Journal of Economic Psychology 105 (2024) 102758

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Decision-makers self-servingly navigate the equality-efficiency trade-off of free partner choice in social dilemmas among unequals[☆]

Luuk L. Snijder [a,*], Mirre Stallen [a,b], Jörg Gross [c]

[a] *Social, Economic and Organisational Psychology, Leiden University, Leiden, the Netherlands*
[b] *Poverty Interventions, Center for Applied Research on Social Sciences and Law, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands*
[c] *Department of Social and Economic Psychology, Institute of Psychology, University of Zurich, Zurich, Switzerland*

ARTICLE INFO

Keywords:
Cooperation
Inequality
Segregation
Partner choice
Decision-making

ABSTRACT

Cooperation is more likely upheld when individuals can choose their interaction partner. However, when individuals differ in their endowment or ability to cooperate, free partner choice can lead to segregation and increase inequality. To understand how decision-makers can decrease such inequality, we conducted an incentivized and preregistered experiment in which participants (n=500) differed in their endowment and cooperation productivity. First, we investigated how these individual differences impacted cooperation and inequality under free partner choice in a public goods game. Next, we calculated if and how decision-makers *should* restrict partner choice if their goal is to decrease inequality. Finally, we studied whether decision-makers actually *did* decrease inequality when asked to allocate endowment and productivity factors between individuals, and combine individuals into pairs of interaction partners for a two-player public goods game. Our results show that without interventions, free partner choice, indeed, leads to segregation and increases inequality. To mitigate such inequality, decision-makers *should* curb free partner choice and force individuals who were assigned different endowments and productivities to form pairs with each other. However, this comes at the cost of lower overall cooperation and earnings, showing that the restriction of partner choice results in an equality-efficiency trade-off. Participants who acted as third-parties were *actually* more likely to prioritize inequality reduction over efficiency maximization, by forcing individuals with unequal endowment and productivity levels to form pairs with each other. However, decision-makers who had a 'stake in the game' self-servingly navigated the equality-efficiency trade-off by preferring partner choice interventions that benefited themselves. These preferences were partly explained by norms on public good cooperation and redistribution, and participants' social preferences. Results reveal potential conflicts on how to govern free partner choice stemming from diverging preferences 'among unequals'.

☆ **Data and code availability:** All data, analyses codes, and materials are openly available in an OSF repository (https://doi.org/10.17605/OSF.IO/3WRSU). There are no restrictions to accessing the data. Additional information can be requested from the corresponding author at l.snijder@fsw.leidenuniv.nl.

* Corresponding author at: Department of Social, Economic and Organisational Psychology, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands.

E-mail addresses: l.snijder@fsw.leidenuniv.nl (L.L. Snijder), m.stallen@fsw.leidenuniv.nl (M. Stallen), gross@psychologie.uzh.ch (J. Gross).

https://doi.org/10.1016/j.joep.2024.102758
Received 30 January 2024; Received in revised form 9 August 2024; Accepted 11 August 2024
Available online 13 August 2024
0167-4870/

1. Introduction

Cooperation is crucial for the functioning of groups and society at large (Hardin, 1968; Ostrom, 1998; Pfeiffer & Nowak, 2006; van Dijk & De Dreu, 2021). However, individuals can be tempted to free-ride on others' cooperation to maximize their own outcomes, creating a social dilemma between the optimal choice for an individual and the optimal choice for the group.

Previous research showed that cooperation is more likely upheld when individuals can choose their interaction partner (André & Baumard, 2011; Barclay, 2013, 2016; Barclay & Willer, 2007; Baumard et al., 2013; Raihani & Barclay, 2016). Under free partner choice, individuals willing to cooperate can find others who are also willing to cooperate, thereby avoiding the risk of being exploited by free-riders that try to take advantage of cooperators. As a result, free-riders have an incentive to start cooperating as well, to avoid isolation, and to render themselves attractive interaction partners again. Via this mechanism, partner choice can create a 'market' for the best interaction partners. It has been argued that such a mechanism has contributed to the evolution of fairness and morality in humans (Baumard et al., 2013; Raihani & Barclay, 2016; van Winden, 2023). Partner choice is also related to 'voting by foot', the idea that free mobility can allow people to sort themselves most efficiently (Bayer & McMillan, 2012; Gürerk et al., 2014; Tiebout, 1956). Theoretically, voting by foot can lead to considerable benefits for the provision of public goods, as people will continuously abandon less functional environments for environments governed by institutions that are better performing or harbour more cooperators.

Yet, what most partner choice models neglect, is that individuals not only differ in their willingness to cooperate but also in their *ability* to cooperate. Specifically, people can differ in the resources they have available for cooperation (inequality in endowments) and in their productivity when they do cooperate (inequality in productivity) (Hauser et al., 2019; Melamed et al., 2022; Stallen et al., 2023). Under such inequalities, some individuals can contribute more to public goods than others and are better positioned to reciprocate cooperation, making them more attractive interaction partners regardless of their willingness to cooperate. Recent studies showed that the presence of such structural individual differences, i.e., those that cannot be controlled or easily changed by the individual, leads to a segregation of the population based on endowment and productivity levels under free partner choice (Melamed et al., 2022; Stallen et al., 2023). In a nutshell, the advantaged (those with a high endowment and a high productivity) flock together and cooperate with each other to maximize their wealth, leading others to increasingly lag behind (Henrich & Boyd, 2008; Merton, 1988; Reardon & Bischoff, 2011; Stallen et al., 2023). Outside of the lab, such segregation has been linked to political polarization and intergroup conflict (Bhavnani et al., 2014).

To combat such segregation, societies may actively try to force individuals who are 'well-off' to interact with those who are 'worse-off' by restricting partner choice 'among unequals'. For example, policies can restrict free school choice such that children from different social classes are more likely to interact with each other compared to when school choice is not restricted (Boterman et al., 2019). Likewise, social housing programs may restrict housing choice to a group of select individuals in order to promote the development of more diverse neighbourhoods (Bruch, 2014). Societies may also attempt to pre-empt the emergence of structural individual differences in the first place by, for example, supporting members of disadvantaged groups via public education and healthcare, aiming to 'level the playing field' (Chetty et al., 2014; Heckman, 2006; Roemer & Trannoy, 2016).

How such interventions actually influence inequality and cooperation and to which degree decision-makers are willing to implement them is an open question. To address this, we confront participants with a stylized social dilemma, the two-player public goods game. In this game, each participant can use their endowment to generate a benefit for their partner at a cost to themselves by investing in cooperation. In our setup, participants differ in their endowment (endowment inequality) and how much benefit they generate through cooperation (inequality in productivity). Furthermore, participants can choose with whom they want to preferably interact.

We use this setup to investigate three primary goals: First, we aimed to replicate previous findings (Melamed et al., 2022; Stallen et al., 2023), showing that people better positioned to cooperate (i.e., those with a high endowment and high productivity) preferably choose to interact with each other. Hence, free partner choice 'among unequals' leads to segregation of people by structural characteristics with increased wealth inequality as a result.

Second, based on the choices of our participants, we aimed to derive how people should have chosen partners, if the goal would be to (i) minimize wealth inequality or (ii) maximize overall earnings (regardless how these earnings are distributed, i.e., efficiency). This allowed us to infer: What partner matching policy would lead to the lowest level of wealth inequality or generate maximal overall wealth, and do these two goals exclude each other?

Third, we investigated to which degree decision-makers are willing to implement partner matching interventions. Specifically, we asked uninvolved third-parties, as well as involved decision-makers, how they would pair people with different structural characteristics (i.e., different levels of endowment and productivity factors), if they would have the power to do so. Further, we also asked third-parties and decision-makers how they would distribute these structural characteristics across people, testing whether they would prefer a 'levelling of the playing field', if they would have the power to do so.

