PLOS ONE

RESEARCH ARTICLE

Playing nice in the sandbox: On the role of heterogeneity, trust and cooperation in common-pool resources

Fijnanda van Klingeren [ID]*

Nuffield College, University of Oxford, Oxford, United Kingdom

* fijnanda.vanklingeren@nuffield.ox.ac.uk

Abstract

The increasing heterogeneity of populations affects cooperation in common-pool resources in a time where the depletion of natural resources is a growing problem. This study investigates the effects of economic and sociocultural heterogeneity on trust and cooperation in common-pool resources using a laboratory experiment. The experiment comprises two Investment Games and a Common-Pool Resource Game, with a sample of 344 subjects from the United Kingdom and the Netherlands. By measuring the effects of economic and sociocultural heterogeneity separately as well as combined, this study disentangles the effects of specific heterogeneity types on cooperation in common-pool resources; something that has not been done before. Higher levels of trusting behaviour are found to have a positive effect on cooperation on the micro- and macro-level over time. While theory suggests negative effects of both forms of heterogeneity on cooperation through decreased levels of trust, the results show a surprising positive effect of economic heterogeneity on cooperation, but a negative effect if economic and sociocultural heterogeneity are combined. This study concludes that economic inequality can promote cooperation in CPRs, unless this inequality is lined up with sociocultural differences.

Introduction

Societies are becoming more diverse on ethnic, cultural and economic dimensions due to growing migrant populations all over the world, especially in Northern Africa, Western Asia and Sub-Saharan Africa [1]. This increasing heterogeneity may pose a challenge for the successful management of common-pool resources [CPRs] as it can diversify interests and decrease trust between appropriators [2]. Especially when there are multiple larger groups of different sociocultural backgrounds living together, intergroup antagonism becomes stronger and coordination between groups becomes harder [3]. The deforestation of tropical forests and overfishing of the seas are examples of appropriators failing to work together on the preservation of natural resources. However, how and to what extent economic and sociocultural heterogeneity, and importantly the combination of the two, affect cooperation is still contested [4–12]. In particular, experimental research looking into the effect of heterogeneity on

cooperation in CPRs is still relatively rare [13]. The aim of this paper is to provide insights into the relation between economic and sociocultural heterogeneity and sustainable cooperation in CPRs, both on the individual and the collective level. To do so, this study employs an Investment Game and a CPR game in a computerised laboratory experiment. Since part of the theoretical mechanism is often suggested to be the negative influence of heterogeneity on trust [14–18] and the beneficial influence of trust on society [19–25], this paper will consider trust as an important variable in the theoretical framework. Disentangling effects of economic and sociocultural heterogeneity using experiments in CPR settings has, to the extent of my knowledge, never been done before.

Two key characteristics of CPRs are their non-excludability and high subtractability: it is hard to exclude potential users from accessing CPRs, and the resource may run out [26]. These characteristics make CPRs vulnerable to the 'tragedy of the commons' as famously described by Hardin [27]: a situation in which the CPR is locked into a system that provides each resource-user incentives to use the limited resource unlimitedly, which will lead to its inevitable decay. From a game theoretic perspective, this tragedy will always take place, as a (myopic) rational individual will free-ride and overexploit the resource despite the long-term benefits of cooperation. Although there are plenty of examples of unsustainable appropriation of CPRs (see for instance [28, 29]), Hardin's prediction is challenged by evidence from field research suggesting that CPR users are able to self-organise using institutions for collective action [30, 31]. In addition, an extensive body of research suggests more complex behaviour than predicted by rational choice theory [13, 29, 32–35]. With examples of both successful and failed CPRs, it is interesting and important to investigate the role that economic and sociocultural heterogeneity in communities surrounding CPRs play in the success or failure of a CPR.

Empirically analysing the influence of heterogeneity in real-life CPR settings can be a challenge due to the number of confounding factors that influence success and failure; one can never be entirely sure that one variable or a set of variables caused an outcome [36]. This can be solved to a large extent by using laboratory experiments. While laboratory experiments score lower on external validity than field experiments, they score high on internal validity: they allow one to test causality by measuring the impact of an isolated variable or set of variables repeatedly, in the same controlled setting—something that is not possible in field research [36, 37].

For the current study, subjects first played an Investment Game [IG] to measure general trust and trustworthiness and then played a CPR game as introduced by Janssen, Holahan and Ostrom [31], to measure behaviour in a CPR setting under different levels of heterogeneity. The CPR game mimics a fishing ground which the subjects can appropriate in return for money in groups of four under different conditions of economic and sociocultural heterogeneity. Economic heterogeneity is introduced as unequal endowments of players, and sociocultural heterogeneity by means of a Minimal Group Experiment [MGE]. An MGE is a method to create artificial identities based on a trivial criterion (see amongst others [38–41]). An extensive explanation of the MGE used in this paper is given in the section on experimental treatments.

Given the increasing depletion of CPRs world-wide and the rising levels of sociocultural and economic heterogeneity, the subject of heterogeneity and sustainable cooperation is gaining importance. The results of this study may be relevant not only for classic resources such as fishing grounds, but also for the growing number of contemporary commons such as food cooperatives and green energy initiatives [30, 42, 43].

The paper is structured as follows. First, existing literature on the research topic will be explored and hypotheses based on the literature are derived. Second, the experimental proceedings, games and treatments will be described. Third, descriptive plots, non-parametric test

and multilevel regression are employed to analyse the data. Lastly, expectations are revisited, the findings are discussed and a conclusion is formulated.

Existing literature

Heterogeneity and cooperation

Experimental research looking into the effects of economic and sociocultural heterogeneity on cooperation suggests that asymmetrical endowments—i.e. an unequal division of money or points to spend between players in a group—lead to unequal contributions, unequal payoffs, and Pareto suboptimal outcomes [44]. Furthermore, it is argued that economic inequality leads to an increase of transaction costs [25] and diversification of interests among individuals, which makes cooperation less likely to happen [45–47]. For Public Good games it was found that heterogeneity in endowments indeed leads to a lower contribution to the public good [13, 35, 48]. This may be caused by an "anticipated reciprocity" effect: a situation in which subjects with a lower endowment expect the subjects with higher endowments to invest more, since they have more means available to invest, while subjects with higher endowments do not in fact do so [13].

Olson [9] suggests an opposite effect of economic heterogeneity: he argues that when groups are sufficiently small, and inequality sufficiently large, economic inequality leads to inequality of incentives, which makes the rich incentivised enough to bear the burden of cooperation by themselves. However, Bardhan and Dayton-Johnson [6] argue that this will only hold if there are "non-convexities" in the CPR, such as maintenance of the CPR, restraints on appropriation or large start-up costs. Most case-study literature still suggests a negative effect of economic heterogeneity [49, 50], despite the theoretical possibility as sketched by Olson.

Theory on the relation between sociocultural heterogeneity and cooperation suggests that individuals are more likely to cooperate with others from their ingroup: individuals with whom they share strong, multi-stranded relationships and common interests [15, 26, 51, 52– 60]. Several case studies show that heterogeneity between CPR appropriators in terms of ethnicity, use of the resource, and view on the resource can be cause for conflict and hampers the development of regulation [29, 61]. Experimental research shows that (induced) group identity leads to positive behaviour towards ingroup members relative to outgroup members [62] and to the prioritising of group interest over individual interests [63]. Timilsina, Kotani and Kamijo [64] show in their research on sustainability of CPRs that subjects from urban areas show less prosocial and sustainable behaviour in a CPR game than subjects from rural areas. They suggest that subjects from urban areas are less prosocial as they come from a more heterogeneous and anonymous environment, whereas subjects from rural areas are more homogeneous and have a long tradition of necessary mutual cooperation. Habyarimana, Humphreys, Posner and Weinstein [65] suggest that people with the same ethnicity cooperate due to an increased "findability" in the ingroup through norms and tight social networks, and thus a higher probability of being punished for defection. Their research, identifying subjects as specific player "types", suggests that homogeneity increases cooperation levels even for player types that are least likely to cooperate. Next to this, research shows that people have strong expectations of cooperation when interacting with ingroup members as opposed to outgroup members, which makes them more likely to cooperate themselves [66, 67]. However, Yamagishi and Kiyonari [67] show that in order for people to have these higher expectations of cooperation and to act upon it, it is necessary for players to 1) know the ingroup or outgroup identity of the other players; and 2) to know that the other players are aware that everyone knows everyone's identity.

There is also research illustrating that sociocultural heterogeneity does not always have a negative impact on cooperation. Gehrig, Schlüter and Hammerstein [68] for instance, show in their study of fishermen from a small-scale fishery in Zanzibar that heterogeneous groups of fishermen from different villages do not cooperate less than homogeneous groups, despite a history of conflict between the villages. They argue that the effect of sociocultural heterogeneity on cooperation may be dependent on the institutional scope in the economic domain. In addition, Varughese and Ostrom [60] argue that heterogeneity does not influence cooperation when the right institutional arrangements are in place. However, the ease with which institutional arrangements can be set in place may depend on the type of heterogeneity. Bazzi, Gaduh, Rothenberg and Wong [3] show in their study of a population resettlement program in Indonesia that in the context of polarisation (a situation with a few larger groups with different sociocultural backgrounds) public goods provision is reduced, the likelihood of ethnic conflict is increased and economic development is hampered. In the context of fractionalisation (a situation with many smaller groups) these negative effects of sociocultural heterogeneity are not found.

Trust

To get a better understanding on how heterogeneity affects cooperation, trust is considered to be a mediating variable. There is extensive evidence that heterogeneity reduces trust [14, 69–72], while higher trust yields higher levels of cooperation [2, 20, 21, 73]. This implies that individuals trust others with a similar identity—for instance religion, ethnicity, culture, social identity or something else—more than others with a different identity [14, 16, 18, 20, 24, 25, 74].

In the particular case that is studied here—a fishing ground—trust is necessary to maintain cooperation. Real fishermen do not know how much fish the other fishermen are catching during a day out fishing at sea, and will only see or hear about each other's catch when all the ships have returned. In the same way, players in the experiment do not know what the others are doing during the appropriation stage and only receive information about this at the end of the period. Like fishermen, players will have to trust each other to behave cooperatively during the appropriation stage due to a lack of mutual monitoring.

Hypotheses

Based on the majority of the discussed literature, the following hypothesis is deduced with regard to the direct effect of heterogeneity on cooperation in CPRs:

Hypothesis 1 *(a) Economic and (b) sociocultural heterogeneity have a negative direct effect on cooperation over time and (c) the coincidence of economic and sociocultural heterogeneity has an even stronger negative direct effect on cooperation over time.*

The following hypotheses are deduced with regard to the indirect effect of heterogeneity on cooperation in CPRs, through trust:

Hypothesis 2 *(a) Economic and (b) sociocultural heterogeneity have a negative effect on trust and (c) the coincidence of economic and sociocultural heterogeneity has an even stronger negative effect on trust.*

Hypothesis 3 *Trust has a positive effect on cooperation over time.*

Materials and methods

Experimental sessions

A computerised laboratory experiment was designed and programmed in z-tree [75]. A total of 344 subjects of age 18 and older participated in the experiment, which was conducted at the

Centre for Experimental Social Sciences [CESS] at Nuffield College, University of Oxford between October 2018 and November 2019, and at the Experimental Laboratory for Sociology and Economics [ELSE] at Utrecht University in April 2019. The subjects for both laboratories were recruited from the Online Recruitment System for Economic Experiments [ORSEE] [76]. After a pre-test with 16 Oxford students, the experiment was held in 13 sessions at CESS containing 248 subjects, and 5 sessions at ELSE containing 96 subjects. Sessions contained 16, 20 or 24 subjects. 81% of the subjects were students, from varying disciplines and years/stages. 60% of the subjects were female, and the average age was 26. This research, including the pre-test, has obtained ethics clearance from the Research Ethics Committee of Department of Sociology (DREC) at the University of Oxford (Ref. SOC_R2_001_C1A_18_30), the CESS ethics committee (Ref. LE_0044) and was covered by the ethical approval (Ref. FETC17-028, Buskens) of the ethical committee of the Faculty of the Social and Behavioural Sciences at Utrecht University. Written consent was obtained from all subjects in the study, including the pre-test, before the start of each experimental session. The data was anonymised before the analyses.

The laboratories are very similar in physical setup: a room with 25 to 30 computers, separated by privacy dividers, with a designated area for the experimenter computer. The experimental procedures, including having participants wait in a waiting room, dividing subjects randomly over computers, and handing out instructions with help of a research assistant, were formalised in the exact same way for sessions in the UK and the Netherlands.

The sequence of the experiment is as follows. When subjects enter the lab, they are given general written instructions in English for the first part of the experiment, which entails (1) a one-shot Investment Game; (2) a basic, practise version of the Fishing Game without any treatments for 3 periods; (3) a Minimal Group Experiment; (4) a binary other-other Dictator Game; and (5) a quiz to strengthen group bonds; and lastly (6) two other one-shot rounds of the Investment Game, once with an ingroup member and once with an outgroup member. An extensive explanation of each of these parts is provided in the upcoming sections. After completing the first part of the experiment, subjects receive instructions for the second part of the experiment: the Fishing Game with heterogeneity treatments. Subjects receive instructions specific to their treatment; economic heterogeneity [EH], sociocultural heterogeneity [SH], economic and sociocultural heterogeneity [EHSH] and the control treatment with no heterogeneity [NH]. The general and treatment-specific instructions can be found in S1 and S2 Text respectively. Completing the experiment took 50 to 70 minutes, of which 20 to 30 minutes were spent on the games before the Fishing Game. Fig 1 shows a schematic overview of the separate parts of the experiment. Subjects played for real money (GBP in the UK and EUR in the Netherlands) under an exchange rate of 500 units = 1 GBP/EUR. The average earning was 15.11 GBP/EUR.

The Investment Game

The game that is used to measure trust before the main experiment is a variation of an Investment Game, as designed by Berg, Dickhaut and McCabe [77]. The Investment Game [IG], also called the Trust Game, is the most frequently used game to study trust [78].

