I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16179

When Fairness Is Not Enough: The Disproportionate Contributions of the Poor in a Collective Action Problem

Eugene Malthouse
Charlie Pilgrim
Thomas Hills
Daniel Sgroi

MAY 2023

IZA – Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16179

When Fairness Is Not Enough: The Disproportionate Contributions of the Poor in a Collective Action Problem

Eugene Malthouse
University of Warwick

Charlie Pilgrim
University College London

Thomas Hills
University of Warwick

Daniel Sgroi
University of Warwick and IZA

MAY 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

When Fairness Is Not Enough: The Disproportionate Contributions of the Poor in a Collective Action Problem

Many of our most pressing challenges, from combating climate change to dealing with pandemics, are collective action problems: situations in which individual and collective interests conflict with each other. In such situations people face a dilemma about making individually costly but collectively beneficial contributions to the common good. Understanding which factors influence people's willingness to make these contributions is vital for the design of policies and institutions that support the attainment of collective goals. In this study we investigate how inequalities, and different causes of inequalities, impact individual-level behaviour and group-level outcomes. First, we find that what people judged to be fair was not enough to solve the collective action problem: if they acted according to what they thought was fair, they would collectively fail. Second, the level of wealth (rich vs. poor) altered what was judged to be a fair contribution to the public good more than the cause of wealth (merit vs. luck vs. uncertain). Contributions during the game reflected these fairness judgements, with poorer individuals consistently contributing a higher proportion of their wealth than richer participants, which further increased inequality – particularly in successful groups. Finally, the cause of one's wealth was largely irrelevant, mattering most only when it was uncertain, as opposed to resulting from merit or luck. We discuss implications for policymakers and international climate change negotiations.

JEL Classification: C92, D91, D63

Keywords: public goods, collective action, cooperation, meritocracy

Corresponding author:
Eugene Malthouse
University of Warwick
Coventry
CV4 7AL
United Kingdom

E-mail: eugene.malthouse@warwick.ac.uk

Introduction

Humans are a highly cooperative species (Henrich & Henrich, 2007; Bowles & Gintis, 2011). Across the world people engage in collective action with others every day (Nowak & Highfield, 2011). Sometimes just a handful of individuals are involved, for example when a group of researchers come together to conduct a study. Other times many millions of individuals are involved, for example when citizens vote or work together to reduce global warming.

Regardless of how many individuals are involved, human cooperation is vulnerable to free riding (Olson, 1965). This term describes the temptation for each person to "free himself of the trouble and expense, and...lay the whole burden on others" (Hume, 1740, p. 590). Just as a researcher might be tempted to leave a tedious task to their co-authors, organisations might be tempted to avoid taking the costly actions required to reduce their greenhouse gas emissions. The problem is that if every individual involved succumbs to this temptation, the group will inevitably fail to achieve its collective goal – an outcome known as a 'tragedy of the commons' (Hardin, 1968).

There are countless factors that may influence one's temptation to free ride, but one is particularly important and relevant to this study: the behaviour of others (Stroebe & Frey, 1982; Ledyard, 1995; Kopelman et al., 2002; Ostrom, 2008; Ostrom, 2010). Much research has shown that a significant proportion of people generally act as conditional cooperators – contributing while others contribute and free riding when others free ride (Kelley & Stahelski, 1970; Sugden, 1984; Keser & van Winden, 2000; Fischbacher et al., 2001; Croson et al., 2005; Gächter, 2006; Fischbacher & Gächter, 2010).

Cooperating on the condition that others are cooperating requires us to judge what constitutes a fair contribution to the joint effort (Rabin, 1993; Fehr & Schmidt, 1999; Elster, 2006; Van Segbroeck et al., 2012). When contributions are financial, two factors have been proposed as important. The first is a person's total wealth, because people intuitively judge financial contributions in proportional rather than absolute terms (Laming, 1984; Stewart et al., 2005; Stewart et al., 2006). In the UK, for example, The Sunday Times Giving List ranks donors according to the proportion of their total wealth donated to charity. This is why the footballer Marcus Rashford topped the list in 2021, despite giving less than a tenth of the sum donated by the runner-up Lord Sainsbury (£229m) (The Sunday Times, 2021).

The second factor proposed to be important is the cause of one's wealth. In recent years, many researchers (e.g., Piketty, 1995; Konow, 2000; Fong, 2001; Alesina et al., 2001; Alesina & Glaeser, 2004; Frank, 2016; Alan & Ertac, 2017; Alesina et al., 2018; Markovits, 2019; Piketty, 2020; Sandel, 2020; Koo et al., 2022) have noted the importance of luck and merit in society. They have generally shown that belief in luck rather than merit as the primary determinant of life outcomes generally correlates with stronger individual preferences and public policies in favour of wealth redistribution. One explanation for this is that people are generally more accepting of inequalities arising from merit than from good fortune alone (Adams, 1965; Walster et al., 1976; Starmans et al., 2017) – particularly in Western capitalist cultures (Son Hing et al., 2011; Gonzalez et al., 2022). According to this logic, many people might think it fair for lottery winners to contribute more to the common good than self-made millionaires. But it is unclear whether this logic generally holds and influences behaviour in collective action problems.

In reality, wealth is almost always determined by an incalculable combination of luck and merit. Yet there is no consensus in the literature about the impact of this uncertainty on people's preferences for wealth redistribution. On the one hand, uncertainty regarding the cause of wealth inequalities has been shown to generate an 'egalitarian pull' on people with meritocratic preferences (Cappelen et al., 2022). One explanation for this is that unlike luck and merit, uncertainty does not provide an easily justifiable reason to deviate from equality (Samuelson & Allison, 1994). In other words, luck and merit may be seen as legitimate causes of wealth inequality, whereas an uncertain mixture of luck and merit is not – and people in such a world may seek to redistribute wealth more evenly. On the other hand, uncertainty may bring out self-serving interpretations of wealth whereby people attribute successes to their own efforts and failures to external forces (Miller & Ross, 1975). If this is the case, richer and poorer individuals may become more and less accepting of wealth inequalities, respectively.

In the present study we empirically investigated the impact of different levels and causes of wealth inequality on beliefs about fairness, individual-level contributions and group-level outcomes in a collective action problem. We did this by adapting a public good game known as the ‘collective risk social dilemma’ (introduced by Milinski et al., 2008) in two ways. First, we introduced different levels of wealth by randomly assigning participants to groups of four made up of two richer participants who received an endowment of £20 and two poorer participants who received an endowment of £10. Second, we introduced different causes of wealth by randomly assigning groups to one of three treatment conditions: one in which participants’ endowments were caused by merit (the *merit* treatment); one in which they were caused by luck (the *luck* treatment); and one in which they were caused either by merit or luck, but participants did not know which (the *uncertain* treatment). Below, we explain our adaptation of the collective risk social dilemma in more detail, as well as our main research questions and hypotheses.

Table 1. *Participant Types*

The table summarises the different levels (rich vs. poor) and causes (merit vs. luck vs. uncertain) of wealth in our experiment design, and the resulting labels for different types of participants. Treatments were assigned at the group level such that a group in the merit treatment consisted of two Deserving Rich and two Deserving Poor players; a group in the luck treatment consisted of two Lucky Rich and two Unlucky Poor players; and a group in the uncertain treatment consisted of two Uncertain Rich and two Uncertain Poor players.