To further understand what drives these choices, we investigated what beliefs underlie investments in cooperation. Additionally, we examined how the general inclination of participants to act pro-socially (their 'social preferences') might influence their willingness to implement different partner matching interventions, as individuals with prosocial preferences may be more likely to implement interventions aimed to reduce inequality (Fehr & Charness, 2023; Fehr & Fischbacher, 2002).

As expected, our data show that (1) under free partner choice, segregation and wealth inequality increase 'among unequals' (i.e., when participants' endowment and productivity levels differ), because those assigned a high endowment and high productivity prefer to only interact with each other. In order to design effective interventions to prevent this, our data reveals (2) that partner matching interventions involve an equality-efficiency trade-off, such that interventions would either reduce wealth inequality or increase efficiency, but not both at the same time. Lastly, (3) we find that uninvolved third-parties prioritize interventions aimed to reduce wealth inequality, whereas involved decision-makers with a high endowment and high productivity prefer partner matchings that serve themselves. These decisions were partly explained by norms on cooperation and participants' social preferences. Taken together, results reveal potential conflicts on how to govern partner choice in a 'society of unequals'.

2. Methods

2.1. Participants and ethics

Our study was approved by the ethics committee of the Institute of Psychology at the University of Zurich (approval nr. 22.10.5) and did not involve any deception. The study was programmed in Qualtrics. Participants were recruited via Prolific (n=500, 49.80 % female; age ranged between 18 and 56 years; M=35.71, SD=8.36), provided informed consent, and received full debriefing after participating. They received a standard fee of £4.00 and their decisions were fully incentivized (M=£2.76, SD=£1.47; see 2.5. for details). Participation took approximately 40 min. No participants were excluded from the final analyses following our pre-registered exclusion criteria (see 2.6.).

2.2. Study design

Participants entered their Prolific ID, read the information letter, and signed the online informed consent. After giving informed consent, participants were instructed that their decisions, and those of other participants, would influence both their own payment and that of others. Instructions explained all rules to participants using neutral language throughout. After the main task (see below) was explained, participants answered 11 practice questions to probe their understanding of the task. Only after all practice questions were answered correctly, participants could start.

2.2.1. Two-player public goods game

In the main task, participants were confronted with a two-player public goods game. In this game, two players are given a certain number of units (their endowment) and each player, simultaneously, decides how many of their units to contribute to the public good (as a measure of cooperation). Every unit that is contributed to the public good is multiplied by a certain factor and is then evenly distributed among pairs. This factor is always lower than 2 but higher than 1. Thus, the game introduces the social dilemma of cooperation: While mutually contributing all units to the public good (i.e., full cooperation) increases joint earnings, it is always optimal to not cooperate and keep all units from a rational selfish perspective. This is because (i) if a partner is not cooperating, it is best to not cooperate either. If (ii) a partner is cooperating, participants can earn most by withholding their own units, since, in this case, they do not pay the cost of cooperation but can free-ride on the cooperation of the partner.

2.2.2. Inequality in endowments and ability to cooperate ('player types')

Importantly, in our setting, participants differ on two dimensions in the two-player public goods game: The units they have available for cooperation and the factor by which each unit is multiplied when invested into the public good. We refer to the first as endowment inequality (either high, $H_e = 75$ or low, $L_e = 25$) and the second as inequality in productivity (either high, $H_p = 1.7$ or low, $L_p = 1.3$) (replicating Stallen et al., 2023).

Before the public goods game started, each participant was randomly assigned one out of four possible types; the H_eH_p, H_eL_p, L_eH_p, or L_eL_p type. For example, a participant with the H_eH_p type had 75 units (high endowment) and each unit that they dedicated for cooperation was multiplied by 1.7 (high productivity). A participant with an L_eL_p type, on the other hand, only had 25 units (low endowment) and each unit they dedicated for cooperation was multiplied by only 1.3 (low productivity). A participant with a mixed type like L_eH_p, for example, had 25 units (low endowment) but each unit they dedicated for cooperation was multiplied by 1.7 (high productivity). We used neutral labels to refer to these types during the study: type 1, type 2, type 3, and type 4. We collected data of 100 participants per type.

2.2.3. Partner choice preferences

Each participant was asked to indicate which type they preferred to be paired with. Preferences were indicated by ranking the four different types (i.e., H_eH_p, H_eL_p, L_eH_p, or L_eL_p) in the order of their preference, from most preferred (1st choice) to least preferred (4th choice). Participants were instructed that after all participants completed the study, we would pair them with their first choice, but only when there was another participant assigned a type of their first choice who placed the participant's type at the top of their list as well. For example, if a participant with an L_eL_p type preferred to be paired with an H_eH_p type, the participant with the L_eL_p type was paired with their preferred type if there was at least one participant with an H_eH_p type who preferred to be paired with an L_eL_p type. When participants could not be paired with their first choice, because there were no other participants who preferred to be paired with the type of the participant, the participant was paired with the partner type of their second choice. If this was not possible, they would be paired with their third choice, and so on. If multiple participants preferred to be paired with one available type, a random draw determined who got paired with the other participant.

2.2.4. Cooperation choices and beliefs

After making the ranking of their preferred interaction partners, we used the strategy method to measure cooperation rates for each possible partner type. Specifically, participants were asked to indicate how many units they wanted to contribute to the public good for each possible type they could be paired with. They made these decisions in a random order. We further asked how many units they believed each type would contribute to the public good (if paired with them). Participants indicated their cooperation and beliefs at the same time. Using this so-called 'strategy method' allowed us to calculate which decisions participants should make if their goal would be to reduce inequality or increase efficiency, because we gathered data on all possible outcomes, including rare occurrences by choice, such as H_eH_p-L_eL_p pairings. However, the strategy method might not fully capture the immediacy and pressure of real-time decision making. Despite such limitations of the strategy method, many treatment effects using direct-response methods do replicate effects found with the strategy method (Brandts & Charness, 2011).

2.2.5. Creation of types ('type intervention')

Subsequently, we probed the implementation of different possible interventions of free partner choice. First, we asked participants how they would create different types if they could assign endowments and productivity levels. Specifically, participants were presented with two entirely undefined ('empty') types. They were instructed to assign one endowment level ($H_e = 75$ or $L_e = 25$) and one productivity level ($H_p = 1.7$ or $L_p = 1.3$) to each empty type by selecting and dragging either a high or low endowment and a high or low productivity level to each type. For instance, participants could create two maximally unequal types by assigning one type a high endowment and high productivity (H_eH_p) and the other type a low endowment and low productivity (L_eL_p). Alternatively, they could choose to 'level the playing field' and distribute endowment and productivity equally; creating one L_eH_p and one H_eL_p type. Note that each level (i.e., high endowment, low endowment, high productivity, low productivity) could only be assigned once. Hence, it was not possible to simply create two H_eH_p types, or any combination in which both types had a high (or low) endowment or high (or low) productivity (e.g., H_eH_p and H_eL_p types or L_eH_p and L_eL_p types).

This method was aimed to test whether participants would rather prefer creating a population of types in which differences in structural characteristics were minimized (i.e., one type has a high endowment but low productivity while the other type has a low endowment but high productivity) or a population of types in which differences in structural characteristics were maximized, creating a population of 'unequals' (i.e., one type has a high endowment and high productivity while the other type has a low endowment and low productivity).

2.2.6. Creation of pairs ('partner matching intervention')

Second, we asked participants to think about their preferred pairings in the public goods game by creating four pairs out of a group of eight types ($2 \times H_eH_p$, $2 \times H_eL_p$, $2 \times L_eH_p$, $2 \times L_eL_p$) based on what they believed to be a 'good combination'. Specifically, participants saw each of the four types (i.e., one H_eH_p, one H_eL_p, one L_eH_p, and one L_eL_p type) but without an assigned partner. Instead of a partner, an empty spot was displayed next to each type. They could then drag and drop the four partner types to these empty spots, with each partner type only being available once (so there was one H_eH_p type, one H_eL_p type, one L_eH_p type, and one L_eL_p type). The types that participants could drag-and-drop were displayed in a random order. Thus, they could, on the one extreme, create completely 'segregated' pairs: H_eH_p paired with H_eH_p, H_eL_p paired with H_eL_p, L_eH_p paired with L_eH_p and L_eL_p paired with L_eL_p. On the other extreme, they could also decide to create maximally mixed pairs; H_eH_p paired with L_eL_p, H_eL_p paired with L_eH_p, L_eH_p paired with H_eL_p, and L_eL_p paired with H_eH_p. It was also possible to create any combination in between.