The current version of the game is played as follows. There are two roles: the sending and the receiving role. Both players first adopt the sending role, and are given an endowment of 10 points. Both players are given the choice of sending points to another player, ranging from 0 to 10 points. This amount will be tripled once before it reaches the other player. Next, both players are put in the receiving role; they are asked how many of the points received by a sending player they would send back, for every possible amount of points received. This amount ranges from 0 to 30 points in increments of 3. 30 points would be the maximum amount to be
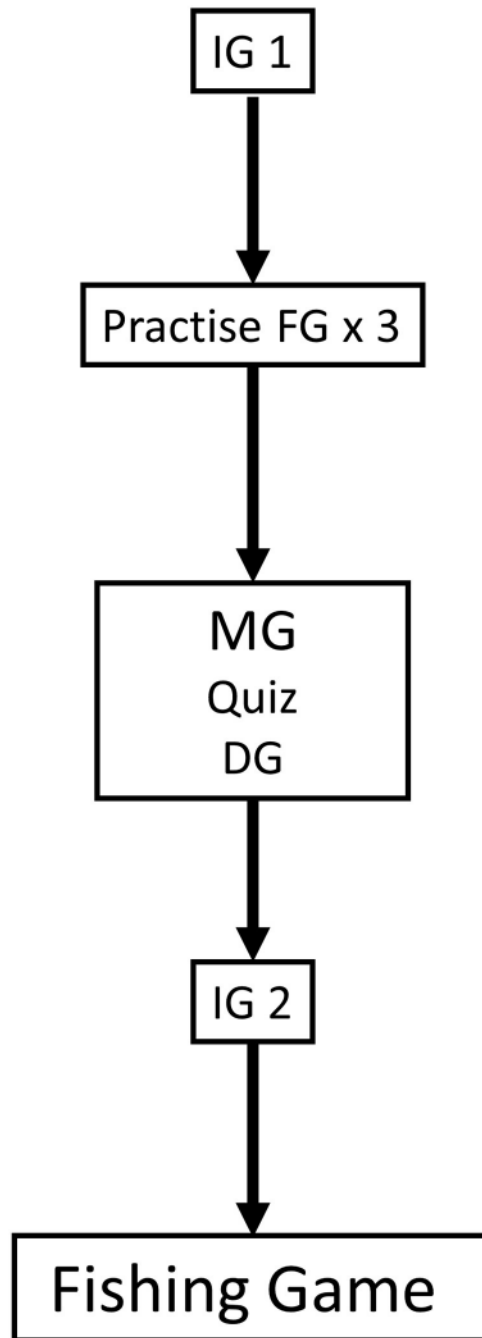
Fig 1. Course of the experiment.

received by the receiving player, since the maximum amount of points that the sending player can send is 10, and $3 \times 10 = 30$. Making players choose how many points to return for every possible received amount is called the strategy method [79, 80]. This method provides the advantage to the researcher of having data on trust and trustworthiness for all subjects, as the player 2 decision can be made separately from another subject's player 1 decision. After playing the game in each of the roles, the subjects are randomly assigned the role for which they will receive their payoff. They are matched to another player with the other role. Their

previous actions decide their final payoff. The material utility functions for the players are as follows: For player 1, the sender/trustor, the general material utility payoff function is:

$$U_i = E_i - S_{ij} + R_{ji}$$

Where E_i is the initial endowment of sender i, S_{ij} is the amount of points sent from the sender i to the receiver j, and R_{ji} is the amount of points returned from the receiver to the sender. For player 2, the receiver/trustee, the general material utility function is:

$$U_j = E_j + 3 \times S_{ij} - R_{ji}$$

Where E_j is the initial endowment of receiver j, S_{ij} is the amount of points sent from the sender i to the receiver j that is multiplied by 3 by the experimenter, and R_{ji} is the returned amount from the receiver to the sender. In the current game, $E = 10$ for both players. The two trust variables, operationalised following Johnson and Mislin [80] will be measured as follows:

$$Trust = \frac{number\ of\ points\ sent\ by\ i}{endowment\ of\ i} = \frac{S_{ij}}{E_i}$$

$$Trustworthiness = \frac{number\ of\ points\ returned\ to\ i\ by\ j}{number\ of\ points\ available\ to\ return\ to\ i\ by\ j} = \frac{R_{ji}}{E_j + 3 \times S_{ij}}$$

A graphic representation of an interaction between two matched players in the Investment Game is presented in S1 Fig. Half of the subjects will be paid for the sending role, and half for the receiving role in each Investment Game. Trustworthiness for all subjects will be calculated as the average trustworthiness over the 10 decisions every subject makes as the receiving player, as facilitated by using the strategy method. There are many variations of the Investment Game. A list and explanation of the specific characteristics of the Investment Game used in this experiment can be found in S3 Text.

The Fishing Game

The game to be played by participants of the experiment is a CPR game. While it is common to measure cooperation, trust and heterogeneity in games such as Public Good [PG] games (see for instance Aksoy [81]), there are fundamental differences between CPRs and PGs that should be taken into account when looking specifically at CPR situations. Two characteristic features of a CPR situation are that exclusion of the collective good is infeasible—for instance, it is very costly to fence off part of an ocean—and that subtractability is high—the resource is finite and can run out [26, 29]. Table 1 shows the classification of different types of goods as shown by Ostrom, Walker and Gardner [26].

The CPR in the current game is a fishing ground. In the game, there are four appropriators that use the CPR, who will play the game with each other for the entire session. All subjects first practise the basic CPR game without treatments for three periods, without any consequences for their payoff, to get to know the game. The real game with treatments is played for

Table 1. A classification of goods (Ostrom et al., 1992, p.7).

		Subtractability	
		Low	**High**
Exclusion	Difficult	Public Goods	common-pool Resources
	Easy	Toll Goods	Private goods

40 periods. A time span of 40 periods is long enough for subjects to see the resource fall into decay if they overexploit it systematically, and there is enough time for subjects to adjust their investments to regrow the resource again.

Appropriation of the resource. At the beginning of each period t, the appropriators all receive an endowment E of units to invest in appropriation of the resource, R. Since appropriation of the CPR is a costly activity—e.g. it takes time and requires maintenance of the boat and fishing nets—the appropriation effort a, where $0 \leq a \leq E$, represents the amount of effort an appropriator can invest in appropriation of the CPR.

The actors can choose how much they want to invest in appropriation of the resource each period. All appropriators make their appropriation choice at the same time, without knowing what the other appropriators do that period. They see how many fish there are in the lake and how many fish they receive per invested unit a. The appropriators all receive the same return $\left(\frac{4}{R_0} R_{t-1}\right)$ per appropriation effort unit of a. The material utility function for the appropriators per period is as follows:

$$U_{it} = \left(\frac{4}{R_0} R_{t-1}\right) a_{it} + (E_i - a_{it})$$

U_{it} is the total material utility of an appropriator i at timepoint t. In the function, a_{it} is the invested appropriation effort of appropriator i at timepoint t, and E_i is the endowment of appropriator i, which is the same every period. R_0 is the original resource size of the CPR (i.e. the maximum number of fish in the lake) for which $R_0 = 600$ is taken. R_{t-1} is the resource size at time $t - 1$. The profit per invested appropriation effort unit of a is thus dependent on the current size of the resource, relative to its original size. If $R_{t-1} = R_0$, which is the case in the first period of the game, the return is $4 - 1 = 3$ units per invested unit of a. When $R_{t-1} < R_0$, the return will be lower than 3 units. The amount of appropriators' endowment not used for fishing is reflected by $(E_i - a_{it})$; players will thus keep the part of the endowment that they did not invest in appropriation as profit for that period.

At the end of the period, the players see how much they invested themselves; their profit from investing; and how much was invested in appropriation of the resource in total as a group. At the beginning of the next period, they also see how much each individual player in their group invested in previous periods.

Resource renewal. Just like real natural resources, the resource in the game has a renewal rate. The renewal rate per period is modelled as follows:

$$R_t = \min\left(600, 1.25\left(R_{t-1} - \left(\frac{R_{t-1}}{R_0}\right)\sum_{i=1}^{4} a_{it}\right)\right)$$

Here, 1.25 is the renewal rate of the resource and R_t is the resource size at timepoint t. The amount of fish in the lake is thus multiplied by 1.25 after each period. The maximum resource size is $R_t = 600$; this is the maximum amount of fish in the lake and the resource cannot grow beyond this size. The sum of appropriation effort of all four appropriators is indicated by $\sum_{t=1}^{4} a_{it}$.

Overexploitation. The CPR is overexploited—that is, the pool of fish in the lake is smaller than in the previous period—when $R_t < R_{t-1}$, so when the resource size in timepoint t is smaller than in the previous timepoint. This happens if $\sum_{t=1}^{4} a_{it} > 120$ because this is the limit of sustainable appropriation based on $R_0 - \frac{R_0}{1.25}$. The CPR is thus overexploited when the four appropriators have invested on average 30 units in appropriation effort per person $\left(600 - \frac{600}{1.25} = 120\right)$. Investing stops being profitable if $R_t = \frac{R_0}{4}$, so if the resource size decreased
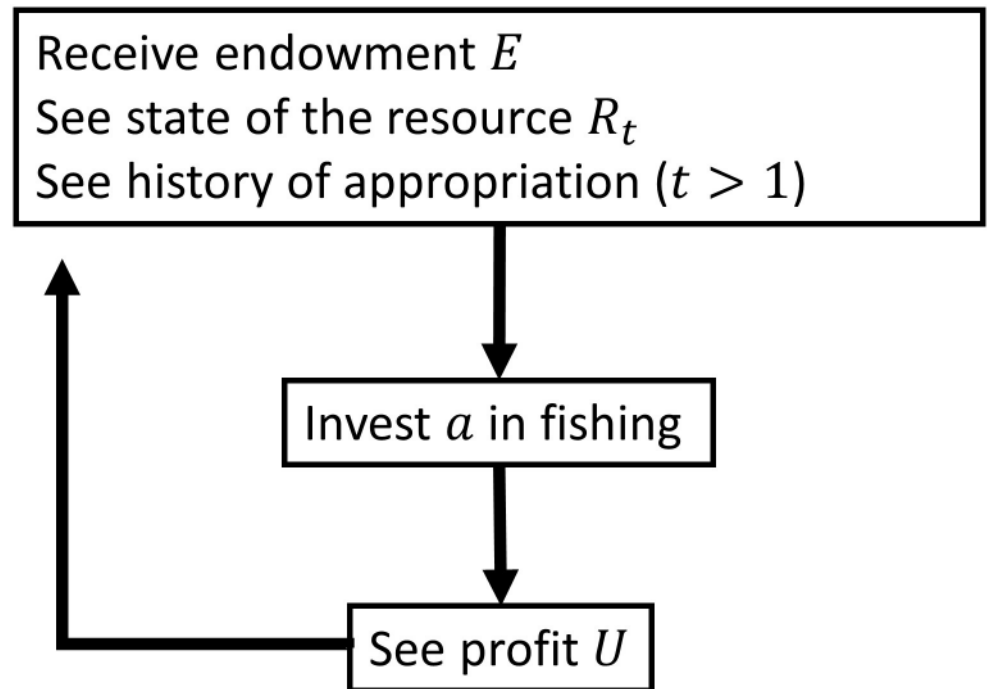
Fig 2. The different stages in a Fishing Game period.

to 25 per cent of the original resource size ($R_t = 150$), because:

$$U_{it} = \left(\frac{4}{600} 150 \right) a_{it} + (E_i - a_{it}) = E_i$$

When this happens, any amount the appropriator invests in appropriation of the resource will result in a return of exactly that amount, and the profit consisting of the return plus the left-over endowment will thus result in $U_{it} = E_i$. For example, if the appropriator has $E_i = 50$ and $a_{it} = 50$, the return will be $50 + 0 = 50$; if the appropriator invests 10 the return will be $10 + 40 = 50$ etc. When the resource size drops below 25 per cent of the original size, appropriators make loss by investing in appropriation. Only when the size of the resource increases again will the multiplication of the invested unit of a increase and fishing will become relatively more profitable. Graphs visualizing the development of the resource size under different levels of appropriation over time, and the decreasing marginal profits of overexploitation are shown in S2 and S3 Figs respectively.

In real life, fishermen may not notice the decline of the resource size within a day's time. However, having limited time for a laboratory experiment, the game unfolds as if the process of depletion and renewal was sped up. With a total number of 40 periods the experiment covers enough time to capture long-term behaviour. Fig 2 schematically visualises the different stages in a period of the Fishing Game.

Trust. Trust is measured at three points in the experimental session: 1) by the one-shot Investment Game with a random other subject at the very beginning of the experimental session, before the MGE and the main game; 2) after the MGE when subjects play the one-shot Investment Game once with an ingroup member and once with an outgroup member; and 3) by asking questions on trust in a post-experimental survey. The one-shot Investment Game with a stranger will be used to measure general trust. Even if this type of trust does not

resemble the mutual trust necessary for repeated games (such as the current CPR game) exactly, it is still useful as an indicator of general trustfulness of subjects before they enter the main game.

Experimental treatments

Sociocultural heterogeneity: The MGE.   To test the effect of sociocultural heterogeneity on cooperation, an MGE is used to create artificial identities which are based on a trivial criterion [38–40, 67]. Following the approach of amongst others Tajfel, Billig, Bundy and Flament [39], Masella, Meier and Zahn [41] and Aksoy [40, 81], the subjects are shown five paintings by two artists, Paul Klee and Wassily Kandinsky, after which they are asked to express their preference of either picture, resulting in a score of 1 to 5 for Klee-preference. Based on the median preference of the particular experimental session, subjects are assigned to the Klee or Kandinksy group. Tajfel et al. [39], in one of the first published studies using the MGE, divide subjects randomly over groups, regardless of their Klee or Kandinsky preferences. Even random allocation led to subjects trying to maximise ingroup outcomes. However, following Aksoy [40] and specifically using the method of Masella et al. [41], subjects in this study were divided on the actual outcomes of their choices. This approach was taken for several reasons. Firstly, this was done to avoid deception of subjects. Secondly, letting subjects go through a process where they are divided into groups based on a real characteristic, namely preference for Klee or Kandinsky, could add to the feeling of belonging to a group. Thirdly, in the rare case a subject would have a profound preference for either painting, knowing which painter painted which painting, allocating them by chance in the wrong group would decrease the power of the experiment.

After grouping subjects into Klees and Kandinskys, this study follows Aksoy [40] by enhancing group identities with a quiz and an other-other Dictator Game. In the quiz, players have to guess the painter (Klee or Kandinsky) of three paintings. High group performance is profitable: if more than half of the answers of the ingroup are right and/or if the ingroup has more right answers than the other group, players from that group get extra points. The extra points from this part of the experiment will be shown at the end of the CPR game, in order to avoid that low group performance in the quiz influences behaviour in the CPR game [40]. Next, subjects play three periods of a binary other-other Dictator Game, as described by amongst others Aksoy and Weesie [82] and Bilancini, Boncinelli, Capraro, Celadin and Di Paolo [83]. In this game, players face three scenarios in which they have to divide points between a Klee and a Kandinsky player (an ingroup and an outgroup member). In each of the three scenarios, they can opt to divide points equally or unequally between both other players. The purpose of this game is to make players aware of their group membership. Each experimental session, one scenario is randomly picked to be a paid interaction in which the choice that players make influences the amount of points that will be sent to a Klee and a Kandinsky player. Subjects know that any one of the three scenarios will be paid. The three scenarios with the available options in the other-other DG are provided in S4 Text.

An alternative way to induce sociocultural heterogeneity would be to use natural identities based on for instance gender, religion or ethnicity [84, 85]. However, it is not necessarily known to what extent, if at all, a subject identifies with their natural identity. Next to that, it is unpredictable how natural identities will respond to experimental manipulations, and there are many other factors that may vary with natural identity that may influence behaviour [40]. Induced identities on the other hand, are fully controllable and unambiguous, allowing for a bigger confidence that any behavioural differences between subjects are indeed caused by the treatment itself [40]. Even though the groups are based on an artificial criterion, plenty of

research shows that the feeling of belonging to a group, no matter on what basis categorisation takes place, is enough to create social identity [39, 40, 59, 67, 86].

It was checked and confirmed that being a Klee or a Kandinsky itself did not significantly affect the outcomes of interest, that is, no underlying behaviour was associated with becoming a Klee or Kandinsky.

Economic heterogeneity. Economic heterogeneity is induced by varying the endowment E_i that players receive at the start of each period. Under economic homogeneity, all appropriators receive $E_i = 50$ to invest in appropriation. Under economic heterogeneity, however, two appropriators receive $E_i = 40$ and two appropriators receive $E_i = 60$ (See [13] for a similar operationalisation of economic heterogeneity based on variations in endowment). The total endowment of the group is 200 for all groups in the experiment.

Four combinations. The four treatments that are applied in the experiment are shown in Table 2. Subjects know their own endowment and the endowment of others in their group; just as they know their own preference identity (Klee or Kandinsky) and the preference identity of others in their group. They see all this information in a box on the screen every period.

