orders of responders. Two responders with most common neighbours will be adjacent to each other. The red color represents rational behaviours and the blue color represents irrational behaviours.



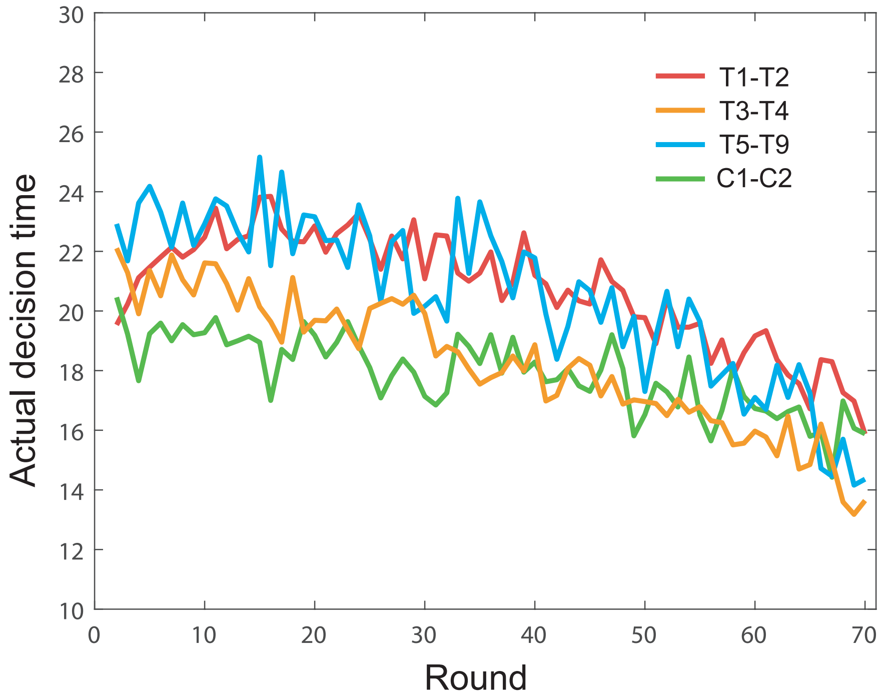
Supplementary Figure 11: Time evolution of fractions of rational behaviours. a, The mean values of fractions of rational behaviours of proposers (i.e., best response) are 0.3047 and 0.2264 in the treatment groups and control groups, respectively. The Pearson correlation coefficient between proportions of rational behaviours of proposers are 0.9699 and 0.9011 in the treatment groups and control groups, respectively. **b**, The mean values of fractions of rational behaviours of responders (i.e., q = 0 or 1) are 0.4095 and 0.5628 in the treatment groups and control groups, respectively. The Pearson correlation coefficient between proportions of rational behaviours of responders are -0.7367 and -0.0346 in the treatment groups and control groups, respectively.