This question was incentivised, such that participants learned that there was a fifty percent chance that participants, who previously indicated their partner preferences and made their cooperation decisions (see 2.2.3. and 2.2.4.), would not be paired based on their own partner preferences, but instead would be paired based on the pairs they created in this task.

2.2.7. Cooperation norms and redistribution preferences

After creating types and pairs, participants were asked how many units they believed each type should contribute to the public good for each possible pair (i.e., 10 pairs in total). For example, participants were shown a pair in which an H_eH_p and L_eL_p type were paired, and the participant had to indicate how many units they believed each type should contribute to the public good (i.e., "how many units do you think each type should contribute to their common account?").

Finally, we asked participants to redistribute earnings within each unique pair. That is, participants were instructed to imagine that both types within a pair contributed their full endowment to the public good. Participants were then asked how they would redistribute the total number of units in the created public good between both types (i.e., "how many units do you think each type should receive from the common account?").

2.2.8. Decision-makers' motives and social preferences

After each 'intervention question' (i.e., creating types, 2.2.5.; creating pairs, 2.2.6.; norms on cooperation, 2.2.7.; and redistribution preferences, 2.2.7.), participants were asked whether their decisions were primarily motivated by a desire to reduce inequality, maximize efficiency, based on (dis)similarity between types, random choice, or other considerations. For this purpose, participants were presented with a multiple-choice question in which they were asked to select one of these predefined motivations. Participants were limited to selecting only one motivation per question.

At the beginning of the experiment (i.e., before the two-player public goods game), we also measured individual-level social preferences to test whether general inclinations to act pro-socially influences partner choice preferences and/or partner matching interventions. Specifically, participants completed the incentivized fifteen-item social value orientation slider measure (SVO; Murphy et al., 2011). In this measure, participants decide how to allocate units between themselves and an unknown other person. Units can be allocated self-servingly (maximizing personal gain at the expense of the other person) or pro-socially (sacrificing own units to benefit the other person). Based on their decision pattern, we calculated participants' continuous measure of pro-sociality (the 'social value orientation angle'), with a higher social value orientation angle indicating more prosocial preferences, and a lower one indicating more selfish preferences.

2.2.9. Demographics and additional questions

At the very end of the study, participants answered questions regarding their general beliefs about the causes of inequality and regarding their attitudes to current inequality in their country (Almås et al., 2022). At last, they were asked to provide information about their demographics (gender, age, education, country, and political ideologies), and then were debriefed. Results on political ideologies can be found in Appendix A1.

2.3. Third-parties

Out of our 500 participants, 100 participants performed the study from a 'third-party' perspective. This means that they were not assigned a type at the beginning of the study (see 2.2.2.), that they did not indicate who they would like to be paired with (see 2.2.3.), and that they also did not make actual cooperation decisions (see 2.2.4.). Instead, they received instructions to the public goods game like the other participants, but were told that they would be "asked their opinion about the task, meaning that they would not engage in the task themselves." They then created types and pairs (see 2.2.5. and 2.2.6.), like the other participants, with the same possible consequences for decision-makers and indicated their cooperation norms (i.e., how much they believed decision-makers should cooperate) and redistribution preferences (see 2.2.7.). This allowed us to measure how people who have no 'stake in the game' would create types and pairs and to which degree their choices differ in comparison to participants who actually interacted in the public goods game and were assigned a type.

In the results, we refer to participants who were not assigned a type and were not involved in the game as 'third-parties'. We refer to participants who were assigned a type, made cooperation decisions, and for whom partner pairings potentially influenced their own earnings as 'decision-makers'. Letting uninvolved third-parties make decisions or evaluate certain actions has been widely used in the literature to test how choices or attitudes change when people have versus lack self-serving motives (see, e.g., Fehr & Fischbacher, 2004; Reuben & Riedl, 2013).

2.4. Attention checks

To determine serious participation, we included three attention checks and notified participants that failing two out of three attention checks would exclude them from data analysis (as pre-registered). First, after the practice questions, participants were asked to select the option Correct in response to the multiple-choice question: 'Please select Correct' with the options 'Correct' and 'Incorrect'. Second, after the final question on how many units participants believed types should contribute to the public good, there was a second attention check which asked participants to enter the number 20 in response to the question how many units they believed each type should contribute to the public good. Finally, before the demographics questionnaire started, participants were asked to type the word 'green' in a response box (spelling errors or differences in capital letters were not treated as missing this attention check). No participant missed more than one attention check, so we did not exclude any participant from the final analyses.

2.5. Incentives

Participants' decisions were incentivized, such that they could earn up to £6.50 based on their decisions. Participants could (1) earn units in the SVO slider measure, (2) in the public goods game, and (3) by correctly guessing the cooperation rates of other participants. Regarding (1), to incentivize participants' decisions using the SVO slider measure, participants were randomly paired twice with another participant in their group. The choices of both pairs were paid, with each participant once being selected as the allocator and once as the receiver. On average, participants earned £0.76 (SD=0.03, range: £0.64–0.82) in this task. Regarding (2), after all participants completed the study, we calculated how many units they earned in the public goods game. For this purpose, half of the participants were paired based on their free partner preferences (see 2.2.3.), and half of the participants were paired based on the pairs that they created 'top-down' (see 2.2.6.). On average, participants received £2.47 (SD=1.20, range: £0.65–5.55). Regarding (3), we compared participants' expectations with the actual cooperation rate of their partner. If their expectation was correct, participants received £0.20. Participants received, on average, £0.03 (SD=0.07, range: £0.00–0.20).

2.6. Pre-registration and data availability

We pre-registered the experimental design, analysis plan, sample size, and exclusion criteria via AsPredicted (https://aspredicted.org/VWJ_PFJ). Furthermore, all data, analysis codes, and materials are available on Open Science Framework (OSF; doi: 10.17605/OSF.IO/3WRSU). Sample sizes were determined based on previous research, before data collection, and no additional data were collected after the data had been analysed. Experimental instructions can be found in Appendix A2.

2.7. Statistical analyses

Statistical models were fit in R. We used (logistic) regression models to analyse single decisions from participants. Whenever participants made repeated decisions, we used multilevel (logistic) regression models including random intercepts for participants to account for violations of independence. All reported statistical tests were two-tailed. Complete results for all regression models can be found in Appendix A3.

3. Results

3.1. Free partner choice leads to segregation and increases inequality

We first replicated that free partner choice 'among unequals' leads to segregation (Stallen et al., 2023). Participants in the role of decision-makers, independent of their own type, most often preferred to be paired with partners who were assigned a high-endowment, high-productivity (H_eH_p) type (see Table S1), selecting them as their first preference in 80 % of their rankings (see Table 1 for participants' partner preferences). In contrast, decision-makers, independent of their own type, least often preferred to be paired with partners who were assigned a low-endowment low-productivity (L_eL_p) type (see Table S2), selecting them as their last preference in 81.50 % of the rankings (see Table 1).

To replicate that this leads to segregation, we ran 1,000 simulations based on decision-makers' partner preferences and cooperation decisions, and found that, on average, similar types most often (in 88.02 % of the pairs) were paired with each other under free partner choice (i.e., H_eH_p types were paired with each other, L_eL_p types were paired with each other, etc.). Partner preferences led to segregation by active choice (e.g., H_eH_p types preferred to be paired with other H_eH_p types and were paired together) or resulted from this choice because participants' first choice (an H_eH_p type) was already taken (e.g., decision-makers least often preferred to be paired with L_eL_p types, so L_eL_p types were forced to form pairs with each other). Pair formation across types (i.e., mixed pairs, e.g., H_eH_p-L_eL_p pairs) only happened in 11.98 % of the cases under free partner choice.