Randomisation of subjects across groups in the main game was organised as follows. First, a rank order was created with subjects ranked on how many Kandinsky paintings they preferred out of five paintings. To prevent subjects sharing a rank, a random number between 0 and 1 was added to their rank, and the list was ordered again. From this ordered list, the bottom half was declared a Klee, and the top half Kandinsky. Groups were then assembled according to the treatment for that group; homogeneous for the EH and NH treatments (all Klees or Kandinskys) or heterogeneous for the SH and EHSH treatments (two Klees and two Kandinskys). Groups are assigned to treatments manually, depending on the number of subjects in a session and the total number of subjects in each treatment from earlier experiments.

Note that in the EHSH treatment, economic heterogeneity is lined up with sociocultural heterogeneity; that is, two members of the same identity group receive $E_i = 60$ and the two other members, who are both members of the other identity group, receive $E_i = 40$. In the EH treatment, two randomly chosen players receive $E_i = 40$ and the other two receive $E_i = 60$. This design does not include a treatment where economic and sociocultural heterogeneity are not lined up: economic heterogeneity is only found within and between groups, not both. This was done for two reasons. Firstly, by only varying heterogeneity within and between groups, the experimental design is kept simpler; adding a fifth treatment would result in fewer subjects in each treatment, or the need for a larger sample size to retain statistical power. Secondly, lining up economic and sociocultural heterogeneity enables one to distil whether it is economic or sociocultural heterogeneity that plays a role in certain conflicts between appropriators, such as the conflicts between Japanese and Chinese fishermen between 1920 and 1930 and after World War II: two groups who differed greatly in culture and in fishing assets [87, 88]. However, this

Table 2. Overview of treatments.

	Treatment	Operationalisation
EH	Economic heterogeneity	Different endowments
	Sociocultural homogeneity	Same preference identities
SH	Economic homogeneity	Same endowments
	Sociocultural heterogeneity	Different preference identities
EHSH	Economic heterogeneity	Different endowments
	Sociocultural heterogeneity	Different preference identities
NH	Economic homogeneity	Same endowments
	Sociocultural homogeneity	Same preference identities

does not mean that there are no cases where economic and sociocultural heterogeneity are not lined up. An interesting experiment that varies the overlap between economic and sociocultural heterogeneity is Aksoy [81].

Analytical strategy

Functions. A multilevel regression framework is deployed to test the hypotheses outlined above. For the macro-outcome resource size, a two-level multilevel model will be fitted with period-level outcomes (level 1) and random intercepts for groups (level 2). Individual characteristics will be aggregated to the group level—i.e. the group mean of general trust from the IG, sex, age, game theory experience etc. will be taken. The model is represented in the following function:

$$
\begin{aligned}
y_{tj} = \alpha_j \quad & + \sum_{k=1} \beta_k x_{jk} \\
& + \sum_{\tau=1} \psi_\tau z_{j\tau} + \omega g_j \\
& + \phi_1 t + \phi_2 t^2 \\
& + \sum_{\tau=1} \theta_\tau (z_{j\tau} \times t) + \sum_{\tau=1} \lambda_\tau (z_{j\tau} \times t^2) + \xi(g_j \times t) \\
& + e_{tj};
\end{aligned}
$$

with $\alpha_j \sim N(\mu^\alpha, \sigma_\alpha^2)$ and $e_{tj} \sim N(0, \sigma^2)$. α_j indicates the intercept for groups. There are t periods for j groups; k control variables x with coefficient β; τ treatments z with coefficients ψ. g_j represents average general trust as measured by the IG per group j with coefficient ω. Finally, an interaction of treatments with period and the quadratic term of period with respectively coefficients θ and λ, and an interaction of average general trust with period with coefficient ξ are included.

For the micro-outcome individual appropriation effort, a three-level multilevel model will be fitted with period-level outcomes (level 1) and random intercepts for individuals (level 2) and groups (level 3), as represented in the following function:

$$
\begin{aligned}
y_{tij} = \Psi_i + \alpha_j \quad & + \sum_{k=1} \beta_k x_{ijk} \\
& + \sum_{\tau=1} \psi_\tau z_{ij\tau} + \omega g_{ij} \\
& + \phi_1 t + \phi_2 t^2 \\
& + \sum_{\tau=1} \theta_\tau (z_{ij\tau} \times t) + \sum_{\tau=1} \lambda_\tau (z_{ij\tau} \times t^2) + \xi(g_{ij} \times t) \\
& + e_{tij};
\end{aligned}
$$

with $\Psi_i \sim N(\mu^\Psi, \sigma_\Psi^2)$, $\alpha_j \sim N(\mu^\alpha, \sigma_\alpha^2)$ and $e_{tij} \sim N(0, \sigma^2)$. Ψ indicates the intercept for individuals, and α indicates the intercept for groups. There are t periods for i individuals in j groups; k control variables x with coefficient β; τ treatments z with coefficient ψ. g_{ij} represents individual general trust as measured by the IG for person i in group j with coefficient ω. An interaction of treatments and period and the quadratic term of period with respectively coefficients θ and λ, and an interaction of individual general trust with period with coefficient ξ are included.

Controls. The macro-model on resource size controls for within-group levels of average general trust as measured by the IG, average age, average number of real-life acquaintances in the experimental session, average game theory experience, percentage of students, percentage

of women, and whether a session took place in the Netherlands (1) or not (0). The micro-model on appropriation effort controls for the group means as listed above, plus the individual measures of general IG trust, age, sex (male 0, female 1), number of acquaintances, game theory experience and being a student. All of the controls are present in the models, but only the ones that impact the dependent variable significantly will be reported in the tables.

In addition, the micro-model controls for resource size at period $t - 1$, to tease out how much of the individual behaviour is due to the actual treatment instead of the resource size in the previous period; one can expect a general tendency of individuals to appropriate more when the resource size is bigger, and less when the resource size is smaller. The macro-model on resource size will not control for this, for two reasons. Firstly, resource size is a direct and absolute measure of success at the macro level. Controlling for resource size in the previous period would change the interpretation of the dependent variable into periodic change in resource size, which is unintuitive as a measure of success. Secondly, as change in resource size is potentially affected by the treatment, controlling for resource size in the previous period could create endogeneity, which would complicate the interpretation of treatment effects from the model.

Instead of controlling for lagged resource size, period by treatment interactions are introduced up to a second order polynomial to account for the explicitly dynamic nature of the treatment effects. A critic may hold that such a model specification may be ad-hoc or otherwise mis-specified; the response can be found in S4 and S5 Figs in the supporting information, showing that non-parametric regression splines fitted to the purged residuals using Generalised Additive Models [89, 90] come to similar conclusions regarding the nature of the dynamic functional forms of the treatments.

While it is common in analyses of experimental data to control for endgame effects, this is not the case in the current analyses. Endgame effects are found when participants behave in a purely selfish and profit optimizing way near the end of the game, as they know that the interaction will end and thus there will be no consequences of defection in the future [91, 92]. In the current game, however, there is no clear endgame behaviour visible, as subjects did not know how many periods of the game they would have to play. As González, Güth and Levati [92] show, not informing participants when the experimental interaction ends does not alter behaviour in the game but does reduce the frequency of endgame effects. In addition, even though the experiment consists of many parts and experimental sessions took a significant amount of time, no sign of exhaustion by players is visible—that is, behaviour throughout the game seems constant with the exception of the first fifteen periods. An effect can be expected for these first few periods of the game: subjects may not know the game well enough to understand the consequences of their behaviour until after the first couple of periods. However, in the current game the initial drop of the resource size is not necessarily an artefact of participants misunderstanding the game but rather a behavioural pattern common in CPR games (see for instance [31] and [93]). Thus, as this is a process that is not an unnatural behavioural response to the game but a process that can be expected to be found in any CPR, the models will not control for startgame effects.

Results

Descriptive plots

Fig 3 shows an interesting difference in resource size over time between the treatments. The first thing that catches the eye is the steep drop of resource size in the first ten to fifteen periods. When the resource is at its fullest, all groups in all treatments seem to overappropriate the resource. This trend is very similar to the trend shown in the CPR game by Janssen, Holahan,
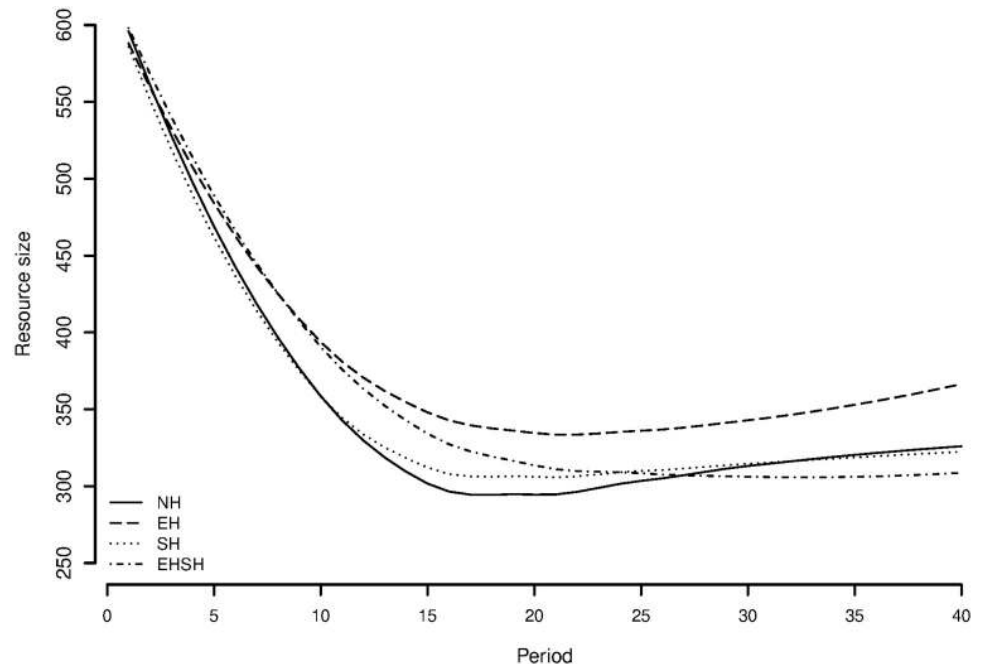
Fig 3. Mean resource size per treatment over time.

Lee and Ostrom [31], under treatments of no communication and punishment and costly punishment. A steep initial drop in resource size is also visible in the first 10 periods of the CPR fishing experiment by Hey, Neugebauer and Sadrieh [93], in a treatment where information resource growth and stock size are available.

Looking at differences between treatments, it stands out that the EH treatment seems to do better—that is, has a higher resource size and thus a higher profit per invested unit of appropriation—throughout the game than the other treatments, including the NH treatment. Strikingly, this only holds for EH under sociocultural homogeneity, and not for the EHSH treatment, in which the resource size keeps decreasing throughout the game and ends up being lower than the NH treatment. Given the curve visible in the graph, treatment effects may vary over time, suggesting an interaction with the quadratic term of time.

To get a more detailed idea of how groups performed in the various treatments, Fig 4 shows the resource size in each group per treatment. This graph shows that all treatments have well and poorly performing groups. However, it is clearly visible that the EH treatment has the highest concentration of groups that maintain a resource size above 400 throughout the game. The EHSH treatment has the highest concentration of groups that have a resource size below 300 throughout the game.

Fig 5 shows a behavioural pattern consistent with Fig 3; the EH treatment starts off with and keeps up a decreasing appropriation effort throughout the game. The EHSH treatment decreases too, although with a smaller slope, resulting in overexploitation and a lower resource size as is visible in Fig 3. In all treatments, the appropriation effort stabilizes around 30 units of *a*, which is the cooperative amount to invest in the resource per period per player.

Fig 6 shows boxplots of general, ingroup and outgroup trust as measured by the general IG, the ingroup IG and the outgroup IG, converted to a measure between 0 and 1. Each section of the box represents a quarter of the observations, and the middle line is the median. The boxplots show that while the trusting behaviour is similar in the general IG and the ingroup IG, it
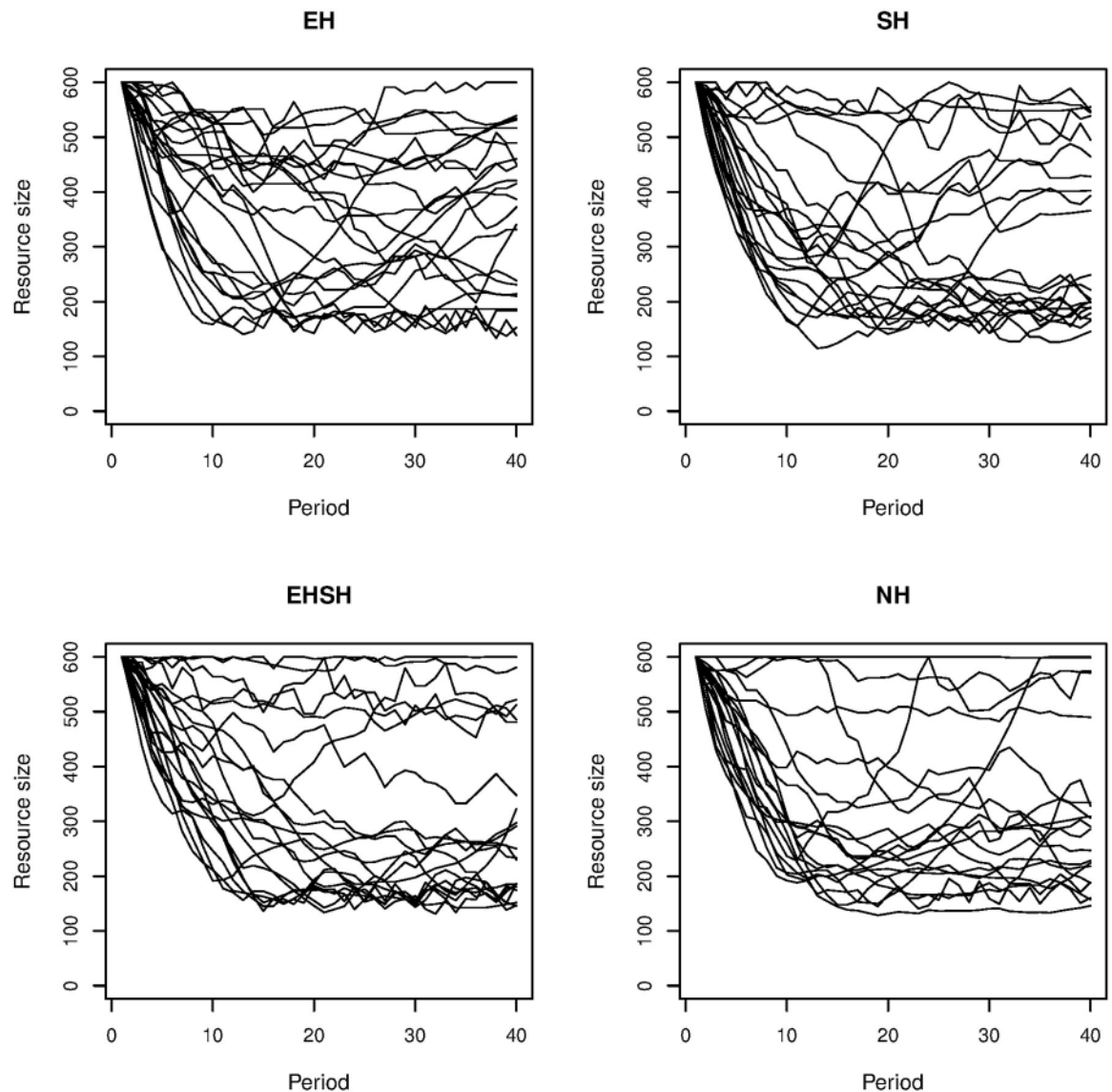
Fig 4. Resource size per group per treatment.

is generally lower for the outgroup IG. The latter illustrates at least partially the implications of the MGE; subjects treat outgroup members differently from ingroup members, even if group membership is based on an artificial criterion.