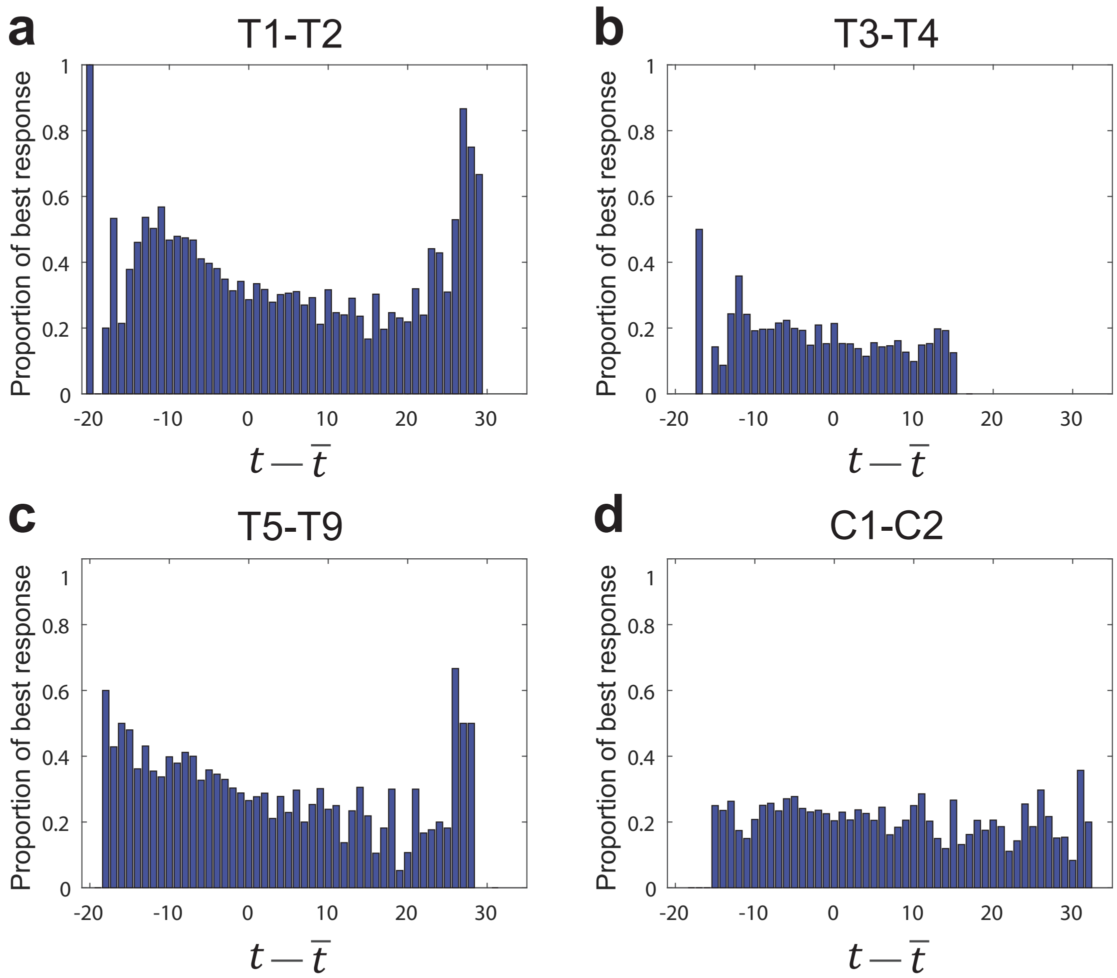
Supplementary Figure 12: Time evolution of the proportions of the three types of behaviours, namely, altruistic behaviours $p > q$, empathic behaviours $p = q$ and selfish behaviours $p < q$. **a**, In the treatment groups, the mean proportion of altruistic behaviours $p > q$, empathic behaviours $p = q$ and selfish behaviours $p < q$ are 76.37%, 14.72% and 8.91%. **b**, In the control groups, the mean proportion of altruistic behaviours $p > q$, empathic behaviours $p = q$ and selfish behaviours $p < q$ are 83.84%,10.55% and 5.61%. Most subjects in our experiments adopt altruistic behaviours.



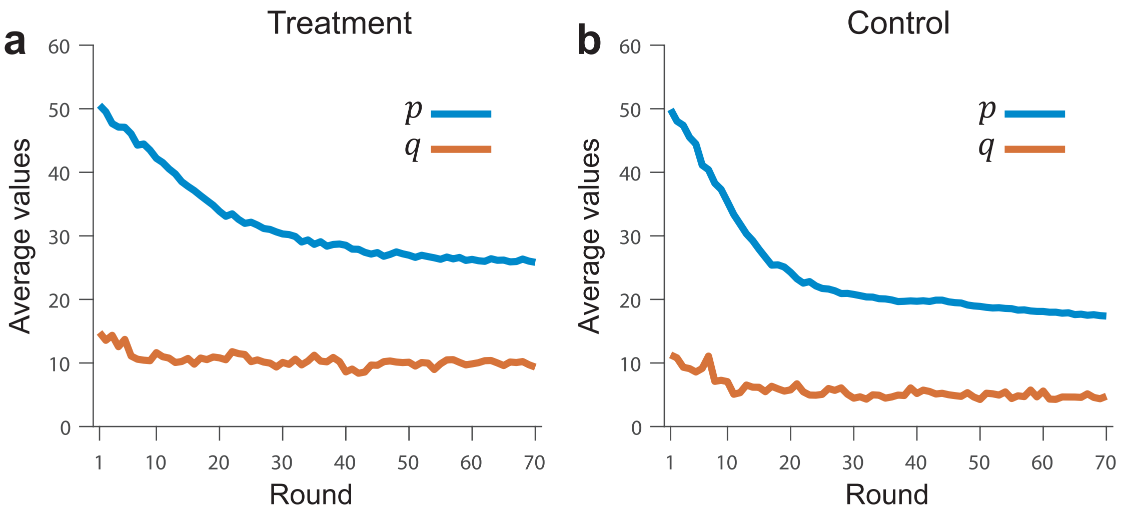
Supplementary Figure 13: Mean payoffs of the three types of behaviours, namely, altruistic behaviours $p > q$, empathic behaviours $p = q$ and selfish behaviours $p < q$. a, In the treatment groups, the mean payoffs of altruistic behaviours $p > q$, empathic behaviours $p = q$ and selfish behaviours $p < q$ are 90.99, 92.31 and 82.99. **b**, In the control groups, the mean payoffs of altruistic behaviours $p > q$, empathic behaviours $p = q$ and selfish behaviours $p < q$ are 92.61, 92.71 and 84.06. The mean payoff of altruistic behaviours $p > q$ is less than that of empathic behaviours $p = q$ in the treatment groups (Mann-Whitney U-test, P-value< 0.001), however, there is no big difference between the mean payoff of altruistic behaviours $p > q$ and empathic behaviours $p = q$ in the control groups (Mann-Whitney U-test, P-value$= 0.057$). Clearly, the mean payoffs of selfish behaviours are lowest in both treatment groups and control groups.