Cooperation levels were also impacted by partner type. Decision-makers cooperated relatively more when their partner was an H_eH_p type (*M*=49.42 %, *SE*=1.53) than when their partner was a less advantaged type ($b = 5.55$, $p < 0.001$, 95 % CI [3.61, 7.48]; see Table S3), and decision-makers cooperated relatively less when their partner was an L_eL_p type (*M*=40.77 %, *SE*=1.54) compared to when their partner was an H_eL_p or L_eH_p type ($b = -4.65$, $p < 0.001$, 95 % CI [-6.70, −2.60]; see Table S3). These cooperation patterns further increased, rather than decreased, pre-existing inequalities. To quantify this, we computed the average Gini coefficient. The Gini coefficient measures income inequality in a population based on average earnings, and ranges from 0 (perfect equality) to 1 (maximum inequality). We calculated the average Gini coefficient by computing the average earnings based on participants respective average cooperation rates with their respective partner type in each given pairing scenario across 1,000 simulations. Before decision-making in the public goods game under free partner choice, the Gini coefficient was 0.25, while after decision-making, inequality increased to 0.30. These results emphasize room for interventions to reduce inequality stemming from free partner choice.

3.2. Optimal interventions: restricting free partner choice leads to an equality-efficiency trade-off

Since we asked decision-makers how much they wanted to invest in cooperation with each possible partner type, we are able to calculate how decision-makers *should* choose partners ('partner matching intervention') if the goal was to either (a) maximally reduce wealth inequality or (b) maximize overall cooperation/efficiency (i.e., total earnings resulting from voluntary cooperation, regardless of how these earnings are distributed across types). Furthermore, we can calculate wealth inequality and overall cooperation/efficiency if structural characteristics would be distributed unequally (i.e., assuming only H_eH_p and L_eL_p types exist) or equally (i.e., assuming only L_eH_p and H_eL_p types exist). We, thus, investigated the combination of two classes of possible interventions: (i) assigning structural characteristics unequally (i.e., creating an environment with only H_eH_p / L_eL_p types) or equally (i.e., creating an environment with only H_eL_p / L_eH_p types) and (ii) forcing these types to 'mix' (i.e., H_eH_p with L_eL_p or L_eH_p with H_eL_p) or pair with similar others ('similar'; H_eH_p with H_eH_p and L_eL_p with L_eL_p or L_eH_p with L_eH_p and H_eL_p with H_eL_p; see Fig. 1a).

For each of these possible 'interventions' based on the creation and pairing of types, we computed an equality and an efficiency coefficient. For the equality coefficient, we first calculated the Gini coefficient by computing the average earnings based on their respective average cooperation rates with their respective partner type in each given pairing scenario. To ease interpretation, we reversed the Gini coefficient by subtracting it from 1 such that numbers closer to 1 indicate higher equality. The efficiency coefficient in each pair was computed by scaling the average generated earnings (i.e., the average earnings of each type minus their endowment) between 0 and 1, where 1 refers to the highest observed generated earnings and 0 refers to the lowest observed generated earnings in each given pairing scenario.

Interestingly, the best intervention to maximize equality and efficiency was not, as one might expect, to distribute structural characteristics (high/low endowment and high/low productivity levels) equally, creating H_eL_p and L_eH_p types (see Fig. 1a). Instead,

Table 1
Frequency of participants' first and last partner preferences based on their own type (in percentages).

	first preference				last preference			
own type	H_eH_p	H_eL_p	L_eH_p	L_eL_p	H_eH_p	H_eL_p	L_eH_p	L_eL_p
H_eH_p	84	7	4	5	3	5	8	84
H_eL_p	76	18	1	5	1	5	14	80
L_eH_p	79	10	5	6	3	8	9	80
L_eL_p	81	5	8	6	8	5	5	82

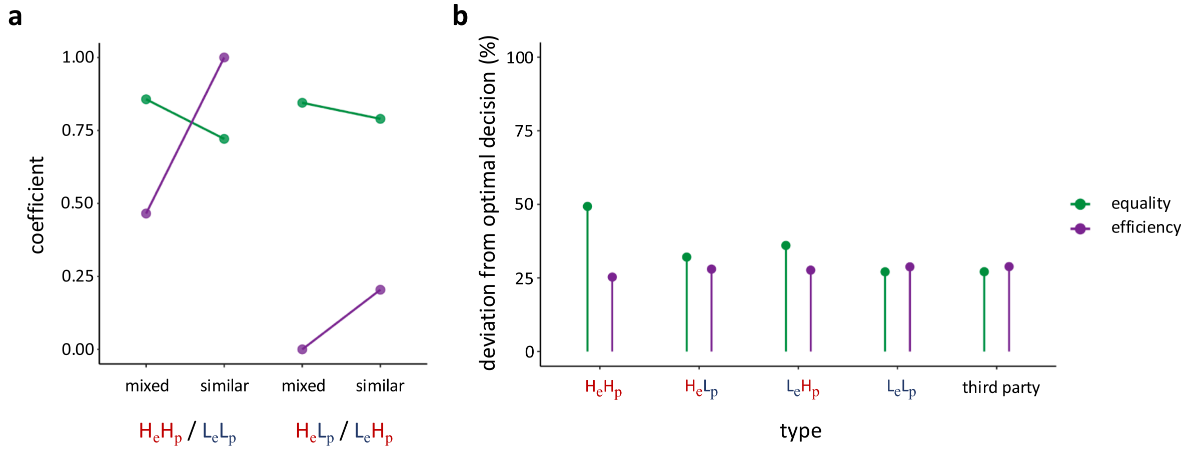

Fig. 1. Optimal and actual equality and efficiency in social dilemmas among unequals. a The equality (1 minus the Gini coefficient; in green) and efficiency coefficient (in purple; i.e., generated earnings scaled between 0 and 1) if participants created H_eH_p and L_eL_p types versus H_eL_p and L_eH_p types and, subsequently, created pairs of mixed or similar types. **b** The discrepancy between optimal choices (aiming either for maximal equality, green, or efficiency, purple; panel a) and actual decisions. Actual decisions were calculated as follows. For each type of participant (x-axis), we determined how often they created each possible pair. Based on the frequency of these created pairs, we calculated the weighted average earnings for each type. Using these weighted average earnings, we then computed the equality coefficient and the average generated wealth. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

wealth equality was highest when creating types that were structurally unequal (H_eH_p and L_eL_p types) but forced into mixed (H_eH_p-L_eL_p) pairs (equality coefficient = 0.86). Compared to free partner choice (equality coefficient = 0.70, see 3.1.), this increased equality by 0.16 points. In contrast, highest overall efficiency was observed when participants of similar types were paired (i.e., H_eH_p-H_eH_p and L_eL_p-L_eL_p pairs), creating 'segregated communities'. However, pairings consisting of similar types also resulted in lowest equality (equality coefficient = 0.72). The creation of H_eL_p and L_eH_p types yielded intermediate levels of equality, and the lowest levels of overall wealth. In this scenario, the formation of mixed (H_eL_p-L_eH_p) pairs resulted in higher equality levels compared to the formation of same type pairs. Yet, creating mixed pairs came at the expense of lowest overall efficiency.

This reveals two insights: First, distributing structural characteristics unequally across individuals (creating H_eH_p and L_eL_p types) maximizes wealth efficiency (i.e., highest cooperation levels) and equality. Yet, efficiency is maximized when those who are best-off (H_eH_p types) and those who are worst-off (L_eL_p types) are segregated, resulting in similar (H_eH_p-H_eH_p and L_eL_p-L_eL_p) pairs – a situation that also endogenously emerged under free partner choice. The highest level of equality, on the other hand, is observed when different types are paired and interact with each other (H_eH_p-L_eL_p pairs). Hence, we observe an efficiency-equality trade-off: When types generate the highest overall wealth (i.e., when H_eH_p-H_eH_p and L_eL_p-L_eL_p pairs are formed), 78 % of units are earned by H_eH_p types, while only 22 % of units are earned by L_eL_p types.