Similar to Fig 6, the boxplots in Fig 7 show a generally lower level of trustworthiness towards outgroup members than ingroup members, while general trustworthiness is similar to the ingroup IG trustworthiness levels.

Based on these descriptive plots, some interesting areas to investigate further include (a) the difference in effects between treatments over time; (b) a potential quadratic relation between treatment effects and time; and (c) the effect of general trust measured with the Investment Game on main-game behaviour of individuals in the different treatments. Results will be analysed separately for the macro- and the micro-level.
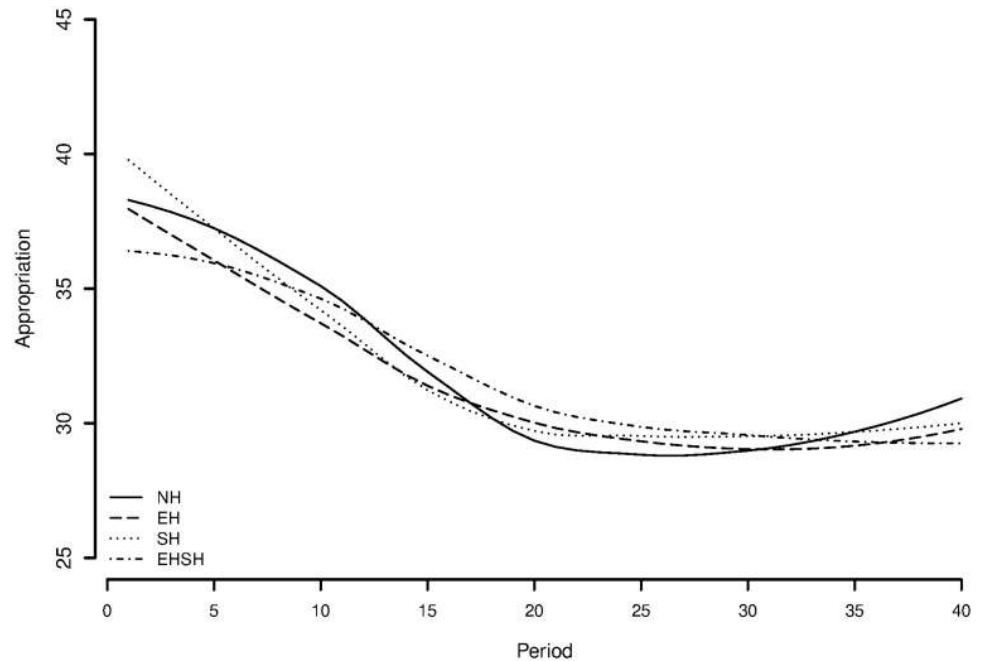
Fig 5. Mean appropriation effort per treatment over time.

https://doi.org/10.1371/journal.pone.0237870.g005

Non-parametric tests

A Kruskal-Wallis test [94] is employed to compare the means of resource size and appropriation effort between the four treatment groups. The test concludes that there is significant evidence that the group means are not equal (X^2 = 27.207; df = 3; p <0.001). Regarding appropriation effort, there is no evidence that the group means are not equal.

To analyse the differences in resource size between specific treatments, Mann-Whitney-Wilcoxon tests are performed for every pair of treatments. The distribution of resource size for EH is found to be significantly different from SH (W = 1233, p <0.001), EHSH (W = 1124, p < 0.001) and NH (W = 1237, p < 0.001). The other treatments do not differ significantly from each other according to the test. However, the trends of resource size and appropriation over time may still differ between treatments, as is apparent from the descriptive plots. Whether this is the case will be investigated using interactions in the multilevel regressions.

Resource size results

Table 3 shows the two-level multilevel regression on resource size. Model 3 is the model including treatments interacted with the linear and quadratic term of period, trust interacted with period, and all the group characteristic control variables.

The main effects of the treatments, to be interpreted as the treatment effect in period $t = 0$, are not significant and are in fact not very meaningful for interpretation on their own.

Model 3 shows a significant positive interaction effect of EH with time (B = 3.064; se = 1.204; p = 0.011), indicating that the slope of EH on resource size is significantly higher each period compared to NH. There is also a negative effect of EH interacted with the quadratic term of period (B = −0.063; se = 0.028; p = 0.027), indicating that the positive slope of EH on resource size by period will flatten out over the course of the game. This is a surprising but interesting result, contradicting hypothesis 1a, stating a negative effect of economic
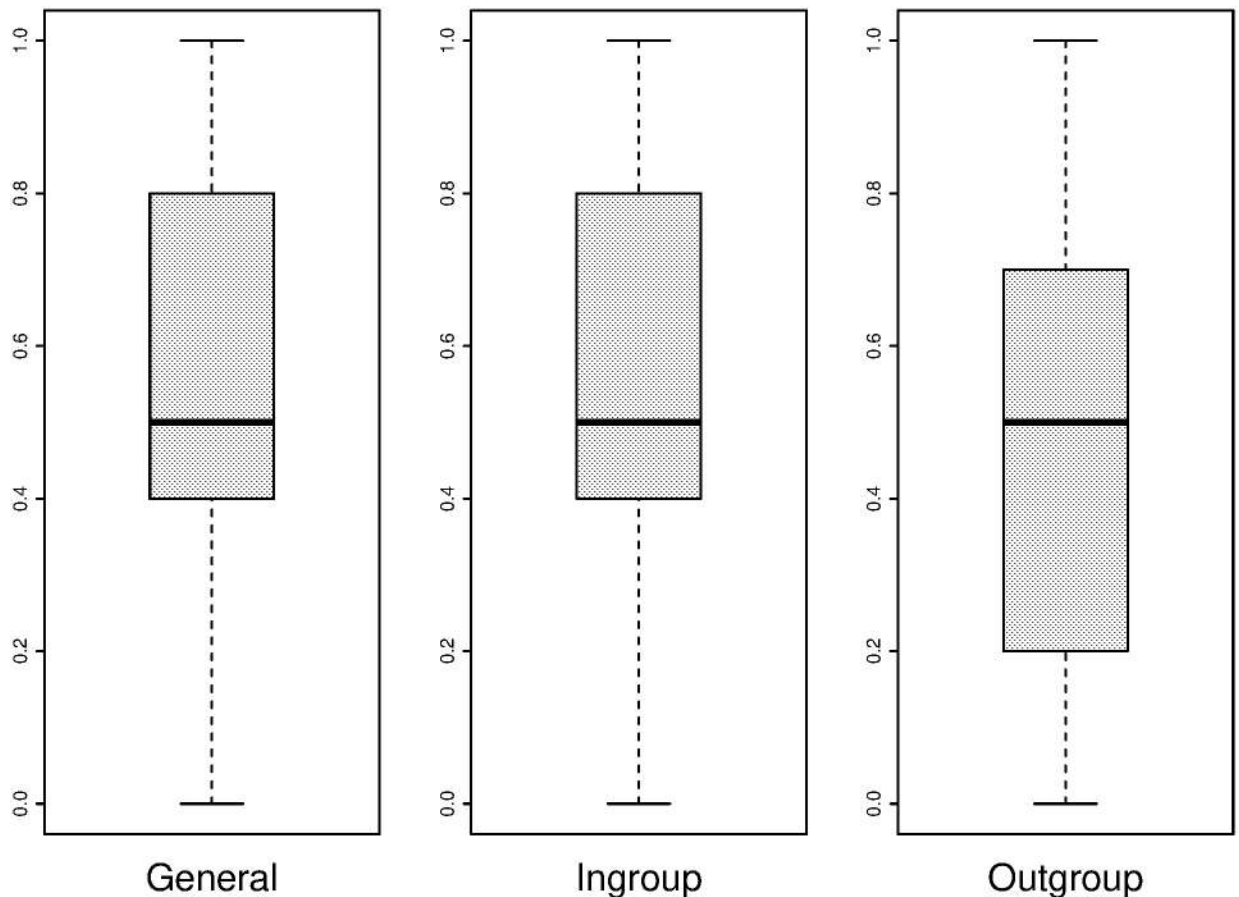
Fig 6. Trust of players in general (left), ingroup (centre) and outgroup (right) players as displayed in the IG.

heterogeneity on cooperation (and thus resource size on the macro-level) over time. Instead, taken together with the negative main effect of EH, results show a more complex relation of the treatment effect over time, starting with a negative effect becoming positive, and then flattening out.

There is a marginally significant positive difference in slope of SH on resource size relative to NH over time (B = 2.110; se = 1.217; p = 0.083) and a marginally significant negative interaction effect of SH with the quadratic term of period (B = −0.048; se = 0.029; p = 0.094). Taken together with the negative main effect of SH, this means that relative to NH, SH starts with a lower resource size, but has a larger slope than NH, which flattens out over time.

Model 3 shows a positive interaction of EHSH with period on resource size (B = 1.576; se = 1.231; p = 0.201), but this effect is not significant. What is significant, however, is the interaction of EHSH with the quadratic term of period (B = −0.069; se = 0.029; p = 0.0119). This indicates that the positive effect of EHSH on resource size per period relative to NH will decrease each period, until the slope becomes smaller than the slope of NH—and thus EHSH performs worse—around period 23. The main effect of period is now to be interpreted as the effect of NH over time. The model shows a significantly negative linear effect (B = −23.479; se = 0.954; p < 0.001), and a positive quadratic effect (B = 0.440; se = 0.021; p = < 0.001), indicating that the slope for NH is lower than the other treatments, but that this difference becomes smaller over time.
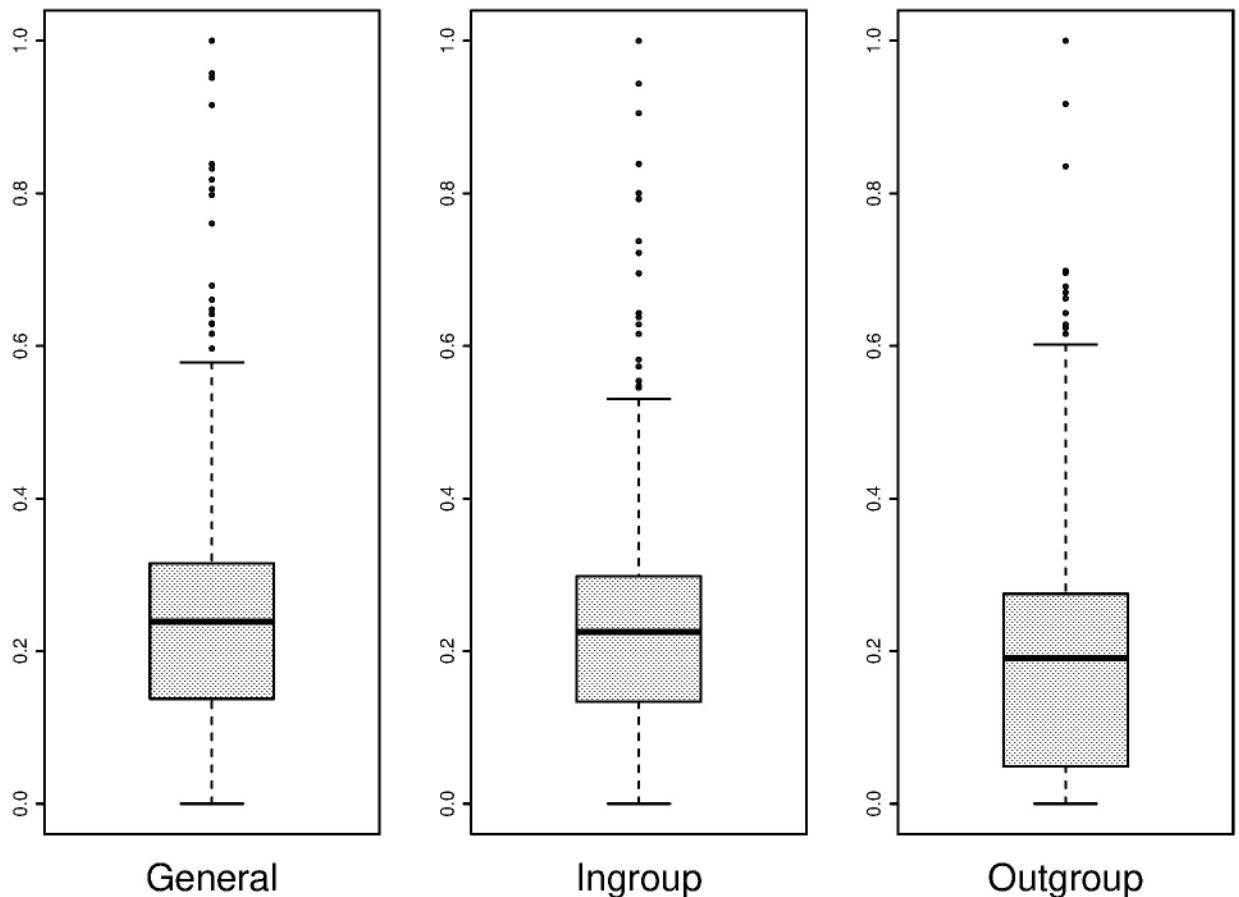
Fig 7. Trustworthiness of players towards general (left), ingroup (centre) and outgroup (right) players as displayed in the IG.

When changing the reference category in the multilevel model, it shows that the treatments do not differ significantly from each other. That is, EH and EHSH differ significantly from NH, but not from each other or from SH. When taking EHSH as a reference category, the interaction of EH with the linear term of period is not significant (B = 1.489, se = 1.189, p = 0.211). However, note that the p-value is low; this reflects a high probability that subjects in EH performed better in terms of resource size over time than the subjects in EHSH.

To visualise the discussed differences in slopes between treatments, Fig 8 shows a graph of treatment effects fitted on the purged residuals of a model with all control variables. This shows the predicted resource size over time per treatment, while controlling for all relevant control variables.

When put together, these results support hypothesis 1c on the negative effect of economic and sociocultural heterogeneity on cooperation over time. The hypothesis stated specifically a stronger negative effect of the combination of economic and sociocultural heterogeneity than either heterogeneity type separately. While technically true, an effect of economic heterogeneity in the opposite direction and a negligible effect of sociocultural heterogeneity were not anticipated. No evidence was thus found for hypotheses 1a and 1b.

The average level of trust in the group—as measured by taking the group mean of general trust displayed in the Investment Game at the beginning of the experiment—has a significant positive interaction with period (B = 1.811; se = 0.658; p = 0.006) meaning that higher average

Table 3. Two-level multilevel regression on resource size with random intercepts for groups.