	Merit Treatment	**Luck Treatment**	**Uncertain Treatment**
Rich (£20)	Deserving Rich	Lucky Rich	Uncertain Rich
Poor (£10)	Deserving Poor	Unlucky Poor	Uncertain Poor

The collective risk social dilemma

The collective risk social dilemma is a specific type of public good game. Public good games have been used for decades to investigate behaviour in situations where individual goals conflict with group goals (e.g., Rapoport, 1988; Sandler, 1992; Ledyard et al., 1995; Camerer, 2003; Van Lange et al., 2013). In brief, participants are given an initial endowment and must decide how much of this to contribute towards a group target in 10 successive rounds. If together they achieve the target sum within the 10 rounds, all players take home the remainder of their endowment (i.e., all funds not contributed to the group account). If the group fails to achieve its target, players face the prospect of losing their remaining endowment. This collective risk creates a social dilemma: the more an individual contributes to the target, the more likely her group is to succeed, but the less she stands to take home at the end of the game (see Figure 1). This game format is typically described as a threshold public good game, since the group either succeeds (by meeting the threshold) or fails (by falling short). It was designed to represent similar real-world collective action problems such as climate change, which can be understood in terms of threshold dynamics in the sense that we either succeed in limiting global warming to 1.5°C above pre-industrial levels or we fail (United Nations, 2015). With the collective risk social dilemma we asked the following three research questions:

1. Is what people perceive as fair sufficient to solve collective action problems?
2. How does what people perceive as fair, and how much they are willing to contribute to a public good, depend on the level and cause of their wealth?
3. If what people perceive as fair is insufficient to solve the problem, under what conditions do groups still manage to succeed?

Before we detail the experiment we used to answer these questions, we first introduce what is known from prior research and our hypotheses associated with each question.

Figure 1. *Group outcomes and player payoffs*

This figure illustrates how each participant's final payoff depends on both their individual contribution decisions during the game and their group's outcome. Each group is made up of two richer participants (Players 1 and 3 here) starting the game with £20 and two poorer participants (Players 2 and 4 here) starting with £10. Players can contribute either £0, £0.75, or £1.50 in each round. If the group succeeds in achieving its target sum of £30 within 10 rounds, players take home what is left of their initial endowment. If the group fails to achieve this target sum, all players face a 50% chance of losing their remaining funds.

Is what people perceive as fair sufficient to solve collective action problems?

In many real-world collective action problems, what is judged to be fair may be insufficient to achieve a collective goal. For example, since the Paris Agreement (United Nations, 2015) countries have outlined how they intend to contribute to the reduction of global emissions via Nationally Determined Contributions (NDCs), which are based partly on what they judge to be fair (Davide et al., 2017). However, according to a recent report from the United Nations Framework Convention on Climate Change (2021) if all 193 governments fulfilled their NDC targets then global greenhouse gas emissions would actually *increase* by 13.7% by 2030 – falling far short of the estimated 45% reduction required to limit global warming to 1.5°C. Another example of fairness not being enough can be found in the European Union Common Fisheries Policy: between 2001 and 2015 the European Council set national quotas that exceeded scientific advice regarding sustainability by an average of 20% per year (Carpenter et al., 2016).

We anticipated that participants in our experiment might similarly struggle to recognise that what is fair might not be enough (Hypothesis 1). In other words, we predicted that participants' judgements about fair contributions would be insufficient to solve the collective action problem.

How does what people perceive as fair, and how much they are willing to contribute to a public good, depend on the level and cause of one's wealth?

One reason why fairness judgements may be insufficient relates to people's tendency to hold self-serving beliefs about what is fair (e.g., Baumeister, 1982; Joireman et al., 1984; Hine & Gifford, 1996; Wade-Benzoni et al., 1996; Babcock & Loewenstein, 1997; Diekmann et al., 1997; Bernard et al., 2012). We therefore anticipated that richer participants would be more likely than poorer participants to judge it fair that richer players contribute a lower proportion of their wealth than poorer players. And, conversely, poorer participants would be more likely than richer participants to judge it fair that poorer players contribute a lower proportion of their wealth than richer players (Hypothesis 2).

In turn, we expected participants' contributions towards the group target within the game to reflect these self-serving fairness judgements. Based on prior research (e.g., Marwell & Ames, 1979; Van Dijk & Wilke, 1993, 1994; Chan et al., 1996; De Cremer, 2007; Tavoni et al., 2011; Van Lange et al., 2013; Vasconcelos et al., 2014; Heap et al., 2016; Vicens et al., 2018; Martinangelia & Martinsson, 2020) we predicted that richer participants would contribute more than poorer participants in absolute terms; but less in proportional terms (Hypothesis 3).

Our remaining hypotheses relate to the effect of different causes of wealth (merit vs. luck vs. uncertainty). Based on the research discussed above we anticipate two possibilities: a) merit is seen as the primary indicator of deserved wealth; or b) both merit and luck (but not uncertainty) are seen as justifiable criteria for wealth inequalities. We believed that (a) was more likely, and therefore predicted that poorer participants in the merit treatment (versus those in the luck and uncertain treatments) would be expected to contribute a higher proportion of their wealth (Hypothesis 4).

In turn, we anticipated that participants' actual contributions towards the group target during the game would reflect these expectations – with the Deserving Rich contributing a lower proportion of their wealth than the Uncertain and Lucky Rich (Hypothesis 5).

Lastly, we anticipated that these predicted differences in contributions between treatments would have a knock-on effect on group outcomes. We therefore predicted that groups in our luck treatment would achieve the target of £30 with a higher success rate than groups in the uncertain and merit treatments (Hypothesis 6).

Together, these made up our main hypotheses – all of which are summarised in the table below (H3, H4, H5, and H6 were formally pre-registered, available at https://osf.io/4expt). If our prediction that fairness judgements would be insufficient turned out to be accurate, participants would not contribute enough to achieve the group target. In the next section, therefore, we identify certain factors that might explain the difference between group success and failure.

Table 2. *Main Hypotheses*

The table summarises our main hypotheses, including the theme and our predictions for each.

Hypothesis	Theme	Prediction
1	Fairness: Insufficiency	Participants' will on average judge it fair that richer and poorer players contribute less than 50% of their wealth;
2	Fairness: Level of Wealth (Rich vs. Poor)	Richer (poorer) participants will judge it fair that they contribute a lower proportion of their wealth than poorer (richer) participants;
3	Contributions: Level of Wealth (Rich vs. Poor)	Richer participants will contribute more than poorer participants in absolute terms, but less in proportional terms;
4	Fairness: Cause of Wealth (Merit vs. Luck vs. Uncertain)	Individuals in the merit treatment will expect the poor to contribute a higher proportion of their wealth;
5	Contributions: Cause of Wealth (Merit vs. Luck vs. Uncertain)	The Deserving Rich will contribute a lower proportion of their wealth than the Uncertain and the Lucky Rich;
6	Group Outcomes	Luck-based groups would be more successful than merit-based and uncertain groups

If what people perceive as fair is insufficient to solve the problem, under what conditions do groups still manage to succeed?

The ability of groups to succeed despite insufficient views of fairness (H1) will depend on whether certain individuals step up to fill the gap between what is fair and what is required for success. To investigate whether richer or poorer participants stepped up in this way to help their groups succeed, we compared their contributions in successful and unsuccessful groups. If it was the latter then wealth inequalities within groups would increase – particularly within groups that were successful. It remains unclear how richer and poorer participants in luck-based, merit-based, and uncertain groups might respond to such a development because, to our knowledge, the intersection of cause of wealth and wealth inequality has not been examined in a collective risk game.

We also identified two other factors that might help to explain group success: participants' contributions in the first round and their response to their group not contributing at the rate required to achieve the target (£3 per round). We explain how we intend to test these, along with all of our other hypotheses, in the section that follows.

Methods

Participants

We sourced a total of 240 participants via Prolific Academic and Mechanical Turk. We arrived at this sample size via power calculations based on effect sizes detected in previous similar studies (see rationale in our pre-registration: https://osf.io/4expt). We collected data initially from 124 participants in April 2021 and (after peer review) from an additional 116 participants in May 2022. We generally collected data from four groups at a time, depending on participant availability, by publishing the study online and accepting participation on a first-come-first-served basis. We originally planned to recruit participants roughly evenly from Prolific and MTurk to avoid any biases associated with either pool (Paolacci et al., 2010; Palan & Schitter, 2018; Litman et al., 2021). However, grouping people up was much more straightforward on Prolific due to greater participant availability and so our final sample consisted of 188 participants from Prolific and 52 participants from MTurk. All results reported below reflect pooled Prolific Academic and Mechanical Turk data; in the supplementary materials we separate results from these two sample populations and note differences between the two (see Figure S3).