Supplementary Figure 14: Time evolution of actual decision time. Mean actual decision time from round 2 to round 70 in the four categories. Overall, the decision time decreases over rounds. Furthermore, the mean actual decision time in T3-T4 is 3 seconds shorter than T1-T2 and T5-T9, but there is no large difference between T3-T4 and C1-C2.



Supplementary Figure 15: Relationship between best-response behaviours and relative decision time. t denotes the actual decision time and $\bar{t}$ is the mean actual decision time in that round. In T1-T2 (a) and T5-T9 (c), we find there are U-shaped relationships between the proportion of best response and $t-\bar{t}$, i.e., both faster and slower decisions are more likely to be best-response. However, in T3-T4 (b), subjects do not have enough time to make a slow decision given the 30 seconds time limit. This explains why the proportion of best-response behaviours in T3-T4 is lower than T1-T2 and T5-T9. Finally, in C1-C2 (d), the correlation between best-response behaviours and $t - \bar{(t)}$ is not significant (Pearson correlation coefficient$= 0.0981$, P-value$= 0.4933$). This implies that in decision time may not affect best-response behaviours in a well-mixed population.



Supplementary Figure 16: Reinforcement learning simulation results. **a**, Mean acceptance levels q are calculated by randomly picking from the treatment database. Mean offers p are reproduced by using reinforcement learning model with fixed interaction structures. **b**, Mean acceptance levels q are calculated by randomly picking from the control database. Mean offers p are reproduced by using reinforcement learning model with well-mixed populations. The results are obtained by averaging 100 independent simulations.

3 Supplementary Tables

Supplementary Table 1: Details of experimental conditions. Our dual-role experiments include four conditions: two large treatment groups (T1 and T2) with structured populations, two median treatment groups (T3 and T4) with structured populations, five small treatment groups (T5-T9) with structured populations, and two control groups (C1 and C2) with well-mixed populations. In our analyses, the data of two subjects in C1 and one subject in T3 are excluded, thus there are only 48 subjects in C2 and 29 subjects in T3 to be analyzed.

	T1	T2	T3-T4	T5-T9	C1	C2
Number of subjects per group	50	53	30	10	50	58
Number of rounds	70	80	100	70	70	80
Decision time per round (sec)	45	45	30	45	45	45
Average income (Yuan)	94	103	119	94	95	104

Supplementary Table 2: A comparison between the theoretical predictions and empirical results Both the single-role and dual-role UG have the same (subgame perfct) Nash equilibrium. Single role MAO UG experiments found $(p, q) \approx (40, 30)$ no matter the population is well-mixed or structured [3]. Most theoretical studies based on evolutionary game theory considered dual-role MAO UG, and found that fixed interaction structures can promote fairness [6, 7].

	Well-mixed	Structured
Single-role Nash equilibrium	$(p, q) = (0, 0)$	$(p, q) = (0, 0)$
Single-role experiments	$(p, q) \approx (40, 30)$	$(p, q) \approx (40, 30)$
Dual-role Nash equilibrium	$(p, q) = (0, 0)$	$(p, q) = (0, 0)$
Our dual-role experiments	$(p, q) \approx (15, 5)$	$(p, q) \approx (20, 10)$
Dual-role evolutionary game model	$(p, q) = (0, 0)$	$(p, q) \approx (30, 30)$

Supplementary Table 3: Statistics results for decision time. Mean actual decision time in T1-T2 ($n = 103$), T3-T4 ($n = 59$), T5-T9 ($n = 50$) and C1-C2 ($n = 106$) from round 2 to round 70. Corr represents the correlation coefficient between actual decision time and rounds. The symbol "*" denotes that the correlation is strong, i.e. P-value$<$0.05.