	(1)	(2)	(3)
EH	21.131	16.152	−1.921
	(36.152)	(35.073)	(35.780)
EH × Period	0.497	0.483	3.064*
	(0.355)	(0.355)	(1.204)
EH × Period²			−0.063*
			(0.028)
SH	−1.645	−9.761	−23.581
	(36.532)	(35.365)	(36.081)
SH × Period	0.143	0.135	2.110+
	(0.359)	(0.359)	(1.217)
SH × Period²			−0.048+
			(0.029)
EHSH	35.099	21.953	2.274
	(36.944)	(36.227)	(36.943)
EHSH × Period	−1.236***	−1.236*	1.576
	(0.363)	(0.363)	(1.231)
EHSH × Period²			−0.069*
			(0.029)
Period	−4.437***	−5.445***	−23.479***
	(0.260)	(0.513)	(0.954)
Period²			0.440***
			(0.021)
General Trust [GT]		106.000	106.000
		(88.263)	(87.787)
GT × Period		1.811*	1.811**
		(0.795)	(0.658)
Controls			
Netherlands		88.930**	88.930**
		(30.551)	(30.551)
% Female		94.096+	94.096+
		(49.836)	(49.836)
Constant	439.853***	471.407**	597.644***
	(26.440)	(165.859)	(165.862)
Observations	3,440	3,440	3,440
Groups	86	86	86
Log Likelihood	−20,323.720	−20,283.180	−19,659.490
Akaike Inf. Crit.	40,667.450	40,602.360	39,362.980
Bayesian Inf. Crit.	40,728.860	40,712.850	39,498.000

Standard errors in parentheses.

+p<0.1;

*p<0.05;

**p<0.01;

***p<0.001

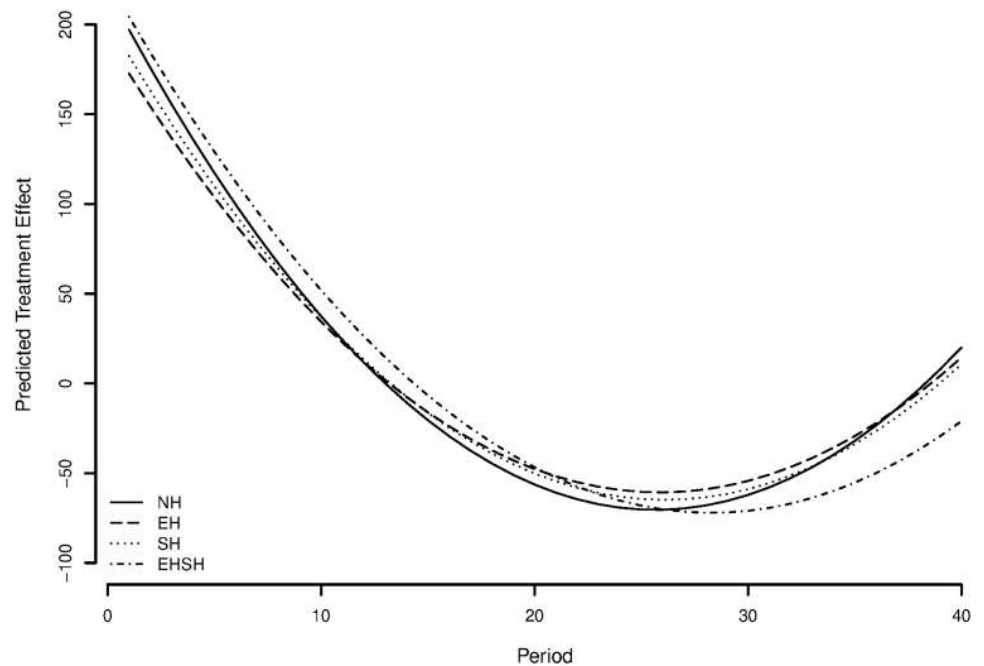Tables produced with *Stargazer* [95]

Fig 8. Predicted treatment effects on resource size with multilevel regression coefficients.

trust in the group will yield higher resource size each period. This supports hypothesis 3 on the positive effect of trust on cooperation over time. Controlling for the maximum individual level of general trust in the group, the interaction between average level of trust and period is still significant, indicating that it is not just one high trust player in the group that facilitates good outcomes for the group, but that more players with higher trust will lead to better results for the group.

As for the significant control variables, the models show that the subjects in the Netherlands managed to keep up higher levels of the resource size (B = 88.930; se = 30.551; p = 0.005), which is an interesting finding indicating that even though the subject pools from Oxford and Utrecht contain international students and residents of approximately the same age range, the location of the experimental sessions (or perhaps the country of residence of subjects) matters. A marginally significant and positive effect is found for a higher percentage of females in the group (B = 94.096; se = 49.096; p = 0.063).

Appropriation effort results

Table 4 shows the three-level multilevel regression on the micro-level variable appropriation effort. None of the models show a significant main effect of the treatments, but as these effects are interpreted as treatment effects at period $t = 0$, these coefficients are not very meaningful.

Model 3 is the complete model including treatments interacted with the linear and quadratic term of period, trust interacted with period, and control variables. The model shows a significant positive effect of EHSH on appropriation per period relative to NH (B = 0.290; se = 0.114; p = 0.011) but a negative effect when interacted with the quadratic term of period (B = −0.006; se = 0.003; p = 0.020). Taken together with the negative main effect (B = −2.517; se = 1.668), this suggests that the difference in slopes between EHSH and NH on appropriation increases and finally flattens out. Subjects in EHSH thus appropriate more for a large part of the game, until appropriation becomes more similar to NH.

Table 4. Three-level multilevel regression on appropriation effort with random intercepts for subjects and groups.

	(1)	(2)	(3)
EH	−0.910	−1.222	−2.103
	(1.287)	(1.430)	(1.620)
EH × Period	0.023	0.017	0.148
	(0.025)	(0.026)	(0.112)
EH × Period^2			−0.003
			(0.003)
SH	−0.086	0.141	0.368
	(1.300)	(1.450)	(1.644)
SH × Period	0.004	0.003	−0.033
	(0.025)	(0.027)	(0.114)
SH × Period^2			0.001
			(0.003)
EHSH	−0.548	−0.700	−2.517
	(1.315)	(1.475)	(1.668)
EHSH × Period	0.024	0.036	0.290*
	(0.026)	(0.027)	(0.114)
EHSH × Period^2			−0.006*
			(0.003)
Period	−0.242***	−0.095***	−0.516***
	(0.018)	(0.026)	(0.090)
Period^2			0.009***
			(0.002)
General Trust [GT]		−1.689	−1.728
		(1.846)	(1.807)
GT × Period		−0.095**	−0.092**
		(0.030)	(0.030)
Controls			
Resource Size t-1		0.017***	0.012***
		(0.001)	(0.001)
Sum Appropriation Others t-1		0.042***	0.036***
		(0.004)	(0.004)
Netherlands		−2.085+	−1.625
		(1.179)	(1.153)
Female		−2.006+	−1.992+
		(1.132)	(1.106)
Constant	36.914***	24.938***	30.795**
	(0.941)	(6.503)	(6.421)
Observations	13,760	13,299	13,299
Subjects	344	341	341
Groups	86	86	86
Log Likelihood	−54,226.350	−52,376.450	−52,370.950
Akaike Inf. Crit.	108,474.700	104,804.900	104,801.900

(*Continued*)

Table 4. (*Continued*)

	(1)	(2)	(3)
Bayesian Inf. Crit.	108,557.500	104,999.700	105,026.700

Standard errors in parentheses.

+p<0.1;

*p<0.05;

**p<0.01;

***p<0.001

Tables produced with *Stargazer* [95]

https://doi.org/10.1371/journal.pone.0237870.t004

No significant effects of the other treatments are visible, but it is worth noting that for EH the p-values of the main effect (p = 0.198), interaction with linear period (p = 0.186) and interaction with the quadratic term of period (p = 0.238) are close to 0.2, which reflects a non-significant but relatively high probability that subjects in EH treatment behaved more cooperatively over time than subjects in NH, as was found in the macro-model. The main effect of period is now to be interpreted as the effect of NH over time. The model shows a negative slope over time (B = −0.516; se = 0.09; p < 0.001) which increases each period (B = 0.009; se = 0.002; p < 0.001).

To give a visualisation of the differences in slopes between treatments, Fig 9 shows a graph of treatment effects fitted on the residuals of a model with all control variables. This shows the predicted resource size over time per treatment, while controlling for all relevant control variables. The results provide modest evidence in the opposite direction of hypothesis 1a, suggesting for the individual level a positive instead of a negative effect of economic heterogeneity under sociocultural homogeneity on cooperation over time. A surprising but interesting result that will be reflected on later.
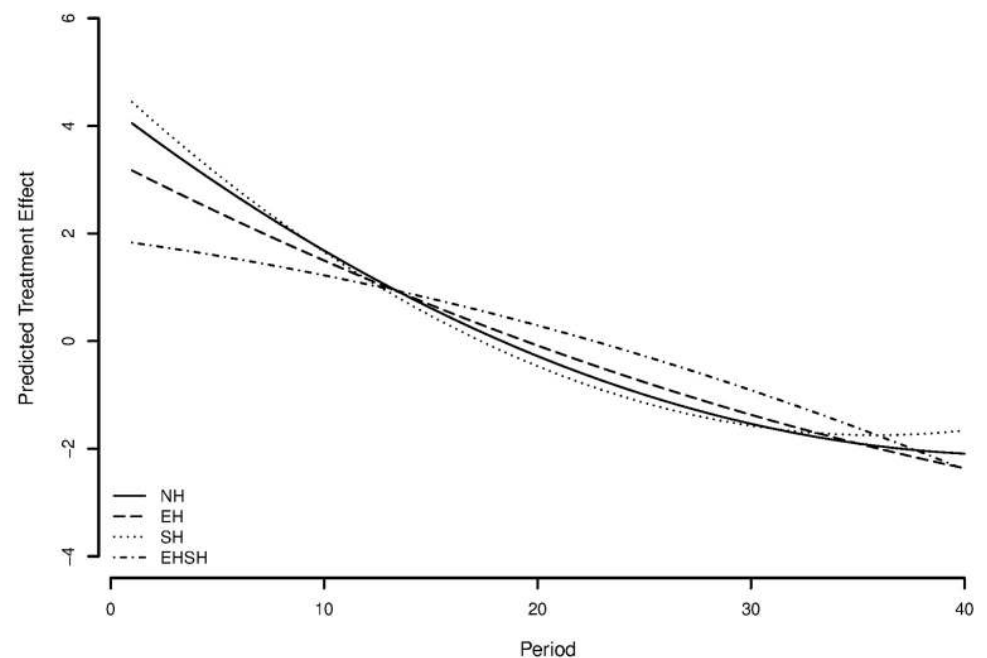

Fig 9. Predicted treatment effects on appropriation effort using multilevel regression coefficients.

https://doi.org/10.1371/journal.pone.0237870.g009

It is worth noting the possibility that one would need more statistical power to detect significant differences on the micro-level. However, even small changes in behaviour on the micro-level can lead to big effects at the macro-level. The lack of significant findings on the micro-level thus does not mean that the effects do not play a role, in particular because the differences in behaviour between treatments found on the micro-level in Table 4 do support the differences that were found in Table 3 on the macro-level.

Regarding trust from the IG, model 3 shows a significant negative interaction effect of trust with time on appropriation effort (B = −0.092; se = 0.03; p = 0.002). This suggests that higher levels of general trust result in lower levels of individual appropriation every period and thus higher levels of individual cooperation over time—after all, a lower appropriation effort yields higher resource size and higher returns for every player in the group. This supports hypothesis 3, stating the positive effect of trust on cooperation over time.

Regarding significant control variables, resource size in $t - 1$ (B = 0.012; se = 0.001; p < 0.001), and the sum of appropriation of other players in $t - 1$ (B = 0.036; se = 0.004; p < 0.001) have positive effects on appropriation effort. Female subjects appropriate marginally less (B = −1.992; se = 1.106; p = 0.073). Model 2 shows that Dutch subjects appropriate marginally less (B = −2.006; se = 1.179; p = 0.081), but this effect disappears in model 3.

Post-experimental trust results

First of all, Fig 10 shows a descriptive plot of trust as measured in the post-experimental survey with the statement "I trusted the other players in my group", referring to the Fishing Game,
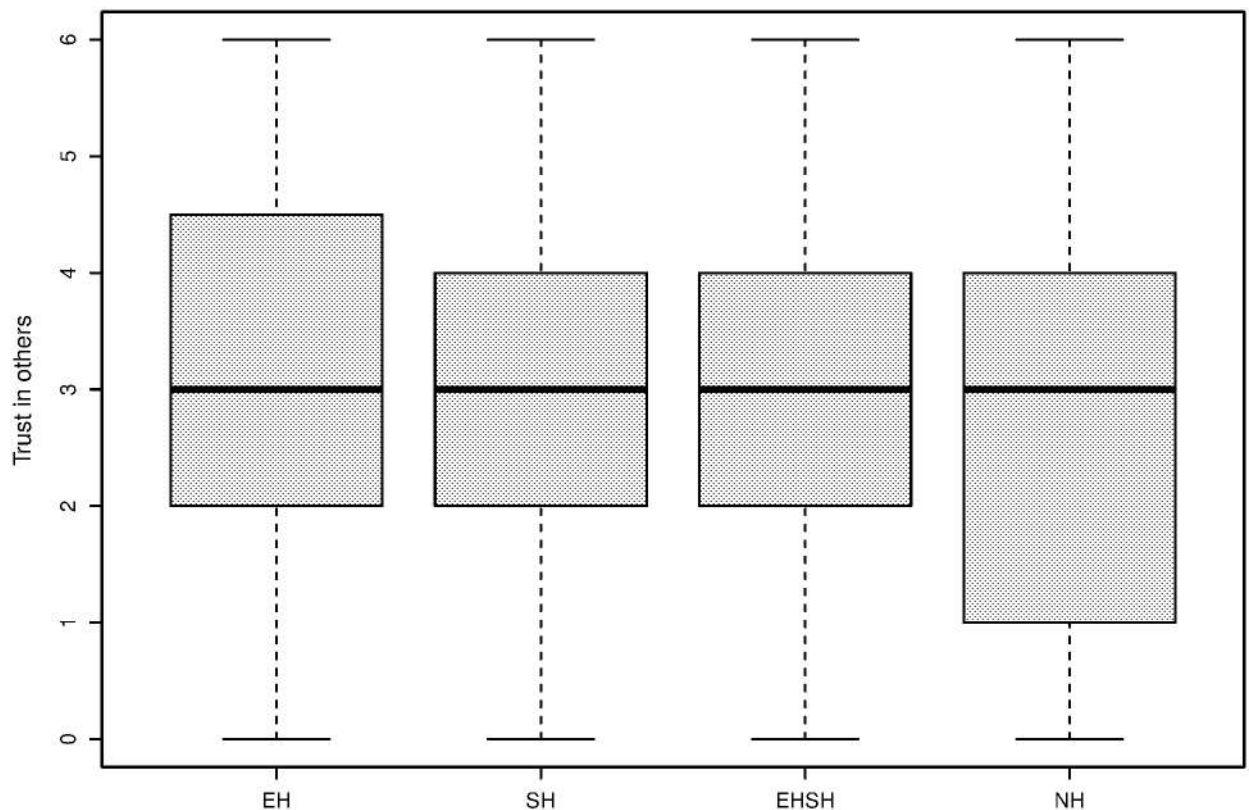