Regardless of platform, all participants were over the age of 18 and entered their age range and gender at the start of the experiment: 34% were aged 18-24; 43% were aged 25-34; 14% were aged 35-44; 7% were aged 45-54; and 3% were aged 55+; while 44% of all participants were female, 55% were male and 1% identified as non-binary. Participants received pro rata payment of £7.50 per hour, as recommended by Prolific Academic. In addition, they had the opportunity to earn a bonus payment depending on the outcome of the experiment (M = £6.69, SD = £4.70). This rate (and the whole experiment) was approved by the University of Warwick's Psychology Department Research Ethics Committee. The experiment was programmed using oTree, a platform that enables researchers to build and run online experiments (Chen et al., 2016).

Experiment Design

After participants read an information sheet and consented to the terms of the experiment (see supplementary materials) they were randomly assigned to groups of four, which in turn were randomly assigned to the merit, luck, or uncertain treatments. Every group was made up of two richer participants, who started the game with an endowment of £20, and two poorer participants who started the game with £10. The level of inequality was therefore identical in each group, but the cause of these inequalities differed between our three treatments:

- In the merit treatment participants' endowments were determined by their performance in the effort task. In each group, the two highest-scoring participants received £20 to start the game and the two lowest-scoring participants received just £10. All participants were explicitly informed about this meritocratic allocation both before and after they completed the task.

- In the luck treatment participants' endowments were determined randomly by a lottery. This meant that their effort task performance had no bearing on whether they started the game with £20 or £10, and we told participants that this was the case both before and after they completed the task. To incentivise completion of the task we gave a £1 bonus payment to the highest-scoring member of each group at the very end of the experiment.
- In the uncertain treatment participants did not know the true determinant of their wealth. In each group two randomly chosen participants' endowments were determined by their performance in the effort task, with the higher-scoring of these two receiving £20 and the lower-scoring player receiving £10. For the other two group members endowments were determined by a lottery, with the winner receiving a £20 endowment and the loser £10. Participants in these groups were told that their endowment was determined either by their task performance or by the lottery – but they were not told which.

The structure of the experiment was exactly the same for every participant. It began with an effort task previously used by Niederle and Vesterlund (2007) and Oswald et al. (2015). This entailed adding up sequences of five random two-digit numbers for five minutes (i.e., 16 + 82 + 51 + 55 + 26 = ?) with participants receiving one point per correct answer. While we asked participants to refrain from using a calculator, we recognised that this request was unlikely to be followed in an online environment. However, as stated in our pre-registration this was not a primary concern because we ultimately wanted participants to believe that their degree of effort was correlated with their rewards. This applies if some people are better at mental arithmetic than others, and even if some people are using a calculator (since using a calculator for five minutes still represents an effortful activity).

Once all participants in a group had completed this task we told them whether they would start the game with £20 or £10 and confirmed whether this was determined by merit, luck, or one of the two (depending on their treatment). We then explained the rules of the collective risk social dilemma game using illustrations similar to Figure 1. We made it clear what would happen if the group succeeded (all players would retain all funds not contributed to the group account) and if it failed (all players would face a 50% chance of losing all funds not contributed). We then tested their understanding with three comprehension questions, the first of which asked how much each player would have to contribute on average for the group to achieve its target (£7.50). We then asked them what, in their opinion, they considered to be a fair total contribution towards the group target from richer and poorer players (see supplementary materials for full pre-game questionnaire). Participants had to answer the comprehension questions correctly before they could proceed to the game, which helped to ensure that their responses to the fairness questions reflected their opinions rather than their understanding of the game. The first round of the game began after every participant in the group had completed these steps.

The collective risk social dilemma was played over 10 rounds. At the start of each round participants were asked how much of their endowment (£0 / £0.75 / £1.50) they would like to contribute towards the group target of £30. We gave participants three contribution options, following Milinski et al. (2008), mainly because it enabled participants to quickly estimate what others had contributed at the end of each round. We set the target at £30 because group success would require each group member to contribute half of their wealth on average. And we set the probability of losing the remainder of one's endowment in the event of group failure at 50% because it meant that players in each group faced the same expected earnings whether they chose to free ride (i.e., contribute nothing to a failing group) or all cooperate (i.e., give half of their endowment). As an example, a poorer participant would, in expectation, stand to take home £10 x 50% = £5 by free-riding and £10 - £5 = £5 by contributing 50% of their wealth, providing others in the group did the same. At the start of each round we showed participants the round number, how much remained of their endowment, how much the group had contributed in total so far, and how much more the group needed to contribute in order to achieve its target. At the end of the game participants were told their group outcome and, in the event of failure, whether or not they had survived the collective risk and therefore retained the remainder of their endowment.

In summary, we created groups made up of two richer and two poorer participants and manipulated the cause of these wealth inequalities (merit vs. luck vs. uncertain). In 10 successive rounds participants decided how much to contribute to the group target. This design enabled us to test our hypotheses regarding fairness judgements, contributions, and group outcomes. We describe our statistical tests, which were carried out in R and JASP (JASP Team, 2020), in the next section.

Main Hypotheses and Statistical Tests

H1) Fairness will be insufficient: participants will on average judge it fair that richer and poorer players contribute less than 50% of their wealth

To test H1 we investigated whether participants' judgements about fair total contributions were on average less than 50% of richer and poorer players' wealth (the level required to achieve the group target). As a secondary measure, we calculated what they judged to be fair for the group as a whole to contribute by summing an individual's responses to these questions and multiplying this figure by two; and then checking the proportion of participants for whom this total was insufficient (i.e., less than the group target of £30).

H2) Richer (poorer) participants will judge it fair that they contribute a lower proportion of their wealth than poorer (richer) participants

To test H2 we first coded participants' responses to our fairness questions according to one of three fairness principles (similar to Reindl, 2022):

- Progressive: if they judged that rich players should contribute a higher proportion of their wealth than poor players;
- Equal: if they judged that rich and poor players should contribute the same proportion of their wealth; or
- Regressive: if they judged that poor players should contribute a higher proportion of their wealth than rich players.

We then conducted chi-squared tests to compare the proportion of richer and poorer participants whose responses reflected progressive and regressive principles. H2a was that a higher proportion of richer (poorer) participants' responses would reflect the regressive (progressive) principle – which would represent a marker for self-serving bias.

We also compared what richer and poorer participants actually judged to be a fair contribution from richer and poorer players. H2b was that richer participants' response to the question of how much richer players should contribute would on average be lower in proportional terms than the response from poorer participants. H2c was the exact opposite: poorer participants' response to the question of how much poorer players should contribute would on average be lower in proportional terms than the response from richer participants. We tested H2b and H2c by conducting standard and Bayesian ANOVAs to test for wealth effects on these responses.

H3) Richer participants will contribute more than poorer participants in absolute terms, but less in proportional terms

We tested H3 by comparing richer and poorer participants' mean absolute and mean proportional total contributions from live rounds (i.e., rounds in which the group target had not already been achieved) with standard and Bayesian ANOVAs. In addition, we conducted two linear multilevel models: the first with absolute contributions as the dependent variable and the second with proportional contributions as the dependent variable; and both with fixed wealth effects and random intercepts for rounds, individuals, and groups. We did this in order to take the hierarchical nature of our data into account (since rounds were nested in individuals and individuals were nested in groups) and did not use random slopes for these factors because they did not significantly improve model fit.