3.3. *Decision-makers self-servingly navigate the equality-efficiency trade-off*

Next, we examined 'intervention' preferences when participants were asked to create and pair types. In other words, we investigated what alternative to free partner choice participants would prefer themselves. Note that decision-makers, i.e., participants who were assigned a type and made cooperation choices, could be motivated to choose self-servingly. Specifically, there was a chance that their decisions to create pairs affected their own (type's) pairing and payoffs. Hence, for decision-makers, some pairings (i.e., being paired with an H_eH_p type) more likely generated value for themselves. Instead, participants in the role of third-parties had no 'stake in the game', since their choices could only affect the payoff of 'decision-makers' but not their own compensation for the study.

3.3.1. *Creation of types*

Third-parties did not significantly differ in how often they created H_eL_p and L_eH_p types (54 % of all decisions) or H_eH_p and L_eL_p types (46 % of all decisions; X^2 (1, N=100) = 0.64, p = 0.424). Furthermore, we found no credible evidence that decision-makers, compared to third-parties, differed in how they created types (see Table S4). This indicates that there was no strong preference on how to distribute structural characteristics (i.e., endowment and productivity to cooperate equally or unequally) and that the choices of how types were created was largely independent of one's own position in the game.

3.3.2. *Creation of pairs*

When participants in the role of third-parties created H_eL_p and L_eH_p types, they were subsequently more likely to create mixed (H_eL_p-L_eH_p) pairs rather than similar (H_eL_p-H_eL_p or L_eH_p-L_eH_p) pairs (see Fig. 2). Decision-makers were, independent of their own type, also most likely to subsequently create mixed pairs (see Table S5; see Fig. 2). Participants thus formed pairs that were complementing each other (i.e., H_eL_p-L_eH_p pairs) and avoided segregation when structural characteristics were divided equally.

However, as previously identified, the most effective strategy to maximize equality or efficiency would involve the creation of H_eH_p

and L_eL_p types (see Fig. 1a). Third-parties who created H_eH_p and L_eL_p types were most likely to create mixed (H_eH_p-L_eL_p) pairs and least likely to create similar (H_eH_p-H_eH_p and L_eL_p-L_eL_p) pairs (see Fig. 2). Hence, third-parties prioritized inequality reduction over efficiency maximization, by forcing H_eH_p types to pair with L_eL_p types. Decision-makers differed in how they formed pairs between H_eH_p and L_eL_p types, depending on their own type. L_eL_p decision-makers were, similar to third-parties, most likely to create mixed (H_eH_p-L_eL_p) pairs (b = -0.15, $p = 0.723$, 95 % CI [-1.01, 0.70]; see Table S7; see Fig. 2), resulting in a prioritisation of inequality reduction. In contrast, compared to third-parties, H_eH_p decision-makers were significantly more likely to create similar (H_eH_p-H_eH_p and L_eL_p-L_eL_p) pairs ($b = 1.12$, $p = 0.021$, 95 % CI [0.20, 2.11]; see Table S8; see Fig. 2) and less likely to create mixed (H_eH_p-L_eL_p) pairs (b = -1.98, $p < 0.001$, 95 % CI [-2.96, −1.08]; see Table S7; see Fig. 2), resulting in a prioritisation of efficiency maximization over inequality reduction. These decisions reveal self-serving choices in pair formation, as H_eH_p decision-makers earned more in similar (H_eH_p-H_eH_p) pairs than in mixed (H_eH_p-L_eL_p) pairs ($b = 21.89$, $p < 0.001$, 95 % CI [18.61, 25.17]; see Table S9), while L_eL_p decision-makers earned more in mixed (H_eH_p-L_eL_p) than in similar (L_eL_p-L_eL_p) pairs ($b = 15.18$, $p < 0.001$, 95 % CI [11.74, 18.62]; see Table S10).

Finally, a substantial proportion of third-parties (27.0 %) and decision-makers (34.75 %) created 'other' pairs (see Fig. 2). These participants did not have a preference for the creation of pairs that were either perfectly segregated or perfectly mixed. Overall, the actual decisions of how to intervene in free partner choice deviated by approximately 25 percent from optimal interventions for inequality reduction or efficiency maximization, as identified in 3.2. (see Fig. 1b). Importantly, H_eH_p decision-makers created types and pairs that resulted in 50 percent more inequality compared to what would be the optimal 'intervention' to reduce inequality. Hence, H_eH_p decision-makers served the interests of their own type, rather than the interests of those who were assigned a lower endowment and productivity level.

3.4. *Additional results*

3.4.1. *Origins of the equality-efficiency trade-off*

To better understand the origins of the equality-efficiency trade-off, we explored norms of third-parties with regards to others' contributions to the public good. Specifically, we asked for their opinions on what level of cooperation they regarded as normatively acceptable for each type in each unique pair, as well as their opinions on how wealth should be redistributed within these pairs (see 2.2.7.). Because structural characteristics (endowment and productivity levels) were randomly assigned to participants, and were thus based on luck, it would be reasonable to assume that third-parties believe that each type should cooperate maximally and that the created public good is shared equally to establish ex-post wealth equality. Interestingly, this is not what we observed. Instead, when third-parties were asked how much others should cooperate, they indicated that individuals should cooperate most when paired with a participant of their own type ($b = 6.91$, $p < 0.001$, 95 % CI [5.13, 8.68]; see Table S11; see Fig. 3a). This implies that H_eH_p types should cooperate more when paired with their own type than when paired with L_eL_p types (b = -7.68, $p < 0.001$, 95 % CI [-11.50, −3.86]; see Table S12), and vice versa ($b = 7.40$, $p < 0.001$, 95 % CI [3.64, 11.16]; see Table S13). This is indeed also what we found in the actual cooperation decisions of participants. This reveals that the lowered cooperation of H_eH_p types in mixed pairs is normatively expected and accepted by third-parties, to some degree.