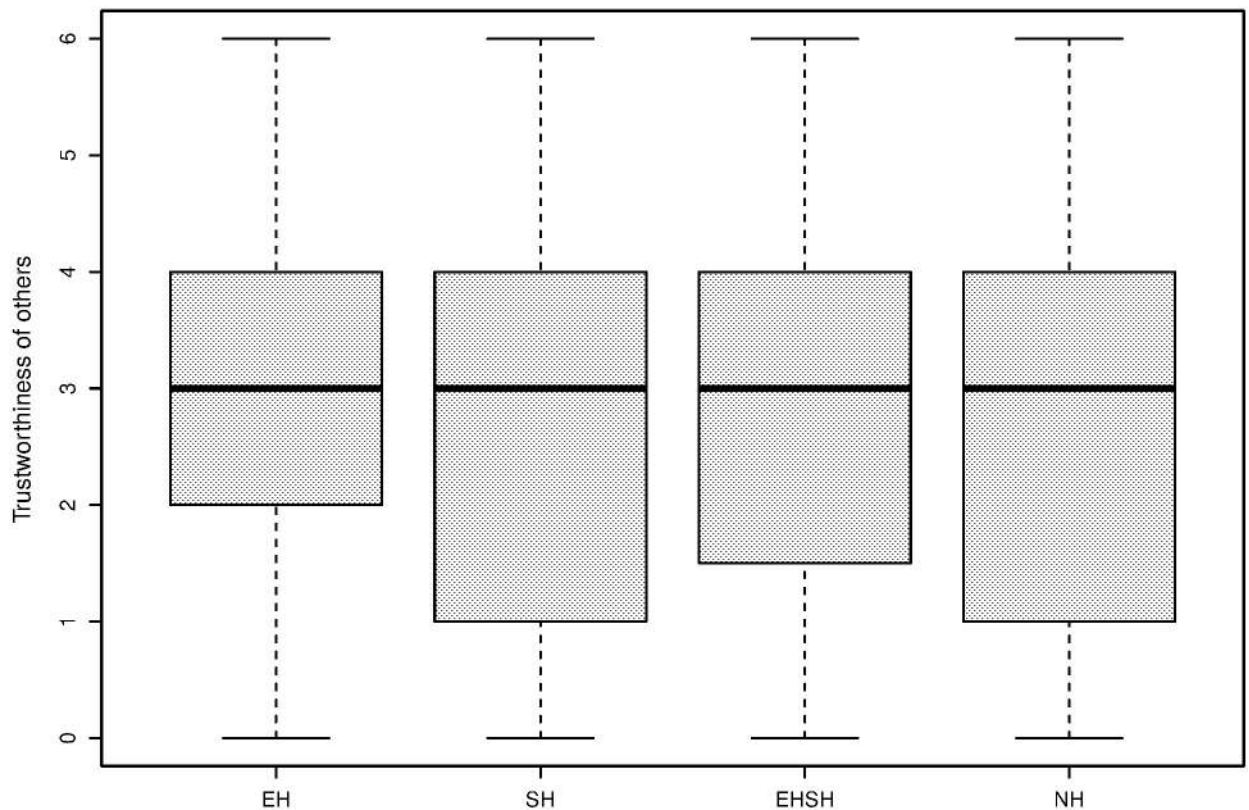**Fig 10. Trust in other players by treatment.**

https://doi.org/10.1371/journal.pone.0237870.g010

Fig 11. Subjective trustworthiness other players by treatment.

https://doi.org/10.1371/journal.pone.0237870.g011

hereafter called trust in other players. The variable is measured on a 7-point Likert scale rang-
ing from 'completely disagree' (0) to 'completely agree' (6). It shows that in especially the eco-
nomic heterogeneity group, the trust question is answered more positively by a higher
percentage of subjects, and for the homogeneity treatment this is lower. Fig 11 shows a
descriptive plot of a statement from the post-experimental survey stating "The other players in
my group were trustworthy", hereafter called subjective trustworthiness of other players, and
measured on the same 7-point Likert scale. It shows slightly lower scores for the sociocultural
heterogeneity treatment and the homogeneity treatment.

An ordinal logistic regression on the two mentioned trust questions from the post-experi-
mental survey can be found in S5 Text. No statistical evidence is found for hypotheses 2a, 2b
or 2c on the negative effect of heterogeneity on trust. Instead it seems that the final individual
profit at the end of the experiment has a strong positive influence on how the survey questions
on trust were answered.

Revisiting expectations

Table 3 shows that economic heterogeneity affects collective action and resource size in a posi-
tive way compared to homogeneity. This is in stark contrast with hypothesis 1a, stating a nega-
tive effect of economic heterogeneity. As not only a non-effect but a significant effect in the
opposite direction in the macro-model, it is worth revisiting the literature for possible
explanations.

The Olson-effect.   Despite the majority of research suggesting a negative effect of hetero-
geneity on cooperation, the positive effect of heterogeneity is theorized by the economist

Mancur Olson in his book *The logic of collective action: public goods and the theory of groups.* [9] describing what is known as the "Olson-effect":

> "In smaller groups marked by considerable degrees of inequality—that is, in groups of members of unequal "size" or extent of interest in the collective good—there is the greatest likelihood that a collective good will be provided; for the greater the interest in the collective good of any single member, the greater the likelihood that that member will get such a significant proportion of the total benefit from the collective good that he will gain from seeing that the good is provided, even if he has to pay all of the cost himself" (p. 34).

Even though Olson does not directly mention a positive effect of economic heterogeneity on cooperation, he does describe a theoretical mechanism of the rich bearing the costs of cooperation for the poor by overinvesting in cooperation. In the context of the current CPR experiment, the group was small enough so that subjects with higher endowments may have invested less than they could have, to provide space for the two other subjects with lower endowments to invest in the resource for profit. The cost for not investing in the resource, and thus not receiving profit from appropriation, is lower for the higher endowed subjects, as everyone can keep the endowment that was not invested, which is higher for them to begin with.

To explore the possibilities of the Olson-effect in the current experiment, Fig 12 shows a plot of the appropriating behaviour of the high and low endowed subjects in the EH and EHSH treatments. It shows that in the EH treatment, the investments in appropriation for the higher endowed subjects are lower than the investments from higher endowed subjects in the EHSH treatment. For both treatments, the lower endowed subjects have on average about the same appropriation over time, with the ones in EH higher from the 16th period on.

Fig 13 shows that the lower endowed subjects in the EH treatment can make more profit than the lower endowed subjects in the EHSH treatment, from about period 10, while higher
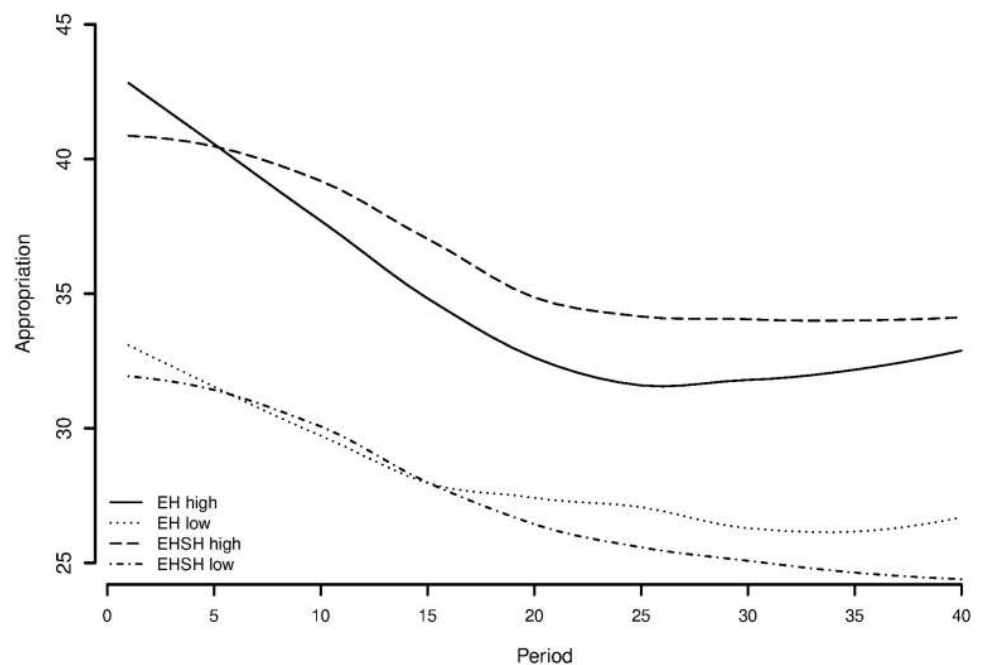

Fig 12. Appropriation of higher (*E* = 60) and lower (*E* = 40) endowed players in economic heterogeneity [EH] and economic and sociocultural heterogeneity [EHSH] treatment (smoothed line).

https://doi.org/10.1371/journal.pone.0237870.g012

Fig 13. Profit of higher (*E* = 60) and lower (*E* = 40) endowed players in economic heterogeneity [EH] and economic and sociocultural heterogeneity [EHSH] treatment (smoothed line).
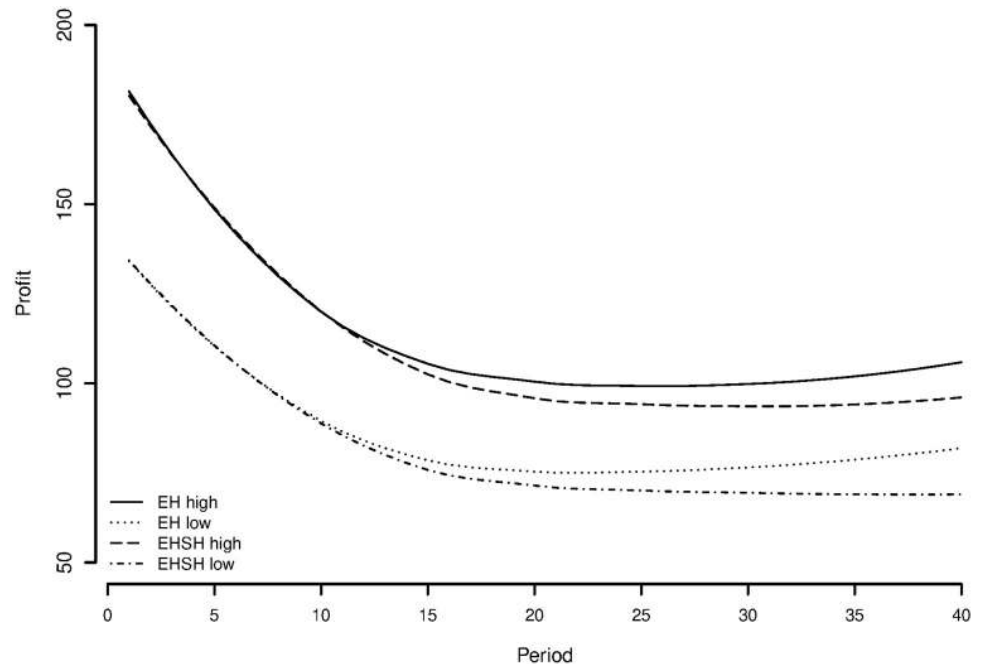
endowed subjects in EH still make more profit than higher endowed subjects in EHSH. It is still the case that the higher endowed in both treatments overappropriate on average, but an unpaired t-test points out that the higher endowed appropriate less in the EH (M = 34.830, SD = 17.116) than in the EHSH (M = 36.390, SD = 17.922) treatment (t (3518) = −2.642, p = 0.004). Taken together these results provide evidence for the Olson-effect in EH, but not in EHSH.

Ingroup favouritism and conditional reciprocity. A closer look can be taken at the effect of the MGE on in-game behaviour. Even though trust itself may not play a role in levels of cooperation, the difference in outcomes of the IG with ingroup and outgroup members could indicate why less cooperative behaviour takes place when economic heterogeneity is combined with sociocultural heterogeneity. Consistent with the descriptive plots of Figs 6 and 7, a paired samples t-test comparing the average trust in the general IG (i.e. playing with a random other player before the MGE) with average trust in the outgroup IG (i.e. playing with an outgroup member after the MGE) shows that there is a significant difference in average trust between a general (M = 0.559, SD = 0.310) and an outgroup (M = 0.455, SD = 0.333) interaction (t (343) = −8.023, p < 0.001). The average trust in the ingroup (M = 0.573. SD = 0.320) and outgroup (M = 0.455, SD = 0.333) IG interaction is also significant (t (343) = 10.010, p < 0.001). These results show that even with a division that is as artificial as painting preferences of painters in the same art discipline, group identities are strong enough to behave differently in different group compositions of ingroup and outgroup members. In their meta-study, Balliet, Wu and De Dreu [96] support the finding that people are more cooperative with ingroup, compared to outgroup, members.

The current results regarding trust indicate derogation towards the outgroup rather than favouritism towards the ingroup, since the difference in behaviour towards a general other and an ingroup member is not significant. In a mixed-design experiment using Dictator Games,

Bilancini et al. [83] show that ingroup favouritism is stronger when group are based on moral preferences rather than non-moral preferences. This finding provides another reason for future research to vary the operationalisation of sociocultural heterogeneity.

Another potential explanation for cooperative behaviour in EH is conditional reciprocity. Numerous studies, both experimental and in the field, have shown that people are conditional cooperators: they cooperate when others cooperate as well (see amongst others [97–102]). In the IG results of the current study, conditional reciprocity is visible: controlling for individual characteristics and location of the experiment, player 2 in the general IG sends a significantly higher percentage of their total points back to player 1, the higher the percentage of the points sent by player 1 of their endowment (B = 0.570, se = 0.091, p = < 0.001). Similarly, the significant positive effect of lagged sum of appropriation of others in Table 4 means that a lower appropriation of other players results in a lower appropriation of the player, which can be thought of as a form of reciprocal cooperation. This behaviour can also be related to informational effects; a phenomenon that describes how individuals will better comply with norms or rules if they see a greater number of other individuals complying as well [103].

However, none of the behavioural theories above explain the difference between EH and EHSH behaviour on their own. The answer may lie in a cross between ingroup favouritism, conditional reciprocity and the Olson-effect.

Discussion

The aim of this paper is to study the impact of economic and sociocultural heterogeneity and the coincidence of the two, through trust on cooperation on the micro- and macro-level in common-pool resources. Using a CPR game in the laboratory allowed for the effects of economic and sociocultural heterogeneity and trust to be disentangled and enabled establishing the causal direction of effects. Existing literature predominantly suggests negative effects of heterogeneity on trust and on cooperation, and positive effects of trust on societal outcomes.

The results show that under the coincidence of economic and sociocultural heterogeneity, groups struggle to converge to a sustainable appropriation of the common-pool resource over time. Surprisingly, the economic heterogeneity treatment is the first to converge to cooperative levels of appropriation. It manages to hold the highest resource size over the 40 periods of the game and is the most successful of all four treatment groups, including the homogeneous control group, over time. A striking conclusion here is thus that it is the presence or absence of sociocultural heterogeneity in common-pool resource settings that makes economic heterogeneity perform respectively worse or better than full homogeneity. The results contribute to the current literature by providing a possible explanation for the emergence of aversion to economic inequality, which may not have a natural origin per se, but could have a sociocultural origin instead. Sociocultural heterogeneity under economic homogeneity, however, differs only marginally from full homogeneity on the macro-level. Regarding trust, it is found that over time, higher average general trust within a group increases resource size, and individual general trust decreases appropriation effort—trust thus increases cooperation on both the micro- and macro-level. In addition, the post-MGE Investment Game results showed that subjects acted more trusting towards ingroup members than towards outgroup members, and less trusting towards outgroup members than to a random other person in the pre-MGE Investment Game. However, no evidence for the mediating role of trust between heterogeneity and cooperation in CPRs was found.

The results support literature suggesting positive effects of economic heterogeneity [9] or U-shaped relations between CPR performance and economic inequality (see [47, 104]). Olson [9] suggests that economic heterogeneity may have a positive effect on cooperation, provided

that the rich act as catalysts for cooperation by bearing the cost of collective action just a little more than the poor. It is shown that subjects with high endowments in the economic heterogeneity treatment invest less on average than subjects with a high endowment in the combination treatment. These results are not unthinkable if one places them in the context of common-pool resources: fishermen who come from the same sociocultural background, speaking the same language and with a similar view on the resource, could be more likely to cooperate with each other, despite economic heterogeneity, than with an unfamiliar actor—such as a large fishing company with a big fleet—that is unlike them on both sociocultural background and economic means. The key here could be a conflict of interests arising at the moment that two groups are unequal in economic means to appropriate a resource without a sociocultural bond to bridge that inequality gap. However, as Bazzi et al. [3] argue, the emergence of coordination depends on whether heterogeneity is shaped by many smaller groups, or a few bigger groups; smaller groups are more prone to find common ground, whereas bigger groups are more prone to ingroup antagonism. The results underline the importance of understanding the influence of various types of heterogeneity, and their interaction, on common-pool resource outcomes.