H4) Individuals in the merit treatment will expect the poor to contribute a higher proportion of their wealth

We tested H4 by conducting standard and Bayesian ANOVAs with treatment and wealth as independent variables and participants' responses to the questions of what would be fair for richer and poorer participants to contribute as the dependent variables. A significant treatment effect would indicate that the cause of wealth influenced participants' fairness judgements.

H5) The Deserving Rich will contribute a lower proportion of their wealth than the Uncertain and the Lucky Rich

We tested H5 by conducting standard and Bayesian ANOVAs to compare proportional contributions from richer participants between our treatments; and again, for additional robustness, by running a multilevel model with fixed wealth and treatment effects and random intercepts for rounds, individuals, and groups to take the variance from these factors into account.

H6) Luck-based groups would be more successful than merit-based and uncertain groups

We tested H5 by comparing the proportion of groups that were successful in each treatment using chi-squared tests.

Exploratory Hypotheses

In our introduction we also highlighted certain factors that might help to explain the difference between group success and failure. The first was whether richer or poorer participants contributed more than their fair share. To investigate this we compared their contributions in successful and unsuccessful groups (similar to Tavoni et al., 2011) with standard and Bayesian ANCOVAs that included proportional contributions as the dependent variable, wealth as the independent variable, and group success as a covariate. For additional robustness, we conducted a multilevel model that included proportional contribution as the dependent variable and fixed wealth and group success effects, as well as random round, individual, and group intercepts. This model enabled us to test for an interaction between wealth and group success, which would show whether richer or poorer participants stepped up to help their groups succeed.

If it ended up being poorer participants who stepped up then wealth inequality within groups would increase – particularly in successful groups. We tested this by calculating and comparing the mean Gini coefficients of successful and unsuccessful groups at the end of the game based on participants' remaining endowments (before those in unsuccessful groups faced the prospect of losing their funds) with standard and Bayesian ANOVAs.

We discussed two other factors that might be important for group success: 1) participants' contributions on the first round; and 2) their response to their group contributing less than the required rate of contribution (£3 per round). We tested the first by comparing mean first round contributions between wealth and treatment levels with standard and Bayesian ANOVAs. We tested the second by calculating the 'slack' before each round (defined as the difference between a cumulative contribution of £3 per round and the current group total) and running multilevel models on data from richer and poorer participants with their contributions as the dependent variable, fixed slack and treatment effects, and random round, individual, and group intercepts. A significant slack x treatment interaction term would indicate that richer or poorer participants in a certain treatment were responding differently to their peers in other treatments.

Transparency and openness

- Citation: all methods developed by others (e.g., oTree, JASP) are appropriately cited in the text and listed in the references section.

- Data and code transparency: anonymised processed data on which study conclusions are based, as well as reproducible computer code used for statistical analyses, are available at https://osf.io/8kn57/.
- Preregistration: the study design, hypotheses, and analysis plan were pre-registered and are available at https://osf.io/4expt.
- Materials transparency: examples of the materials described in the methods section are shown in the Supplementary Materials.

Results

Is what people perceive as fair sufficient to solve collective action problems?

Our first hypothesis (H1) concerned people's beliefs about what was fair and whether this was enough to solve the collective action problem. Before analysing participants' judgements we excluded 12 responses above £15 for the question of how much richer players should contribute because this was not practically possible (since the maximum players could contribute in each of the ten rounds was £1.50). We also excluded four responses above £10 for the question of how much poorer players should contribute (since this exceeded their endowment).

Across all three treatments participants on average judged it fair for richer participants to contribute £7.50 (37.5% of their wealth) and for poorer participants to contribute £4.04 (40.4% of their wealth). This was insufficient to solve the collective action problem – which on average required everyone to contribute 50% of their wealth (see Figure 2). Furthermore, what 42.4% of participants judged to be fair for their group to contribute was insufficient, totalling less than £30. This was higher than the proportion of people (7.6%) whose responses added up to more than £30. We found a similar pattern when including all responses: 39.6% of participants judged to be fair an amount that was insufficient, while 13.8% judged to be fair an amount that exceeded £30. On account of this consistent skew below the group threshold, and on account of the fact that these fairness judgements were elicited immediately after three comprehension questions (one of which directly asked how much participants needed to contribute on average to succeed, and all of which participants had to answer correctly in order to proceed to the game) it seems unlikely that these results represent a fundamental misunderstanding of the game. These results therefore provide support for H1.

How does what people perceive as fair, and how much they are willing to contribute to a public good, depend on the level and cause of their wealth?

H2 related to self-serving bias in fairness judgements from richer and poorer participants. As illustrated in Plot A in Figure 2, we found that after categorising participants' judgements according to one of three fairness principles (progressive vs. equal vs. regressive) a significantly higher proportion of poorer participants' responses (23.0%) were progressive compared to richer participants (7.2%; χ^2 (1) = 9.67, p = .002). And a significantly higher proportion of richer participants' responses (31.5%) were regressive compared to poorer participants (15.0%; χ^2 (1) = 7.64, p = .006). However, as illustrated in Plots B-C in Figure 2 we did not detect a significant wealth effect on participants' responses in proportional terms to the question of what would be a fair total contribution from richer players (F(1, 226) = 0.13, p = .719; BF_{01} = 6.5) or poorer players (F(1, 234) = 2.66, p = .104; BF_{01} = 2.0). These results provide strong evidence for H2a but no evidence for H2b or H2c – indicating that participants' level of wealth did influence which fairness principle their judgements reflected, but did not influence their responses significantly.