	Mean	Corr
T1-T2	20.80	-0.8035^*
T3-T4	17.90	-0.7822^*
T5-T9	20.51	-0.8683^*
C1-C2	18.28	-0.9480^*

Supplementary Table 4: The mean values and standard deviations of offers and acceptance levels. We calculate mean values and standard deviations of p and q for all 70 rounds, and separately for rounds 1 to 35 and rounds 36 to 70. Mean(p) and std(p) represent the mean value and the standard deviation of offers of all proposers, respectively, in which a proposer's offer p is taken as the average of his/her offers p over 1-70 rounds/ 1-35 rounds/ 36-70 rounds. Similarly, mean(q) and std(q) represent the mean value and the standard deviation of acceptance levels q of all responders, respectively, in which a responder's acceptance level is taken as the average of his/her acceptance levels q over 1-70 rounds/ 1-35 rounds/ 36-70 rounds.

	1-70 rounds/ 1-35 rounds/ 36-70 rounds			
	mean(p)	std(p)	mean(q)	std(q)
T1-T2	21.66/25.21/18.10	7.11/10.91/5.91	9.86/10.47/9.25	7.04/9.49/7.37
T3-T4	26.14/31.01/21.27	12.93/14.82/12.55	11.90/13.18/10.63	10.87/12.79/11.00
T5-T9	21.12/23.30/18.94	6.67/8.08/6.36	9.97/9.50/10.44	6.51/6.71/7.29
C1-C2	17.78/22.52/13.03	5.40/8.65/3.27	5.70/6.50/4.90	5.82/7.35/5.32

Supplementary Table 5: Mann-Whitney U-test for offer p and acceptance level q. Statistics results of Mann-Whitney U-test for offer p and acceptance level q ($n = 103$ in T1-T2, $n = 59$ in T3-T4, $n = 50$ in T5-T9, and $n = 106$ in C1-C2). A subject's offer (or acceptance level) is taken as the average of his/her p (or q) over 70 rounds. The symbol "*" denotes that the mean values of two groups are significantly different, i.e. P-value<0.05.

P-value for p	T1-T2	T3-T4	T5-T9	C1-C2
T1-T2	1	0.0288*	0.9133	< 0.001*
T3-T4	0.0288*	1	0.0468*	< 0.001*
T5-T9	0.9133	0.0468*	1	< 0.001*
C1-C2	< 0.001*	< 0.001*	< 0.001*	1
P-value for q	**T1-T2**	**T3-T4**	**T5-T9**	**C1-C2**
T1-T2	1	0.7356	0.8641	< 0.001*
T3-T4	0.7356	1	0.9806	< 0.001*
T5-T9	0.8641	0.9806	1	< 0.001*
C1-C2	< 0.001*	< 0.001*	< 0.001*	1

Supplementary Table 6: Mann-Whitney U-test for proportion of best-response behaviours. Statistics results of Mann-Whitney U-test for proportion of best-response behaviours ($n = 103$ in T1-T2, $n = 59$ in T3-T4, $n = 50$ in T5-T9, and $n = 106$ in C1-C2). The symbol "*" denotes that the mean values of two groups are significantly different, i.e. P-value<0.05.

P-value for p	T1-T2	T3-T4	T5-T9	C1-C2
T1-T2	1	< 0.001*	0.1182	< 0.001*
T3-T4	< 0.001*	1	< 0.001*	0.0028*
T5-T9	0.1182	< 0.001*	1	0.003*
C1-C2	< 0.001*	0.0028*	0.003*	1

4 Supplementary References

[S1]. Ellison G. Learning, local interaction, and coordination. *Econometrica* **61**, 1047–1071 (1993).

[S2]. Brenner T., Vriend N. J. On the behavior of proposers in ultimatum games. *J. Econ. Behav. Organ.* **61**, 617–631 (2006).

[S3]. Han, X. *et al.* Emergence of Communities and Diversity in Social Networks. Natl. Acad. Sci. USA. **114**, 2887–2891. (2017)

[S4]. Fischbacher U. z-Tree: Zurich toolbox for ready-made economic experiments. *Exper. Econ.* **10**, 171–178 (2007).

[S5]. Sutton R. S. & Barto A. G. *Reinforcement learning: An introduction.* (Cambridge: MIT press, 1998).

[S6]. Sinatra, R. *et al.* The ultimatum game in complex networks. *J. Stat. Mech. Theory Exp.* **2009**, P09012 (2009).

[S7]. Iranzo, J., Floría, L. M., Moreno, Y. & Sánchez, A. Empathy emerges spontaneously in the ultimatum game: small groups and networks. *PLoS ONE* **7**, e43781 (2012).