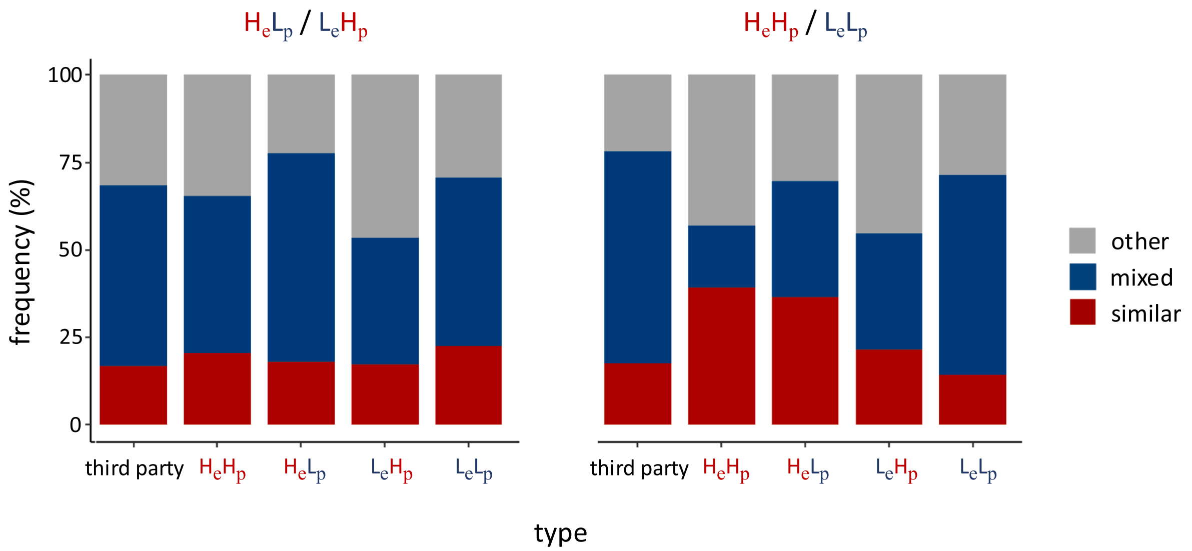

Fig. 2. Population composition following the creation of types and pairs by decision-makers. Stacked bar graph illustrating the relative frequency of created pairs following the creation of H_eL_p and L_eH_p types (left) or H_eH_p and L_eL_p types (right). Each bar represents the relative frequency of pairs being formed consisting of similar types (in red; similar; i.e., H_eL_p-H_eL_p or L_eH_p-L_eH_p pairs in the left panel and H_eH_p-H_eH_p or L_eL_p-L_eL_p pairs in the right), different types (in blue; mixed; i.e., H_eL_p-L_eH_p pairs in the left panel and H_eH_p-L_eL_p pairs in the right), or types not initially created (in grey; other; e.g., H_eL_p-L_eL_p pairs in the left panel and H_eH_p-L_eH_p pairs in the right, etc.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Yet, this (a) restricts overall welfare creation through the public good and (b) hinders the effectiveness of mixed pairs to use the public good for redistribution. Essentially, third-parties 're-created' the equality-efficiency trade-off that we observed in actual decisions on the normative level. In line with this, third-parties also redistributed more units to H_eH_p types than to L_eL_p types ($b = 4.67$, $p < 0.001$, 95 % CI [2.81, 6.52]; see Table S14; see Fig. 3b) in mixed pairs (i.e., H_eH_p-L_eL_p pairs). That is, when third-parties were asked to indicate how they would redistribute wealth between types that cooperated maximally within each hypothetical pair, third-parties did not choose to distribute public goods earnings equally.

3.4.2. Social preferences can reduce inequality via the creation of pairs

Lastly, we investigated how individual-level social preferences (measured using the separate SVO slider measure; see 2.2.8.) impacted partner choice, cooperation rates, and the creation of pairs. Social preferences did not impact partner preferences; decision-makers most often preferred to be paired with partners who were assigned a high endowment and a high productivity ($b = 0.01$, $p = 0.762$, 95 % CI [-0.02, 0.02]; see Table S15), and least often preferred to be paired with partners who were assigned a low-endowment low-productivity type ($b = 0.02$, $p = 0.742$, 95 % CI [-0.02, 0.02]; see Table S16), regardless of their social preferences.

Decision-makers with a higher social value orientation angle (indicating stronger prosocial preferences) did invest a larger part of their endowment in cooperation ($b = 0.52$, $p < 0.001$, 95 % CI [0.34, 0.70]; see Table S17). However, because partner preferences resulted in segregation and participants cooperated most with H_eH_p types, the impact of social preferences did not reduce wealth inequality in our setting.

Interestingly, social preferences did impact participants' willingness to pair more advantaged with less advantaged individuals. Participants with a higher social value orientation angle (indicating stronger prosocial preferences) were more likely to pair H_eH_p types with L_eL_p types ($b = 0.03$, $p = 0.006$, 95 % CI [0.01, 0.05]; see Table S20) and were mostly driven by self-reported motives to reduce inequality ($b = 0.03$, $p < 0.001$, 95 % CI [0.02, 0.05]; see Table S22), while those with a lower social value orientation angle (indicating stronger selfish preferences) were more likely to create similar (H_eH_p-H_eH_p and L_eL_p-L_eL_p) pairs ($b = -0.03$, $p = 0.009$, 95 % CI [-0.05, −0.01]; see Table S21) and reported that their choices were mostly driven by the motive to maximize efficiency ($b = -0.02$, $p = 0.010$, 95 % CI [-0.03, −0.004]; see Table S23) or driven by the (dis)similarity between types ($b = -0.02$, $p = 0.018$, 95 % CI [-0.03, −0.003]; see Table S24).

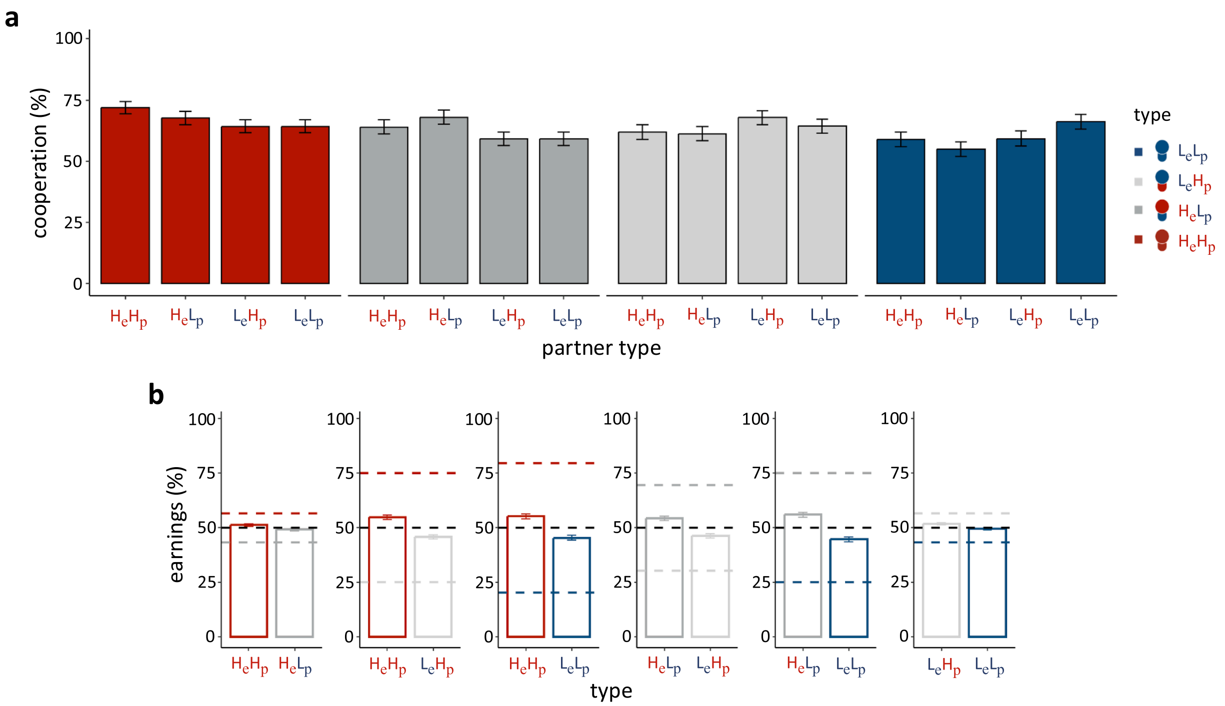

Fig. 3. Perceived norms of cooperation and preferences for public good redistribution by third-parties. a Relative cooperation (measured as average contributions to the public good as a percentage of the endowment of each participant type) considered normatively acceptable by third-parties for each type with each possible partner type. **b** Relative earnings received from the public good (measured as the average number of received units from the public good as a percentage of the total number of units in the public good), indicated as normatively acceptable by third-parties for each type with each possible different partner type. The black dotted line shows how much both types should receive if the public good would be divided equally (50%). The coloured lines show how much each type should receive if the public good would be divided according to perfect merit (i.e., types receive back what they contributed). Error bars indicate the standard error of the mean.

4. Discussion

This study showed that cooperation under free partner choice leads to segregation and can thereby increase wealth inequality (replicating Stallen et al., 2023). Curbing free partner choice could, hence, be beneficial to reduce inequality. Based on cooperation choices in a large sample, we calculated how the combination of type-creation and partner-matching interventions would reduce inequality or increase efficiency. This revealed an equality-efficiency trade-off (Okun, 2015): While the creation of pairs consisting of individuals who differed in their endowment and productivity decreases inequality, the creation of pairs consisting of individuals who had similar endowments and productivities increases overall efficiency.