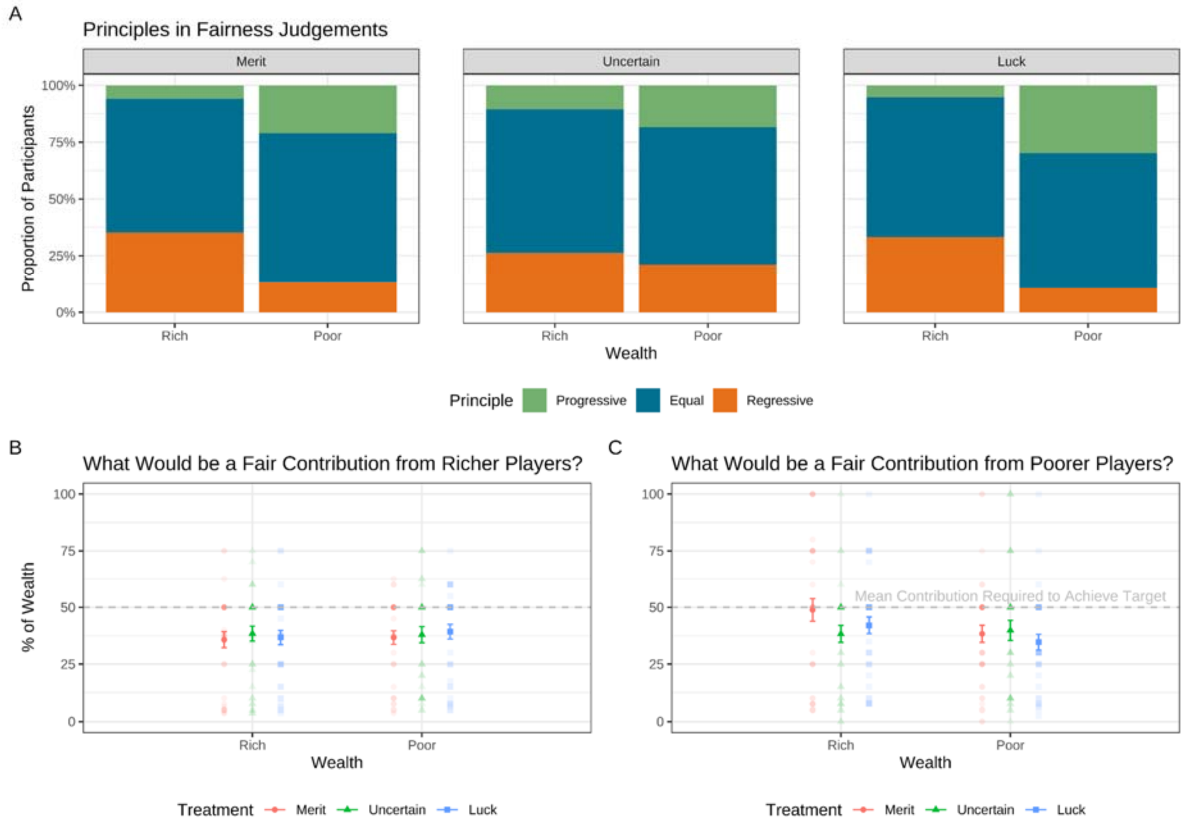


Figure 2. ***Participants' Beliefs About Fair Contributions***

Plots show participants' responses to two questions in the pre-game questionnaire: 'In your opinion, what would be a fair total contribution in £ to the group account during the game?' from players starting with £20 and players starting with £10. Plot A shows the fairness principles that participants' responses reflected across wealth and treatment levels – Progressive meant they believed poorer players should contribute a higher proportion of their wealth than richer players; Regressive meant the opposite; and Equal meant they believed players should contribute equal proportions. Plots B-C illustrate actual responses and have been converted into proportional terms as a percentage of each type of player's wealth. Plot B shows that richer players were on average expected to contribute 37.5% of their wealth while Plot C shows poorer players were expected to contribute 40.4% of theirs. In these plots, points in the background represent raw data and are slightly transparent to show overlapping responses; summary points show mean responses from richer and poorer participants across treatment conditions with bars representing the standard errors. The dashed grey line represents the average level of contribution required (50%) for groups to achieve their target.

H3 was that contributions would reflect fairness judgements: richer participants would contribute more to the group account than poorer participants in absolute terms, but less in proportional terms. Standard and Bayesian ANOVAs indicated that in absolute terms richer participants on average contributed more than poorer participants ($F(1, 238) = 89.25$, $p < .001$; $BF_{10} = 1.257e+15$). In proportional terms, however, poorer participants contributed a higher proportion of their wealth ($M = 62.3\%$) than richer participants ($M = 47.6\%$, $F(1, 238) = 26.52$, $p < .001$, $BF_{10} = 23{,}862$). Our multilevel models similarly indicated that poorer participants contributed less in absolute terms ($t(238) = -10.3$, $p < .001$) but more in proportional terms ($t(238) = 5.4$, $p < .001$). These results supported H3 and also showed that both richer and poorer participants contributed a higher proportion of their wealth than was judged by all to be fair (rich: 47.6% vs. 37.5%; poor: 62.3% vs. 40.4%) – with poorer participants doing so to a greater extent.

Our next set of hypotheses were related to the effect of different causes of wealth on fairness judgements (H4); contributions towards the group target (H5); and group success (H6). For H4, we did not detect a treatment effect on fairness judgements (see Plots B-C in Figure 2). This was equally true for the question of what would be fair for richer participants to contribute ($F(2, 222) = 0.22$, $p = .803$; $BF_{01} = 17.9$) as it was for the question of what it would be fair for poorer participants to contribute ($F(2, 230) = 0.98$, $p = .378$; $BF_{01} = 9.9$). In other words, fairness judgements were not influenced by whether inequalities had been determined by merit, luck, or one of the two.

Similarly, we did not find that the cause of wealth influenced participants' actual contributions during the game (H5). As shown in Figure 3, richer and poorer participants' absolute contributions were similar across the merit, luck, and uncertain treatments (results of standard and Bayesian ANOVAs: (rich) $F(2, 117) = 0.32$, $p = .726$; $BF_{01} = 9.7$; (poor) $F(2, 117) = 3.55$, $p = .702$; $BF_{01} = 9.4$). In fact, the outputs of both Bayesian ANOVAs indicated that these differences were around 10 times more likely to be explained by the null hypothesis. Equally, our multilevel model did not detect a significant treatment effect (see Table S1 for full model output and Figure S5 for model predictions across treatment and wealth levels in Supplementary Materials). These results provide strong evidence that the cause of wealth was largely irrelevant to the poor contributing substantially more than their fair share.

In line with this finding, we did not find that different causes of wealth resulted in significantly different outcomes at the group level (H6). While uncertain groups had a success rate of 90%, merit-based and luck-based groups achieved the target 75% of the time (χ^2 (2) = 1.88, p = .392, BF_{01} = 17.4; see Plot A in Figure 4).