Some critical comments can be made about this study. A well-known criticism of laboratory experiments using mainly students as their subjects is that they are not representative of situations in the real world, while the results are sometimes presented as real-world outcomes. However, if the aim of the research is to investigate relationships between human behaviour and social, biological or economic contextual variables, experiments are a good way of doing so, regardless of the subject pool [36, 105–108]. To point out causality and to show the effect of a treatment the only assumption necessary is appropriate randomisation, which the laboratory setup provides [109]. In addition, a survey-experiment conducted with a representative sample of a city's population by Exadaktylos, Espín and Brañas-Garza [110] points out that students are indeed appropriate subjects to study human behaviour with laboratory experiments. The external validity of experiments can be secured as long as the environment under which the results are generated capture essential characteristics of the real-world version of the phenomenon that is being researched [111]. In the current paper, a CPR experiment was conducted which contained key aspects of the way CPRs, and in particular fisheries, work. A first step improvement on this study could be made by using real forms of identity, as real-life heterogeneous communities using CPRs also have to deal with real-life sentiments towards real ingroup and outgroup members. For instance, gender or nationality of subjects could be used to operationalise sociocultural heterogeneity. This could also serve as an extra test for and comparison with using the MGE to investigate whether subjects behave differently in CPR games if they are grouped by a natural rather than a more artificial characteristic. Lastly, future research could investigate whether the group size of different preference or identity groups influences the outcome of the game, as Bazzi et al. [3] suggest.

For future CPR experiments it would also be interesting to add communication and/or punishment between players (see for instance Janssen et al. [31]) or to add a type of risk to overexploitation that is separate from decreasing income (see for instance Bednarik, Linnerooth-Bayer, Magnuszewski and Dieckmann [112] where overharvesting trees increases the chance of flood damage). Adding elements like this can change behaviour drastically, and can create a more realistic setting. A bolder improvement on this study could be made by setting up an artefactual field experiment, better known as a lab-in-field experiment: a controlled environment where artefactual games (such as the Investment Game or the Fishing Game) are played, but with a subject pool that is more like the population of interest. In the current context, this could be a group of actual fishermen.

With the use of detailed and heavily contextualised games, comes a more detailed—and thus limited—interpretation of the results; the results of this study are valuable for common-pool resources, and especially resources with structures similar to fishing grounds. With the information about the resource size (fish stock) available for players to see at the beginning of every new period, the results may not hold in situations where the total allowable catch cannot be correctly measured because the fish population dynamics is unknown [93]. In addition, whereas players could see the history of other players' actions, this may not always be the case (see for instance Lacomba, Lagos and Perote [113]). Future experiments could vary the amount and accuracy of information about resource growth and stock size, or information on other players' actions, as it is found to influence players' appropriation behaviour [93, 113, 114]. Findings may be different for other resource types, such as irrigation systems, where resource renewal is partially dependent on the weather, appropriation is sequential, and farmers already know if someone took more water than allowed before they make their own decision of how much water to take.

The level of detail in the game is necessary to fully understand the different mechanisms at work in different CPRs; games that are too generalised will not provide directed results. However, hypotheses were not formulated based on resource-specific characteristics, so it may well be that the found results would apply to a wider range of CPRs. The flexibility for researchers to adjust a game to represent any type of CPR is a very valuable asset of experimental research, and one that should be deployed more often in social science.

Conclusion

As far as one should base policy advice on laboratory results alone, the main recommendation is perhaps that management of CPRs cannot be a 'one size fits all' solution, but instead that sustainable cooperation in CPRs can be achieved under a flexible management that is adapted to the level and the type(s) of heterogeneity found in a CPR user community. Laboratory experiments are important to tease out the mechanisms at work in social phenomenons. The mechanisms for which this paper provide evidence should be further studied in real-life contexts, first through lab-in-field experiments, and then via large-scale randomised control trials, to ensure they are externally valid instruments of policy.

Understanding cooperative behaviour and trust under different conditions of heterogeneity is a core question within social sciences. The application of this question to common-pool resource situations is vital especially in a time of increasing depletion of these resources, manifested in overfishing, deforestation and unsustainable use of fresh water. It is crucial to understand the behaviour of humans in these contexts, to prevent overexploitation of resources and to promote cooperation between different actors involved. In addition, the growing number of modern commons such as citizen initiatives for food, green energy and infrastructure provide new incentives to investigate human behaviour in and around CPRs [30, 42, 43]. The investigation of relations between heterogeneity, trust and cooperation in common-pool resources is not only important to advance insights within the social sciences, but also demonstrates the importance of social science research for present-day problems.

Supporting information

S1 Text. General instructions. The general experiment instructions that subjects receive upon entering the lab.
(PDF)

S2 Text. Specific instructions. Treatment-specific instructions for subjects in the different treatments. Treatments are indicated on the top left of the page with an 'E' for economic heterogeneity, 'S' for sociocultural heterogeneity, 'B' for economic and sociocultural heterogeneity ('Both') and 'N' for no heterogeneity (the homogeneous control group).
(PDF)

S3 Text. The Investment Game. A description of all the characteristics of the Investment Game.
(PDF)

S4 Text. Table: Other-other Dictator Game details. A table that shows the two options that subjects faced in each of the three scenarios in the other-other DG.
(PDF)

S5 Text. Post-experimental trust models. Ordinal logistic regression on two survey questions on trust in and trustworthiness of other players.
(PDF)

S1 Fig. The Investment Game. A graphic representation of the Investment Game as played in the current experiment.
(TIF)

S2 Fig. Figure: Resource size under different group behaviour. A graph that shows how resource size develops over time under different appropriation behaviour: from a summed appropriation of 120 to 200 in increments of 10.
(TIFF)

S3 Fig. Figure: Cumulative profit under different group behaviour. A graph that shows how cumulative profit of groups develops over time under different appropriation behaviour: from a summed appropriation of 120 to 200 in increments of 10.
(TIFF)

S4 Fig. GAM with splines for resource size. A graph to compare predicted treatment effects between multilevel [ML] models and generalised additive model [GAM] models, to check whether the interaction effects of the treatments with the linear and quadratic term of period as specified in the multilevel models fit the natural trend of the data. The time variant variables (period and the interactions of treatments with periods) are specified as smooth terms, which are based on low rank version of splines.
(TIFF)

S5 Fig. GAM with splines for appropriation effort. A graph to compare predicted treatment effects between multilevel [ML] models and generalised additive model [GAM] models, to check whether the interaction effects of the treatments with the linear and quadratic term of period as specified in the multilevel models fit the natural trend of the data. The time variant variables (period and the interactions of treatments with periods) are specified as smooth terms, which are based on low rank version of splines.
(TIFF)

Acknowledgments

University together with Dr. Rense Corten. I would also like to thank Dr. Ozan Askoy for providing useful feedback, and Roberto Cerina and other Nuffield College colleagues for discussing my research with me on multiple accounts. Lastly, I would like to thank Prof. Nan Dirk de Graaf for his support.

Author Contributions

Conceptualization: Fijnanda van Klingeren.

Data curation: Fijnanda van Klingeren.

Formal analysis: Fijnanda van Klingeren.

Funding acquisition: Fijnanda van Klingeren.

Investigation: Fijnanda van Klingeren.

Methodology: Fijnanda van Klingeren.

Writing – original draft: Fijnanda van Klingeren.

Writing – review & editing: Fijnanda van Klingeren.

References

1. International Migration report 2019. Department of Economic and Social Affairs, United Nations; 2019.

2. Seabright P. Managing Local Commons: Theoretical Issues in Incentive Design. Journal of Economic Perspectives. 1993 Dec; 7(4):113–134. Available from: https://www.aeaweb.org/articles?id=10.1257/jep.7.4.113.

3. Bazzi S, Gaduh A, Rothenberg AD, Wong M. Unity in Diversity? How Intergroup Contact Can Foster Nation Building. American Economic Review. 2019 Nov; 109(11):3978–4025. Available from: https://www.aeaweb.org/articles?id=10.1257%2Faer.20180174&&from=f.

4. Anderson LR, Mellor JM, Milyo J. Induced Heterogeneity in Trust Experiments. Experimental Economics. 2006; 9(3):223–235. https://doi.org/10.1007/s10683-006-9124-2

5. Andersson K, Agrawal A. Inequalities, Institutions, and Forest Commons. Global Environmental Change. 2011; 21(3):866–875. https://doi.org/10.1016/j.gloenvcha.2011.03.004

6. Bardhan P, Dayton-Johnson J. Unequal Irrigators: Heterogeneity and Commons Management in Large-Scale Multi-Variate Research. In: Ostrom E, Dietz T, Dolsak N, Stern PC, Stonich S, Weber EU, editors. The Drama of the Commons. Washington, D. C.: National Academy Press; 2002. p. 87–112.

7. Flache A, Mäs M. Why Do Faultlines Matter? A Computational Model of How Strong Demographic Faultlines Undermine Team Cohesion. Simulation Modelling Practice and Theory. 2008; 16(2):175–191. https://doi.org/10.1016/j.simpat.2007.11.020

8. Kölle F. Heterogeneity and Cooperation: The Role of Capability and Valuation on Public Goods Provision. Journal of Economic Behavior & Organization. 2015; 109:120–134. https://doi.org/10.1016/j.jebo.2014.11.009

9. Olson M. The Logic of Collective Action: Pubic Goods and the Theory of Groups. Cambridge, MA: Harvard University Press; 1965.

10. Poteete AR, Ostrom E. Heterogeneity, Group Size and Collective Action: The Role of Institutions in Forest Management. Development and change. 2004; 35(3):435–461. https://doi.org/10.1111/j.1467-7660.2004.00360.x

11. Ruttan LM. Economic Heterogeneity and the Commons: Effects on Collective Action and Collective Goods Provisioning. World Development. 2008; 36(5):969–985. https://doi.org/10.1016/j.worlddev.2007.05.005

12. Vedeld T. Village Politics: Heterogeneity, Leadership and Collective Action. The Journal of Development Studies. 2000; 36(5):105–134. https://doi.org/10.1080/00220380008422648

13. Cherry TL, Kroll S, Shogren JF. The Impact of Endowment Heterogeneity and Origin on Public Good Contributions: Evidence from the Lab. Journal of Economic Behavior & Organization. 2005; 57 (3):357–365. https://doi.org/10.1016/j.jebo.2003.11.010

14. Alesina A, La Ferrara E. Who Trusts Others? Journal of Public Economics. 2002; 85(2):207–234. https://doi.org/10.1016/S0047-2727(01)00084-6

15. Anderson CJ, Paskeviciute A. How Ethnic and Linguistic Heterogeneity Influence the Prospects for Civil Society: A Comparative Study of Citizenship Behavior. The Journal of Politics. 2006; 68(4):783–802. https://doi.org/10.1111/j.1468-2508.2006.00470.x

16. Delhey J, Newton K. Predicting Cross-National Levels of Social Trust: Global Pattern or Nordic Exceptionalism? European Sociological Review. 2005; 21(4):311–327. https://doi.org/10.1093/esr/jci022

17. Keller RT. Cross-Functional Project Groups in Research and New Product Development: Diversity, Communications, Job Stress, and Outcomes. The Academy of Management Journal. 2001; 44 (3):547–555. https://doi.org/10.5465/3069369

18. Putnam RD. E Pluribus Unum: Diversity and Community in the Twenty-First Century The 2006 Johan Skytte Prize Lecture. Scandinavian Political Studies. 2007 Jun; 30(2):137–174. https://doi.org/10.1111/j.1467-9477.2007.00176.x

19. Fukuyama F. Trust: The Social Virtues and the Creation of Prosperity. New York: Free Press; 1995.

20. Knack S, Keefer P. Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. The Quarterly Journal of Economics. 1997; 112(4):1251–1288. https://doi.org/10.1162/003355300555475

21. La Porta R, Lopez-de-Silanes F, Shleifer A, Vishny RW. Trust in Large Organizations. The American Economic Review. 1997; 87(2):333–338.

22. Putnam RD. Making Democracy Work: Civic Traditions in Modern Italy. Princeton, NJ: Princeton University Press; 1993.

23. Uslaner EM. The Moral Foundations of Trust. Cambridge: Cambridge University Press; 2002.

24. You Js. Social Trust: Fairness Matters More Than Homogeneity: Social Trust: Fairness Matters. Political Psychology. 2012; 33(5):701–721. https://doi.org/10.1111/j.1467-9221.2012.00893.x

25. Zak PJ, Knack S. Trust and Growth. The Economic Journal. 2001; 111(470):295–321. https://doi.org/10.1111/1468-0297.00609

26. Ostrom E, Walker J, Gardner R. Rules, Games and Common-Pool Resources. The American Political Science Review. 1992; 86(2):404–417.

27. Hardin G. The Tragedy of the Commons. Science. 1968; 162(3859):1243–1248. https://doi.org/10.1126/science.162.3859.1243 PMID: 5699198

28. Myers RA, Hutchings JA, Barrowman NJ. Why Do Fish Stocks Collapse? The Example of Cod in Atlantic Canada. Ecological Applications. 1997; 7(1):91–106. _eprint: https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/1051-0761%281997%29007%5B0091%3AWDFSCT%5D2.0.CO%3B2. Available from: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/1051-0761%281997%29007%5B0091%3AWDFSCT%5D2.0.CO%3B2.

29. Ostrom E. Governing the Commons: The Evolution of Institutions for Collective Action. The Political Economy of Institutions and Decisions. Cambridge; New York: Cambridge University Press; 1990.

30. De Moor T. Homo Cooperans: Institutions for Collective Action and the Compassionate Society [Inaugural Lecture]; 2013.

31. Janssen MA, Holahan R, Lee A, Ostrom E. Lab Experiments for the Study of Social-Ecological Systems. Science. 2010; 328(5978):613–617. https://doi.org/10.1126/science.1183532 PMID: 20431012

32. Axelrod R, Hamilton WD. The Evolution of Cooperation. Science. 1981; 211(4489):1390–1396. https://doi.org/10.1126/science.7466396 PMID: 7466396

33. Chaudhuri A. Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. Experimental Economics. 2011; 14(1):47–83. https://doi.org/10.1007/s10683-010-9257-1

34. De Oliveira A, Croson RTA, Eckel CC. One Bad Apple: Uncertainty and Heterogeneity in Public Good Provision. SSRN Electronic Journal. 2009;. https://doi.org/10.2139/ssrn.1883766

35. Levati MV, Sutter M, van der Heijden E. Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information. Journal of Conflict Resolution. 2007; 51(5):793–818. https://doi.org/10.1177/0022002707302796

36. Ostrom E. The value-added of laboratory experiments for the study of institutions and common-pool resources. Journal of Economic Behavior & Organization. 2006 Oct; 61(2):149–163. Available from: http://www.sciencedirect.com/science/article/pii/S0167268106001041.