We then examined how participants with the power to create types and pairs *actually* navigated this equality-efficiency trade-off. On the one hand, many people are averse to inequality and prefer to reduce inequality (Dawes et al., 2007; Fehr et al., 2008; Fehr & Schmidt, 1999; Kiatpongsan & Norton, 2014; Norton & Ariely, 2011), especially when differences between individuals are based on luck (like in the current study) (Cappelen et al., 2007, 2011; Starmans et al., 2017; Wright & Aldama, 2023). On the other hand, some level of inequality needs to be accepted when collective efficiency is the goal (Engelmann & Strobel, 2004; Hoenig et al., 2024; Komorita & Parks, 1995). We found that uninvolved third-parties were most likely to make choices aiming to reduce wealth inequality. While decision-makers who were assigned a low endowment and low productivity in our artificially generated 'society' were also more likely to impose pairings that would reduce inequality, those who were assigned a high endowment and high productivity tended to prioritize efficiency maximization and segregation. This is in line with previous research showing that, while uninvolved persons with no conflict of interest are often concerned with fairness (Mitchell et al., 1993), individuals who have a 'stake in the game' are driven by other motivations such as self-interest (Gordon-Hecker et al., 2017; Reuben & Riedl, 2013).

Taken together, this illustrates that an individual's tolerance and preference for segregation, to some degree, depends on their own position (Durante et al., 2014). The advantaged benefit from and prefer segregation and the results of free partner choice (i.e., by flocking together and only cooperating with each other they maximize their own wealth), while the disadvantaged benefit from and prefer restrictions on free partner choice and (imposed) mixed pairings that increase equality. Such disagreements between the advantaged and disadvantaged on how to govern free partner choice in 'unequal societies' may lead to intragroup fault-lines and conflict that, outside of the lab, may also underlie disagreements on policies such as free school choice or social housing initiatives. Based on our results, we would expect that in situations with higher structural inequality (i.e., in a world consisting of high/high and low/low types rather than low/high and high/low types), conflicts and preference polarizations also increase. As a result, groups may have difficulties to establish effective policies and shared fairness standards.

In our study, social preferences did not impact partner preferences among individuals most capable of reducing inequality (i.e., those assigned a high endowment and high productivity). This illustrates that social preferences are conditional and not solely driven by altruism (Fehr & Charness, 2023; Thöni & Volk, 2018). That is, the degree to which individuals exhibit prosocial behaviour is contingent upon the associated costs. In our experimental setup, it appears that the cost for participants who were assigned a high endowment and high productivity to engage in prosocial behaviour—such as choosing low-endowment low-productivity others—might have been too high, deterring even individuals with prosocial preferences from such choices. However, while partner preferences did not depend on social preferences, participants' decisions on how to intervene in the formation of partner pairings were impacted by social preferences. Specifically, participants' social preferences did impact their willingness to force individuals to mix: While participants with more *prosocial* preferences were inclined to force advantaged individuals to form pairs with less advantaged individuals, individuals with more *selfish* preferences allowed for segregation.

Findings resonate with previous research in which individuals could choose between multiple public goods, each with distinct attributes – one attractive for its efficiency in delivering relatively high returns, but less appealing due to its unequal distribution of returns among group members and vice versa (Hoenig et al., 2024). Notably, in the study by Hoenig et al. (2024), individuals tended to cooperate more with an efficient, yet unequal public good when they personally gained the most from such inequality (see also Gross et al., 2020; Gross & Böhm, 2020). Conversely, the disadvantaged showed a preference for public goods with equal returns, even though it had a lower efficiency. Furthermore, when decision-makers are asked to punish selfish behaviours and to reward cooperative behaviours, their decisions can also be self-serving (Chen et al., 2023). Indeed, decision-makers often face a dilemma between making legitimate choices and acting in one's self-interest (Gino et al., 2016; Rodriguez-Lara & Moreno-Garrido, 2012).

Interestingly, we found that third-parties believed that advantaged individuals should receive more units from the public good when they were paired with disadvantaged individuals, even when both types contributed their full endowment to the public good and types were assigned randomly. A possible reason for these normative beliefs could be that those who made a higher contribution to the collective outcome were perceived as deserving more. However, this also means that structural differences in the ability to cooperate were, at least to some degree, treated by third-parties as if they were resulting from merit, even though they were completely based on luck in our setting. The misperception also partly explains the observed efficiency-equality trade-off: In our public goods game, each individual receives an equal share of the resources (thereby also introducing the free-rider problem). In mixed pairs, receiving only half of the public good is deemed normatively unacceptable for advantaged individuals. Therefore, advantaged individuals reduce their cooperation when paired with disadvantaged individuals (note that this was considered normatively acceptable by our third-parties). Conversely, in pairs consisting of similar types, receiving half of the public good was perceived normatively acceptable. Therefore, individuals increase their cooperation when paired with partners who had similar endowments and productivities. This results in significant inequality across unequal pairs, fuelling the efficiency-inequality trade-off: Under these norms of cooperation, achieving both optimal efficiency and equality at the same time becomes unattainable.

These findings hold important implications for real-world interventions aimed at addressing inequality, considering that decision-makers are seldom entirely uninvolved, particularly in matters concerning cooperation and structural differences in the ability to cooperate. As a case in point, decision-makers in influential positions may be more likely to possess high endowments and productivity levels rather than lower ones, and may therefore make decisions (e.g., from creating pairs to designing public goods to be equal or efficient) that prioritize overall efficiency, potentially disregarding its distribution. Consequently, interventions designed by these parties might (inadvertently) favour their own position. Our findings suggest that third-parties or disadvantaged decision-makers would be most effective in designing interventions if the aim is to reduce (ex-post) wealth inequality.

As interventions typically not try to influence behaviour that occurs on a one-time basis, but instead involves repeated interactions, future research is needed to study the long-term impact of interventions on cooperation and partner choice 'among unequals'. When decision-makers force individuals to interact with others who differ in their ability to cooperate, individuals may change their cooperation level based on their support of the intervention. Indeed, previous research has shown that assigning partners 'top-down' can decrease cooperation over time (Stallen et al., 2023). Furthermore, endogenously chosen institutions can lead to higher levels of cooperation than institutions that are exogenously imposed upon individuals (Foster & Putterman, 2010; Gallier, 2020). Therefore, the dynamics of ongoing interactions in response to interventions on free partner choice requires further investigation to study their sustained effects on cooperation.

A limitation of our study is that the order in which participants made decisions was fixed and not randomized, across participants. We chose to fix the decision order to increase participants' understanding of the rules of the game and possible consequences of their decisions. That is, participants began by creating types, followed by forming pairs, then indicated their contributions to the public good, and finally indicated their preferences regarding the redistribution of wealth from the public good. The downside of a fixed decision order is that this does not allow us to control for, or investigate, potential order effects. Additionally, structural participant characteristics, such as high/low endowment and high/low productivity factors, were fixed in our study, with the high endowment being 75, the low endowment being 25, the high productivity factor being 1.7, and the low productivity factor being 1.3. We recognize that varying the size of endowments and productivity factors could alter the dynamics of partner choice and cooperation. However, the impact of the magnitude of these characteristics on partner preferences remains largely unexplored, presenting an interesting avenue for future research. Finally, in real-world work teams, individuals may differ in their ability to contribute and cooperate over time. For example, someone might be an independent 'high performer' in one task but needs to rely more on team members in another. Similarly, in societal contexts, people can ascend or descend the societal ladder due to economic shifts. Such flexibility or changes in the ability to cooperate within individuals may allow for tacit coordination of cooperation and achievement of fairness across interactions. Future studies could experimentally examine settings where structural characteristics are not fixed but may change from round to round.