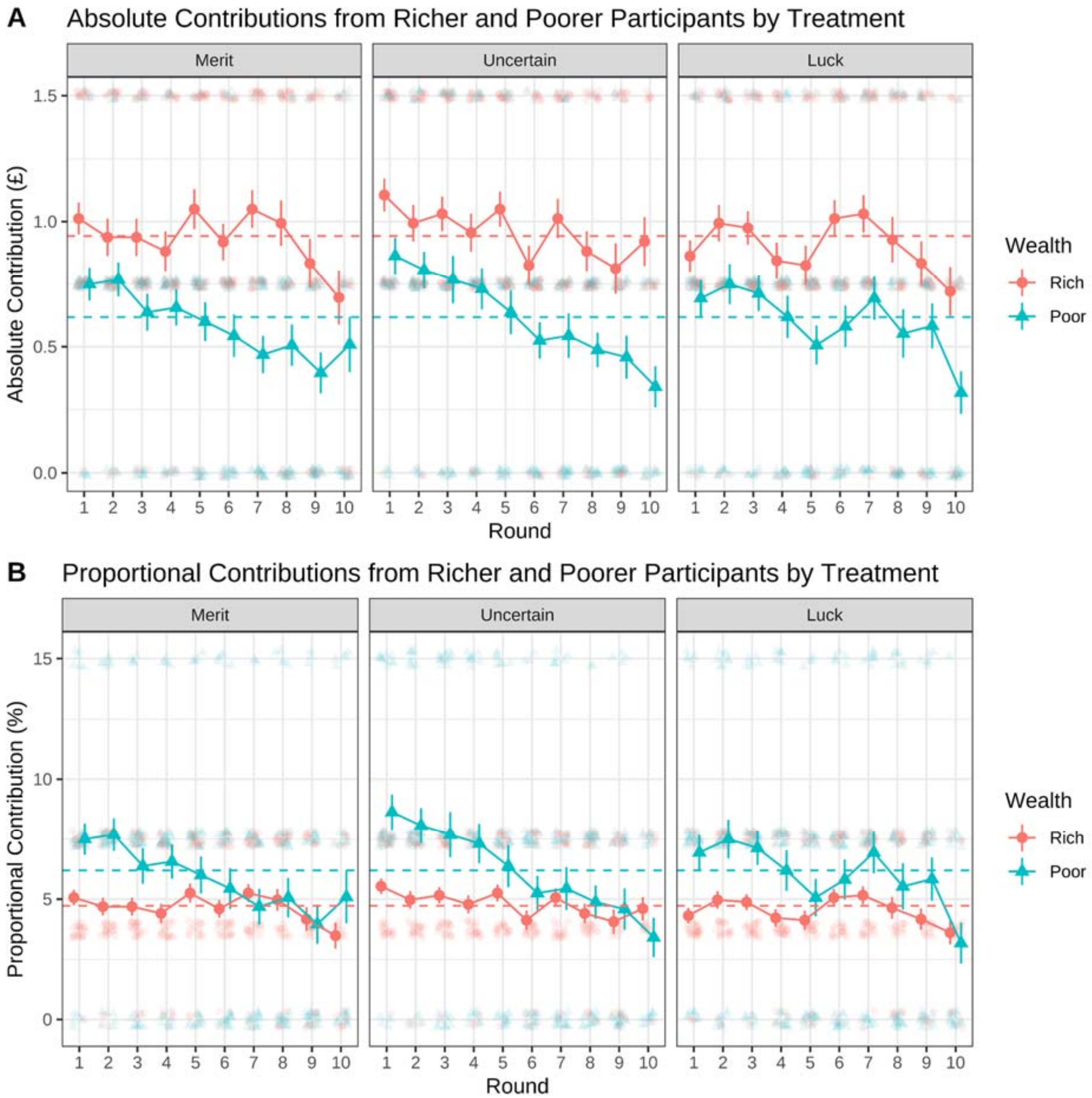


Figure 3. ***Participant Contributions in Absolute and Proportional Terms***

Plots show the mean contributions (excluding rounds in which the group target had already been met) of richer and poorer participants by treatment. Points in the background represent raw data: they are faded and jittered to show overlapping responses. Larger coloured points represent the mean; bars represent the standard error. Dashed horizontal lines represent the mean total contribution from all richer and poorer participants across all three treatment conditions. Plot A shows the absolute value of contributions (£0 / £0.75 / £1.50). Plot B shows contributions in proportional terms: with a £1.50 contribution being 7.5% of a richer participant's and 15% of a poorer participant's endowment; a £0.75 contribution represented as 7.5% and 3.75%, respectively; and a £0 contribution represented as 0%.

If what people perceive as fair is insufficient to solve the problem, under what conditions do groups still manage to succeed?

To answer this question we first compared richer and poorer participants' contributions in successful and unsuccessful groups. We found that richer participants contributed 49.4% of their wealth in successful groups and 40.6% of their wealth in unsuccessful groups – a difference of 8.8 percentage points. On the other hand, poorer participants contributed 67.8% of their wealth in successful groups and 40.0% of their wealth in unsuccessful groups – a difference of 27.8 percentage points. Standard and Bayesian ANCOVAs detected wealth $F(1, 237) = 29.72$, $p < .001$; $BF_{10} = 23{,}862$) and group success effects ($F(1, 237) = 29.65$, $p < .001$; $BF_{10} = 23{,}164$); while our multilevel model highlighted a significant interaction of poor wealth and group success ($t(236) = 3.0$, $p = .003$). This interaction indicated that the effect of poor wealth on proportional contributions depended on group success.

These results suggested that poorer participants' contributions were particularly relevant for group outcomes. We verified this by comparing Gini coefficients within groups at the end of the game (see Plot B in Figure 3), which were higher in successful groups (M = 0.30) than unsuccessful groups ($M = 0.20$; $F(1, 54) = 12.0$, $p = .001$; $BF_{10} = 33.3$) and were not moderated by treatment ($F(1, 54) = 0.22$, $p = .800$; $BF_{10} = 0.007$). Together, these findings indicated that poorer participants stepping up to contribute substantially more than their fair share helped to explain group success.

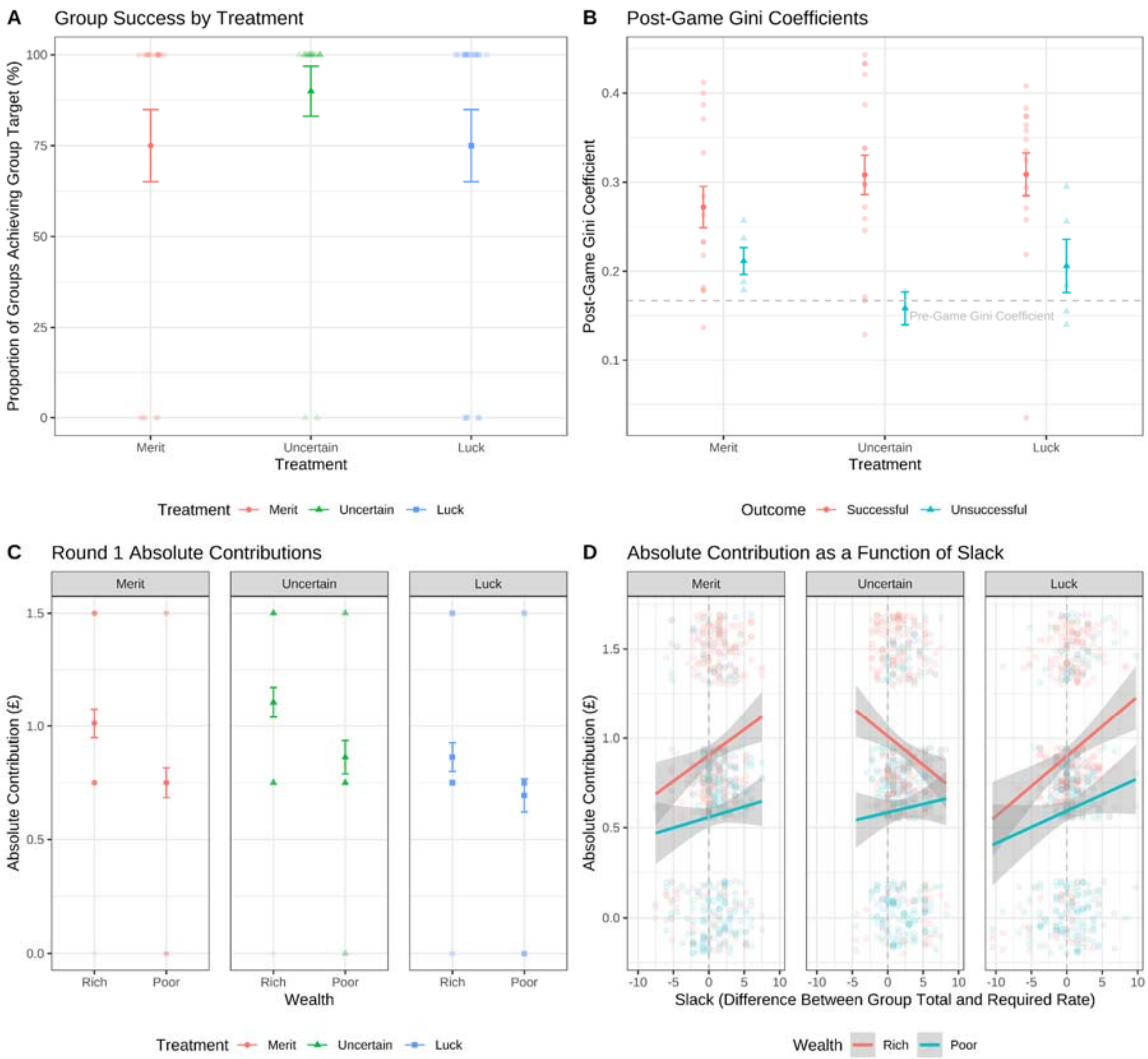


Figure 4. *Group Outcomes, Group Inequality, Cumulative Contributions, and Slack*

Plot A shows the proportion of groups who were successful, with points in the background representing groups and summary points showing the mean success rates and error bars representing the standard error (where 100 represents group success and 0 represents group failure). Plot B shows the Gini coefficients of successful and unsuccessful groups at the end of the game, illustrating how within-group inequality tended to increase over time – particularly in successful groups. The horizontal dashed grey line represents the Gini coefficient of all groups at the start of the game (0.17); points in the background represent groups; solid coloured points and bars representing means and standard errors for successful and unsuccessful groups in each treatment. Plot C shows cumulative group contributions over time in each treatment. Points represent the mean contribution from groups in each treatment in each round and bars represent the standard error. The diagonal grey line illustrates the required rate of cumulative group contributions in order to succeed in reaching the target sum of £30 within 10 rounds. Plot D shows contributions from richer and poorer participants in each treatment as a function of the difference between the current group total and the required contribution rate of £3 per round. For example, a group that has collectively contributed £25 after 9 rounds would be plotted at -2, because they are £2 behind the curve.