37. Meinzen-Dick R, Chaturvedi R, Domènech L, Ghate R, Janssen M, Rollins N, et al. Games for Groundwater Governance: Field Experiments in Andhra Pradesh, India. Ecology and Society. 2016; 21(3). https://doi.org/10.5751/ES-08416-210338

38. Tajfel H. Experiments in Intergroup Discrimination. Scientific American. 1970; 223(5):96–103. Publisher: Scientific American, a division of Nature America, Inc. Available from: https://www.jstor.org/stable/24927662.

39. Tajfel H, Billig MG, Bundy RP, Flament C. Social Categorization and Intergroup Behaviour. European Journal of Social Psychology. 1971; 1(2):149–178. https://doi.org/10.1002/ejsp.2420010202

40. Aksoy O. Effects of Heterogeneity and Homophily on Cooperation. Social Psychology Quarterly. 2015; 78(4):324–344. https://doi.org/10.1177/0190272515612403

41. Masella P, Meier S, Zahn P. Incentives and Group Identity. Games and Economic Behavior. 2014; 86:12–25. https://doi.org/10.1016/j.geb.2014.02.013

42. De Moor T. De Herontdekking van Het Collectief. 2013;.

43. De Moor T. De Toekomst van Bewonersoverlegorganisaties. 2018;.

44. Chan KS, Mestelman S, Moir R, Muller RA. Heterogeneity and the Voluntary Provision of Public Goods. Experimental Economics. 1999; 2(1):5–30. https://doi.org/10.1023/A:1009984414401

45. Shanmugaratnam N. Nationalisation, Privatisation and the Dilemmas of Common Property Management in Western Rajasthan. The Journal of Development Studies. 1996; 33(2):163–187. https://doi.org/10.1080/00220389608422461

46. Adhikari DB, Lovett JC. Institutions and collective action: Does heterogeneity matter in community-based resource management? The Journal of Development Studies. 2006 Apr; 42(3):426–445. Publisher: Routledge _eprint: https://doi.org/10.1080/00220380600576201. Available from: https://doi.org/10.1080/00220380600576201.

47. Bardhan P. Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communities in South India. Economic Development and Cultural Change. 2000; 48(4):847–865. https://doi.org/10.1086/452480

48. Ledyard JO. Public Goods: A Survey of Experimental Research. Division of the Humanities and Social Sciences, California Institute of Technology; 1993.

49. Easter KW, Palanisami K. Tank Irrigation in India and Thailand: An Example of Common Property Resource Management. Minneapolis: Department of Agricultural and Applied Economics Staff, University of Minnesota; 1986.

50. Jayaraman TK. Farmers' Organisations in Surface Irrigation Projects: Two Empirical Studies from Gujarat. Economic and Political Weekly. 1981; 16(39):A89–A98.

51. Becker CD, Ostrom E. Human Ecology and Resource Sustainability: The Importance of Institutional Diversity. Annual Review of Ecology and Systematics. 1995; 26:113–133. https://doi.org/10.1146/annurev.es.26.110195.000553

52. Bowles S, Gintis H. Social Capital and Community Governance. The Economic Journal. 2002; 112 (483):F419–F436. https://doi.org/10.1111/1468-0297.00077

53. Boyd R, Richerson PJ. Culture and the Evolutionary Process. Chicago: University of Chicaco Press; 1985.

54. Ellickson RC. Order without Law: How Neighbors Settle Disputes. Harvard University Press; 1991.

55. Jones EC. Wealth-Based Trust and the Development of Collective Action. World Development. 2004; 32(4):691–711. https://doi.org/10.1016/j.worlddev.2003.10.009

56. Nettle D, Dunbar RIM. Social Markers and the Evolution of Reciprocal Exchange. Current Anthropology. 1997; 38(1):93–99. https://doi.org/10.1086/204588

57. Portes A, Landolt P. Social Capital: Promise and Pitfalls of Its Role in Development. Journal of Latin American Studies. 2000; 32(2):529–547. https://doi.org/10.1017/S0022216X00005836

58. Putnam RD. Bowling Alone: The Collapse and Revival of American Community. New York: Simon & Schuster; 2000.

59. Singleton S. How Distributional Conflicts Shape Co-Management Institutions and Collaborative Environmental Policy in the United States. In: "Economic Inequality, Collective Action and Environmental Sustainability" Workshop. Santa Fe Institute; 2001.

60. Varughese G, Ostrom E. The Contested Role of Heterogeneity in Collective Action: Some Evidence from Community Forestry in Nepal. World Development. 2001; 29(5):747–765. https://doi.org/10.1016/S0305-750X(01)00012-2

61. Johnson RN, Libecap GD. Contracting Problems and Regulation: The Case of the Fishery. The American Economic Review. 1982; 72(5):1005–1022.

62. Chen Y, Li SX. Group Identity and Social Preferences. The American Economic Review. 2009; 99 (1):431–457. https://doi.org/10.1257/aer.99.1.431

63. Eckel CC, Grossman PJ. Managing Diversity by Creating Team Identity. Journal of Economic Behavior & Organization. 2005; 58(3):371–392. https://doi.org/10.1016/j.jebo.2004.01.003

64. Timilsina RR, Kotani K, Kamijo Y. Sustainability of common pool resources. PLoS ONE. 2017 Feb; 12 (2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5315376/.

65. Habyarimana J, Humphreys M, Posner DN, Weinstein JM. Why Does Ethnic Diversity Undermine Public Goods Provision? American Political Science Review. 2007; 101(4):709–725. https://doi.org/10.1017/S0003055407070499

66. Brewer MB, Kramer RM. Choice Behavior in Social Dilemmas: Effects of Social Identity, Group Size, and Decision Framing. Journal of Personality and Social Psychology. 1986; 50(3):543–549. https://doi.org/10.1037/0022-3514.50.3.543

67. Yamagishi T, Kiyonari T. The Group as the Container of Generalized Reciprocity. Social Psychology Quarterly. 2000; 63(2):116–132. https://doi.org/10.2307/2695887

68. Gehrig S, Schlüter A, Hammerstein P. Sociocultural Heterogeneity in a Common Pool Resource Dilemma. PLOS ONE. 2019; 14(1):e0210561. https://doi.org/10.1371/journal.pone.0210561 PMID: 30653546

69. Barr A. Familiarity and Trust: An Experimental Investigation. Centre for the Study of African Economies, University of Oxford; 1999. 1999–23.

70. Coleman JS. Foundations of Social Theory. Cambridge, MA: Harvard University Press; 1994.

71. Glaeser EL, Laibson DI, Scheinkman JA, Soutter CL. Measuring Trust. The Quarterly Journal of Economics. 2000; 115(3):811–846. https://doi.org/10.1162/003355300554926

72. Glaeser EL, Laibson D, Scheinkman JA, Soutter CL. What Is Social Capital? The Determinants of Trust and Trustworthiness. National Bureau of Economic Research; 1999. 7216.

73. Alesina A, La Ferrara E. Participation in Heterogeneous Communities. 2000;.

74. Bouckaert J, Dhaene G. Inter-Ethnic Trust and Reciprocity: Results of an Experiment with Small Businessmen. European Journal of Political Economy. 2004; 20(4):869–886. https://doi.org/10.1016/j.ejpoleco.2003.08.006

75. Fischbacher U. Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. Experimental Economics. 2007; 10(2):171–178. https://doi.org/10.1007/s10683-006-9159-4

76. Greiner B. The Online Recruitment System ORSEE 2.0—A Guide for the Organization of Experiments in Economics. 2004; p. 67.

77. Berg J, Dickhaut J, McCabe K. Trust, Reciprocity, and Social History. Games and Economic Behavior. 1995; 10(1):122–142. https://doi.org/10.1006/game.1995.1027

78. Evans AM, Revelle W. Survey and Behavioral Measurements of Interpersonal Trust. Journal of Research in Personality. 2008; 42(6):1585–1593. https://doi.org/10.1016/j.jrp.2008.07.011

79. Bahry DL, Wilson RK. Confusion or Fairness in the Field? Rejections in the Ultimatum Game under the Strategy Method. Journal of Economic Behavior & Organization. 2006; 60(1):37–54. https://doi.org/10.1016/j.jebo.2004.07.005

80. Johnson ND, Mislin AA. Trust Games: A Meta-Analysis. Journal of Economic Psychology. 2011; 32 (5):865–889. https://doi.org/10.1016/j.joep.2011.05.007

81. Aksoy O. Crosscutting Circles in a Social Dilemma: Effects of Social Identity and Inequality on Cooperation. Social Science Research. 2019; 82:148–163. https://doi.org/10.1016/j.ssresearch.2019.04.009 PMID: 31300075

82. Aksoy O, Weesie J. Beliefs about the Social Orientations of Others: A Parametric Test of the Triangle, False Consensus, and Cone Hypotheses. Journal of Experimental Social Psychology. 2012 Jan; 48 (1):45–54. https://doi.org/10.1016/j.jesp.2011.07.009

83. Bilancini E, Boncinelli L, Capraro V, Celadin T, Di Paolo R. "Do the Right Thing" for Whom? An Experiment on Ingroup Favouritism, Group Assorting and Moral Suasion. arXiv:200212464 [physics, q-bio]. 2020 Feb;ArXiv: 2002.12464. Available from: http://arxiv.org/abs/2002.12464.

84. Bouckaert J, Dhaene G. Inter-Ethnic Trust and Reciprocity: Results of an Experiment with Small Business Entrepreneurs. University of Antwerp, Faculty of Applied Economics; 2002. 2002022.

85. Fershtman C, Gneezy U. Discrimination in a Segmented Society: An Experimental Approach. The Quarterly Journal of Economics. 2001; 116(1):351–377. https://doi.org/10.1162/003355301556338

86. Billig MG, Tajfel H. Social Categorization and Similarity in Intergroup Behaviour. European Journal of Social Psychology. 1973; 3(1):27–52. https://doi.org/10.1002/ejsp.2420030103

87. Keyuan Z. Sino-Japanese joint fishery management in the East China Sea. Marine Policy. 2003 Mar; 27(2):125–142. Available from: http://www.sciencedirect.com/science/article/pii/S0308597X02000866.

88. Muscolino M. The Yellow Croaker War: Fishery Disputes between China and Japan, 1925–1935. Environmental History. 2008 Apr; 13(2):306–324. Publisher: Oxford Academic. Available from: https://academic.oup.com/envhis/article/13/2/306/619103.

89. Wood SN. Generalized Additive Models: An Introduction with R, Second Edition. CRC Press; 2017.

90. Wood SN. Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models. Biometrics. 2006; 62(4):1025–1036. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1541-0420.2006.00574.x. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1541-0420.2006.00574.x PMID: 17156276

91. Andreoni J. Why Free Ride?: Strategies and Learning in Public Goods Experiments. Journal of Public Economics. 1988; 37(3):291–304. https://doi.org/10.1016/0047-2727(88)90043-6

92. González LG, Güth W, Levati MV. When Does the Game End? Public Goods Experiments with Non-Definite and Non-Commonly Known Time Horizons. Economics Letters. 2005; 88(2):221–226. https://doi.org/10.1016/j.econlet.2005.02.010

93. Hey JD, Neugebauer T, Sadrieh A. An Experimental Analysis of Optimal Renewable Resource Management: The Fishery. Environmental and Resource Economics. 2009 May; 44(2):263. Available from: https://doi.org/10.1007/s10640-009-9285-5.

94. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association. 1952; 47(260):583–621. Publisher: [American Statistical Association, Taylor & Francis, Ltd.]. Available from: https://www.jstor.org/stable/2280779.

95. Hlavac M. Stargazer: Well-Formatted Regression and Summary Statistics Tables. R Package Version 5.2.2.; 2018.

96. Balliet D, Wu J, De Dreu CKW. Ingroup favoritism in cooperation: A meta-analysis. Psychological Bulletin. 2014; 140(6):1556–1581. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0037737 PMID: 25222635

97. Fehr E, Fischbacher U. Social norms and human cooperation. Trends in Cognitive Sciences. 2004 Apr; 8(4):185–190. Available from: http://www.sciencedirect.com/science/article/pii/S1364661304000506 PMID: 15050515

98. Herrmann B, Thöni C. Measuring conditional cooperation: a replication study in Russia. Experimental Economics. 2009 Mar; 12(1):87–92. Available from: https://doi.org/10.1007/s10683-008-9197-1.

99. Fischbacher U, Gächter S, Fehr E. Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters. 2001 Jun; 71(3):397–404. Available from: http://www.sciencedirect.com/science/article/pii/S0165176501003949.

100. Chaudhuri A, Paichayontvijit T. Conditional cooperation and voluntary contributions to a public good. Economics Bulletin. 2006; p. 1–14.

101. Ahmed A. Conditional reciprocity in the investment game. The Social Science Journal. 2011 Jun; 48(2):404–408. Publisher: Routledge _eprint: https://doi.org/10.1016/j.soscij.2010.11.005. Available from: https://doi.org/10.1016/j.soscij.2010.11.005.

102. Frey BS, Meier S. Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment. American Economic Review. 2004 Nov; 94(5):1717–1722. Available from: https://pubs.aeaweb.org/doi/10.1257/0002828043052187.

103. Krupka E, Weber RA. The focusing and informational effects of norms on pro-social behavior. Journal of Economic Psychology. 2009 Jun; 30(3):307–320. Available from: http://www.sciencedirect.com/science/article/pii/S0167487008001128.

104. Dayton-Johnson J, Bardhan P. Inequality and Conservation on the Local Commons: A Theoretical Exercise. The Economic Journal. 2002; 112(481):577–602. https://doi.org/10.1111/1468-0297.00731

105. Anderies JM, Janssen MA, Bousquet F, Cardenas JC, Castillo D, Lopez MC, et al. The Challenge of Understanding Decisions in Experimental Studies of Common Pool Resource Governance. Ecological Economics. 2011; 70(9):1571–1579. https://doi.org/10.1016/j.ecolecon.2011.01.011

106. Falk A, Zehnder C. A City-Wide Experiment on Trust Discrimination. Journal of Public Economics. 2013; 100:15–27. https://doi.org/10.1016/j.jpubeco.2013.01.005

107. Falk A, Heckman JJ. Lab Experiments Are a Major Source of Knowledge in the Social Sciences | Science; 2009. https://science.sciencemag.org/content/326/5952/535 PMID: 19900889

108. Levitt SD, List JA. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? 2007; p. 81.

109. Levitt SD, List JA. Field Experiments in Economics: The Past, the Present, and the Future. European Economic Review. 2009; 53(1):1–18. https://doi.org/10.1016/j.euroecorev.2008.12.001

110. Exadaktylos F, Espín AM, Brañas-Garza P. Experimental Subjects Are Not Different. Scientific Reports. 2013; 3:1213. https://doi.org/10.1038/srep01213 PMID: 23429162

111. Fehr E, Fischbacher U, von Rosenbladt B, Schupp J, Wagner GG. A Nation-Wide Laboratory: Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Survey. Rochester, NY: Social Science Research Network; 2003. ID 385120.

112. Bednarik P, Linnerooth-Bayer J, Magnuszewski P, Dieckmann U. A Game of Common-pool Resource Management: Effects of Communication, Risky Environment and Worldviews. Ecological Economics. 2019 Feb; 156:287–292. Available from: http://www.sciencedirect.com/science/article/pii/S0921800918302647.

113. Lacomba JA, Lagos F, Perote J. The Lazarillo's game: Sharing resources with asymmetric conditions. PLoS ONE. 2017; 12(7). Publisher: Public Library of Science. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5509147/ PMID: 28704408

114. Neugebauer T, Perote J, Schmidt U, Loos M. Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. Journal of Economic Psychology. 2009 Feb; 30 (1):52–60. Available from: http://www.sciencedirect.com/science/article/pii/S0167487008000780.