5. Conclusion

Individuals can often choose with whom they want to cooperate or not. However, when individuals not only differ in their willingness but also ability to cooperate, partner choice leads to segregation and increases resource inequality. To mitigate such inequality, curbing free partner choice can help, but can also lead to an equality-efficiency trade-off. When decision-makers do not intervene in partner choice and allow for segregation, collective efficiency is highest. When they force individuals with different cooperation abilities to interact with each other, wealth inequality is reduced at the price of reduced collective efficiency. This is because the advantaged reduce their cooperation when paired with the disadvantaged, which also third-parties seem to perceive as acceptable.

Uninvolved third-parties still prioritized inequality reduction over efficiency maximization. However, involved decision-makers navigated the equality-efficiency trade-off rather self-servingly: Advantaged decision-makers prefer to maximize efficiency for their own benefit, recreating the situation observed under free partner choice, while disadvantaged decision-makers prefer restrictions of partner choice to reduce inequality. Results not only show how free partner choice and its restrictions influence cooperation 'among unequals', but also how, under economic inequality, preferences for governing free (partner) choice can polarize, creating the foundation for political conflict, diverging fairness standards within groups, and opposing demands for freedom versus regulation.

Funding

This work was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data of our experiment are openly available in an OSF repository (https://doi.org/10.17605/OSF.IO/3WRSU).

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.joep.2024.102758.

References

Almås, I., Cappelen, A. W., Sørensen, E.Ø., & Tungodden, B. (2022). Global evidence on the selfish rich inequality hypothesis. *Proceedings of the National Academy of Sciences, 119*(3). e2109690119.

André, J., & Baumard, N. (2011). The evolution of fairness in a biological market. *Evolution, 65*(5), 1447–1456.

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior, 34*(3), 164–175.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology, 7*, 33–38.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences, 274*(1610), 749–753.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences, 36*(1), 59–78.

Bayer, P., & McMillan, R. (2012). Tiebout sorting and neighborhood stratification. *Journal of Public Economics, 96*(11–12), 1129–1143.

Bhavnani, R., Donnay, K., Miodownik, D., Mor, M., & Helbing, D. (2014). Group segregation and urban violence. *American Journal of Political Science, 58*(1), 226–245.

Boterman, W., Musterd, S., Pacchi, C., & Ranci, C. (2019). School segregation in contemporary cities: Socio-spatial dynamics, institutional context and urban outcomes. *Urban Studies, 56*(15), 3055–3073.

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics, 14*, 375–398.

Bruch, E. E. (2014). How population structure shapes neighborhood segregation. *American Journal of Sociology, 119*(5), 1221–1278.

Cappelen, A. W., Hole, A. D., Sørensen, E.Ø., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review, 97*(3), 818–827.

Cappelen, A. W., Hole, A. D., Sørensen, E.Ø., & Tungodden, B. (2011). The importance of moral reflection and self-reported data in a dictator game with production. *Social Choice and Welfare, 36*(1), 105–120.

Chen, J., Lian, Z., & Zheng, J. (2023). Self-serving reward and punishment: Evidence from the laboratory. *Scientific Reports, 13*(1), 13997.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics, 129*(4), 1553–1623.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature, 446*(7137), 794–796.

Durante, R., Putterman, L., & Van der Weele, J. (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association, 12*(4), 1059–1086.

Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review, 94*(4), 857–869.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079–1083.

Fehr, E., & Charness, G. (2023). Social preferences: fundamental characteristics and economic consequences. CESifo Working Paper No. 10488.

Fehr, E., & Fischbacher, U. (2002). Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives. *The Economic Journal, 112*(478), C1–C33.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63–87.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics, 114*(3), 817–868.

Foster, A., & Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review, 100*(5), 2205–2229.

Gallier, C. (2020). Democracy and compliance in public goods games. *European Economic Review, 121*, 103346.

Gino, F., Norton, M. I., & Weber, R. A. (2016). Motivated Bayesians: Feeling moral while acting egoistically. *Journal of Economic Perspectives, 30*(3), 189–212.

Gordon-Hecker, T., Rosensaft-Eshel, D., Pittarello, A., Shalvi, S., & Bereby-Meyer, Y. (2017). Not taking responsibility: Equity trumps efficiency in allocation decisions. *Journal of Experimental Psychology: General, 146*(6), 771–775.

Gross, J., & Böhm, R. (2020). Voluntary restrictions on self-reliance increase cooperation and mitigate wealth inequality. *Proceedings of the National Academy of Sciences of the United States of America, 117*(46), 29202–29211.

Gross, J., Veistola, S., De Dreu, C. K. W., & Van Dijk, E. (2020). Self-reliance crowds out group cooperation and increases wealth inequality. *Nature Communications, 11*(1), 1–9.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2014). On cooperation in open communities. *Journal of Public Economics, 120*, 220–230.

Hardin, G. (1968). The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. *Science, 162*(3859), 1243–1428.

Hauser, O. P., Hilbe, C., Chatterjee, K., & Nowak, M. A. (2019). Social dilemmas among unequals. *Nature, 572*(7770), 524–527.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902.

Henrich, J., & Boyd, R. (2008). Division of labor, economic specialization, and the evolution of social stratification. *Current Anthropology, 49*(4), 715–724.

Hoenig, L. C., Pliskin, R., & De Dreu, C. K. W. (2024). Equality and efficiency shape cooperation in multiple-public-goods provision problems. *Journal of Experimental Psychology: General, 153*(5), 1236–1256.

Kiatpongsan, S., & Norton, M. I. (2014). How much (more) should CEOs make? A universal desire for more equal pay. *Perspectives on Psychological Science, 9*(6), 587–593.

Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology, 46*(1), 183–207.

Melamed, D., Simpson, B., Montgomery, B., & Patel, V. (2022). Inequality and cooperation in social networks. *Scientific Reports, 12*(1), 1–10.

Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis, 79*(4), 606–623.

Mitchell, G., Tetlock, P. E., Mellers, B. A., & Ordonez, L. D. (1993). Judgments of social justice: Compromises between equality and efficiency. *Journal of Personality and Social Psychology, 65*(4), 629–639.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771–781.

Norton, M. I., & Ariely, D. (2011). Building a better America—one wealth quintile at a time. *Perspectives on Psychological Science, 6*(1), 9–12.

Okun, A. M. (2015). *Equality and efficiency: The big tradeoff*. Brookings Institution Press.

Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action: Presidential address, american political science association, 1997. *American Political Science Review, 92*(1), 1–22.

Pfeiffer, T., & Nowak, M. A. (2006). All in the game. *Nature, 441*(7093), 583–584.

Raihani, N. J., & Barclay, P. (2016). Exploring the trade-off between quality and fairness in human partner choice. *Royal Society Open Science, 3*(11), 160510.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology, 116*(4), 1092–1153.

Reuben, E., & Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior, 77*(1), 122–137.

Rodriguez-Lara, I., & Moreno-Garrido, L. (2012). Self-interest and fairness: Self-serving choices of justice principles. *Experimental Economics, 15*(1), 158–175.

Roemer, J. E., & Trannoy, A. (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature, 54*(4), 1288–1332.

Stallen, M., Snijder, L. L., Gross, J., Hilbert, L. P., & De Dreu, C. K. W. (2023). Partner choice and cooperation in social dilemmas can increase resource inequality. *Nature Communications, 14*(1), 6432.
Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour, 1*(4), 1–7.
Thöni, C., & Volk, S. (2018). Conditional cooperation: Review and refinement. *Economics Letters, 171*, 37–40.
Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy, 64*(5), 416–424.
van Dijk, E., & De Dreu, C. K. W. (2021). Experimental games and social decision making. *Annual Review of Psychology, 72*, 415–438.
van Winden, F. (2023). The informational affective tie mechanism: On the role of uncertainty, context, and attention in caring. *Journal of Economic Psychology, 97*, 102625.
Wright, R., & Aldama, A. (2023). Not all luck is created equal: Sources of income inequality and willingness to redistribute. *Journal of Economic Psychology, 97*, 102637.