If the poor contributing a greater proportion of their wealth helped to explain group success across all three treatments, what explained the higher success rate of groups in our uncertain treatment? One explanation, illustrated in Plot C in Figure 4, was that participants in this treatment on average contributed more towards the group target in the first round (M = £0.98) than participants in the merit (M = £0.88) and luck treatments (M = £0.78) ($F(2,237)$ = 4.39, p = .013; BF_{10} = 2.17). This meant that uncertain groups for the most part had to sustain rather than build momentum.

A second explanation, illustrated in Plot D in Figure 4, was that the Uncertain Rich were more likely than the Deserving and Lucky Rich to support their groups when they fell behind the required rate of contribution. We verified this by running separate multilevel models with contributions from richer and poorer participants as dependent variables: the only significant interaction we detected was between the slack variable and the uncertain treatment on richer participants' contributions ($t(168)$ = -2.5, p = .015; see Supplementary Materials, Table S2 for full model outputs and Figure S6 for model predictions across treatment and wealth levels). In summary, the Uncertain Rich picked up the slack when their groups fell behind the required rate of contribution in a way that the Deserving and Lucky Rich did not.

Table 3. *Hypotheses and Results*

The table summarises the statistical tests we used for our main hypotheses (PR = pre-registered) and whether they were supported.

Hypothesis	Theme	Prediction	Statistical Tests	Supported
1	Fairness: Insufficiency	Participants' will on average judge it fair that richer and poorer players contribute less than 50% of their wealth	None	✔
2	Fairness: Level of Wealth (Rich vs. Poor)	Richer (poorer) participants will judge it fair that they contribute a lower proportion of their wealth than poorer (richer) participants	Chi-squared tests (H2a); Standard and Bayesian ANOVAs (H2b-c)	~
3 (PR)	Contributions: Level of Wealth (Rich vs. Poor)	Richer participants will contribute more than poorer participants in absolute terms, but less in proportional terms	Standard and Bayesian ANOVAs + Multilevel Models	✔
4 (PR)	Fairness: Cause of Wealth (Merit vs. Luck vs. Uncertain)	Individuals in the merit treatment will expect the poor to contribute a higher proportion of their wealth	Standard and Bayesian ANOVAs	✖
5 (PR)	Contributions: Cause of Wealth (Merit vs. Luck vs. Uncertain)	The Deserving Rich will contribute a lower proportion of their wealth than the Uncertain and the Lucky Rich	Standard and Bayesian ANOVAs + Multilevel Model	✖
6 (PR)	Group Outcomes	Luck-based groups would be more successful than merit-based and uncertain groups	Chi-squared tests	✖

Discussion

Our main finding is that what many people perceive to be fair is insufficient to solve the collective action problem at hand. Overall, participants judged it fair that richer participants contribute 37.5% and poorer participants contribute 40.4% of their wealth – both of which fell short of the average 50% figure required to solve the problem. Similarly, what a significant proportion of individuals (42.4%) judged to be fair for their group as a whole to contribute was not enough for group success. This was considerably higher than the proportion of individuals (7.6%) who judged it fair that their group should contribute *more* than the target of £30. This finding supports Hypothesis 1 and is highly relevant to a host of real-world collective action problems, including climate change and sustainable fishing, where what is judged to be fair may ultimately be insufficient.

One explanation for this finding was that fairness judgements were often self-serving. This was evident in the fact that 23% of poorer participants (vs. just 7.2% of richer participants) judged progressive wealth redistribution to be fair; while 31.5% of richer participants (vs. just 15% of poorer participants) judged regressive wealth redistribution to be fair. These self-serving interpretations of fairness, which partially supported Hypothesis 2, have been cited as a major barrier in international climate negotiations (Lange et al. 2010; Carlsson et al. 2013; Brick and Visser 2015; Reindl, 2022).

Even when fairness judgements are not self-serving, if they are insufficient then group success will require some members to contribute more than what is deemed to be their fair share. In our study, it was predominantly poorer participants who stepped up in this way and who had a disproportionate influence on group outcomes. Despite having less to give, and despite it generally being judged fair that they contribute 40.4% of their endowment, they consistently contributed a higher proportion of their wealth (M = 62.3%) than richer participants (M = 47.6%). This finding supported Hypothesis 3 and was particularly true in successful groups, in which wealth inequality increased as a result.

The level of people's wealth therefore had an important effect on fairness judgements and contributions, unlike the cause of their wealth. We did not find evidence to support Hypothesis 4 (that fairness judgements would differ between treatments) or Hypothesis 5 (that contributions would differ between treatments). In other words, richer participants generally contributed a lower proportion of their wealth regardless of its cause. The rich did this despite the fact that they had more to lose in financial terms than poorer participants, which might have motivated them to cooperate more.

Contrary to Hypothesis 6, we found that uncertain (rather than luck-based) groups were the most successful. We attributed this to two factors: uncertain participants' higher contributions in round 1 and the response of the Uncertain Rich to their group contributing less than the required rate of contribution. One possible explanation for these differences, discussed in the introduction, is that uncertainty about the cause of inequality can generate an egalitarian pull on the behaviour of meritocrats (Cappelen et al., 2022). The reason for this is that people may view uncertainty as an unfair way of distributing wealth in comparison with merit or luck. In an experiment by Samuelson & Allison (1994), for instance, luck and merit were both viewed as valid causes of inequality. More spurious causes of inequality on the other hand were not. In our experiment, the Deserving and Lucky Rich may have similarly viewed merit and luck as equally valid causes of inequality and been less willing to redistribute wealth as a result – unlike the Uncertain Rich. It is also possible that conflicting beliefs about the legitimacy of inequality between richer and poorer participants in luck- and merit-based groups may have hampered group coordination. We acknowledge, however, that validating these explanations would require further research.

It is also worth highlighting here what we believe to be the main limitations of our experiment, which relate to the generalisability of our findings. Firstly, it is unlikely that our merit and luck manipulations accurately reflect how people think about these phenomena in relation to their life outcomes in natural settings. In practice they are often conflated, as illustrated by the Latin proverb, 'fortune favours the brave' (Flusfeder, 2022). Secondly, failure in real-world collective action tends to involve higher stakes than simply losing one's endowment. It is likely that our participants were therefore less concerned about collective failure than individuals might be in real-world equivalent situations, which may result in different behavioural responses. Future research could therefore further explore attributions of merit and luck in experimental and natural settings, or it could consider increasing the stakes. In addition, it could explore how endowments based on performance shape people's level of trust in others and their feelings about other members of their group. And how these feelings might, in turn, influence both individual-level contribution behaviour and group-level outcomes.

In conclusion, our findings illustrate how what is perceived to be fair in collective action problems may be insufficient for the attainment of group goals. On an individual level our results highlight the general reluctance of richer individuals to sacrifice personal wealth and reduce inequality in order to support joint efforts to avert collective risks. Our findings also highlight the disproportionate influence of the poor in these situations, whose willingness to contribute considerably more than what they themselves deemed as fair was crucial to group success. Our results relating to the effect of different causes of wealth suggest that promoting a meritocratic message about the cause of wealth inequalities is unlikely to support cooperation in collective action problems; nor is telling people that they have been either lucky or unlucky (Frank, 2016; Markovits, 2019; Sandel, 2020). Instead, perhaps there is greater promise in highlighting the uncertainty inherent in any attempt to calculate the relative roles of luck and merit in our respective histories.

On a group level, our results illustrate both the value of early contributions and the value of individuals who are willing to pick up the slack if the group is not contributing at the rate required to solve the collective action problem. These were important factors underlying the higher success rate of groups in our uncertain treatment. The implication for policymakers is that if a group falls behind this required rate it may be unwise to expect its members, richer or poorer, to mitigate the impending disaster later down the line.

Constraints on Generality

Our results do not capture the impact of the high stakes associated with real-world collective action problems such as climate change and pandemics. In addition, our sample was made up of participants available on Prolific Academic and MTurk and consequently reflects the populations signed up to these platforms (see Supplementary Materials Table S3 for details). Real-world collective action problems often involve individuals from many more countries – many of whom are likely to hold different beliefs about both fairness and the relationship between merit, luck, and inequality. Lastly, the contribution options available to participants in each round (£0 / £0.75 / £1.50) were artificially restrictive to force individuals to solve the problem over multiple rounds of interaction, similar to many real world problems; in reality, individuals in collective action problems are likely to have much greater control over their contributions. Beyond these limitations, we have no reason to believe that the results depend on other characteristics of the participants, materials, or context.

References

Adams, J. S. (1965). Inequity in social exchange. In *Advances in experimental social psychology* (Vol. 2, pp. 267-299). Academic Press.

Alan, S., & Ertac, S. (2017). Belief in hard work and altruism: Evidence from a randomized field experiment. *Human Capital and Economic Opportunity Working Paper*.

Alesina, A. F., Glaeser, E. L., & Sacerdote, B. (2001). Why doesn't the US have a European-style welfare system?. *Brookings Paper on Economic Activity*.

Alesina, A., Glaeser, E., & Glaeser, E. L. (2004). *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press.

Alesina, A., Stantcheva, S., & Teso, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, *108*(2), 521-54.

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. *American Economic Review*, *85*(5), 1337-1343.

Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic perspectives*, *11*(1), 109-126.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*(1), 3.

Bernard, M., Reuben, E., & Riedl, A. (2012). Fairness and coordination: the role of fairness principles in coordination failure and success.

Bowles, S., & Gintis, H. (2011). *A cooperative species*. Princeton University Press.

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton university press.

Brick, K., & Visser, M. (2015). What is fair? An experimental guide to climate negotiations. *European Economic Review*, *74*, 79-95.

Cappelen, A. W., Mollerstrom, J., Reme, B. A., & Tungodden, B. (2022). A Meritocratic Origin of Egalitarian Behaviour. *The Economic Journal*, *132*(646), 2101-2117.

Carpenter, G., Kleinjans, R., Villasante, S., & O'Leary, B. C. (2016). Landing the blame: The influence of EU Member States on quota setting. *Marine Policy*, *64*, 9-15.

Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, Å., Qin, P., & Sterner, T. (2013). A fair share: Burden-sharing preferences in the United States and China. *Resource and Energy Economics*, *35*(1), 1-17.

Chan, K. S., Mestelman, S., Moir, R., & Muller, R. A. (1996). The voluntary provision of public goods under varying income distributions. *Canadian Journal of Economics*, 54-69.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88-97.

Croson, R., Fatas, E., & Neugebauer, T. (2005). Reciprocity, matching and conditional cooperation in two public goods games. *Economics Letters*, *87*(1), 95-101.

Davide, M., Parrado, R., & Campagnolo, L. (2017). Fairness in NDCs: comparing mitigation efforts from an equity perspective. *United Nations Framework Convention on Climate Change.* Accessed online at <https://unfccc.int/sites/default/files/resource/312_Fairness%20in%20NDCs.%20Comparing%20mitigation%20efforts%20from%20an%20equity%20perspective.pdf>

De Cremer, D. (2007). When the rich contribute more in public good dilemmas: The role of provision point level. *European Journal of Social Psychology*, *37*(3), 536-546.

Diekmann, K. A., Samuels, S. M., Ross, L., & Bazerman, M. H. (1997). Self-interest and fairness in problems of resource allocation: allocators versus recipients. *Journal of Personality and Social Psychology*, *72*(5), 1061.

Elster, J. (2006). Fairness and norms. *Social Research*, 365-376.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817-868.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397-404.

Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, *100*(1), 541-56.

Flusfeder, D. (2022). *Luck: A Personal Account of Fortune, Chance and Risk in 13 Investigations*. 4th Estate.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, *82*(2), 225-246.

Frank, R. H. (2016). *Success and Luck*. Princeton University Press.

Gächter, S. (2006). *Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications* (No. 2006-03). CeDEx Discussion Paper Series.

Gonzalez, A. M., Macchia, L., & Whillans, A. V. (2022). The developmental origins and behavioral consequences of attributions for inequality. *Journal of Experimental Social Psychology*, *101*, 104329.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 1243-1248.

Heap, S. P. H., Ramalingam, A., & Stoddard, B. V. (2016). Endowment inequality in public goods games: A re-examination. *Economics Letters*, *146*, 4-7.

Henrich, N., & Henrich, J. P. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. Oxford University Press.

Hine, D. W., & Gifford, R. (1996). Attributions about self and others in commons dilemmas. *European Journal of Social Psychology*, *26*(3), 429-445.

Hume, D. (1740). *A Treatise of Human Nature.*

JASP Team (2020). JASP (Version 0.14) [Computer software].

Joireman, J. A., Kuhlman, D. M., & Okuda, H. (1984). Fairness judgements in an asymmetric public goods dilemma. In *Social dilemmas and cooperation* (pp. 99-116). Springer, Berlin, Heidelberg.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*(1), 66.

Keser, C., & Van Winden, F. (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, *102*(1), 23-39.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, *90*(4), 1072-1091.

Kopelman, S., Weber, J. M., & Messick, D. M. (2002). Factors influencing cooperation in commons dilemmas: A review of experimental psychological research.

Koo, H. J., Piff, P. K., & Shariff, A. F. If I could do it, so can they: among the rich, those with humbler origins are less sensitive to the difficulties of the poor.

Laming, D. (1984). The relativity of ‘absolute’ judgements. *British Journal of Mathematical and Statistical Psychology*, *37*(2), 152-183.

Lange, A., Löschel, A., Vogt, C., & Ziegler, A. (2010). On the self-interested use of equity in international climate negotiations. *European Economic Review*, *54*(3), 359-375.

Ledyard, J.O. (1995). Public goods: a survey of experimental research. In *Handbook of Experimental Economics*, ed. JH Kagel, AE Roth, 2:111–81. Princeton, NJ: Princeton Univ. Press

Litman, L., Moss, A., Rosenzweig, C., & Robinson, J. (2021). Reply to MTurk, Prolific or panels? Choosing the right audience for online research. *Choosing the right audience for online research (January 28, 2021)*.

March, J. G. (1994). *Primer on decision making: How decisions happen*. Simon and Schuster.

Markovits, D. (2019). *The meritocracy trap*. Penguin UK.

Martinangeli, A. F., & Martinsson, P. (2020). We, the rich: inequality, identity and cooperation. *Journal of Economic Behavior & Organization*, *178*, 249-266.

Marwell, G., & Ames, R. E. (1979). Experiments on the provision of public goods. I. Resources, interest, group size, and the free-rider problem. *American Journal of Sociology*, *84*(6), 1335-1360.

Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, *105*(7), 2291-2294.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction?. *Psychological Bulletin*, *82*(2), 213.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. *The Quarterly Journal of Economics*, *122*(3), 1067-1101.

Nowak, M., & Highfield, R. (2011). *Supercooperators: Altruism, evolution, and why we need each other to succeed*. Simon and Schuster.

Olson, M. (1965). *The Logic Of Collective Action: Public Goods And The Theory Of Groups*. Harvard University Press.

Ostrom, E. (2008). Tragedy of the commons. *The New Palgrave Dictionary of Economics*, *2*.

Ostrom, E. (2010). Analyzing collective action. *Agricultural economics*, *41*, 155-166.

Oswald, A. J., Proto, E., & Sgroi, D. (2015). Happiness and productivity. *Journal of Labor economics*, *33*(4), 789-822.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22-27.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgement and Decision making*, *5*(5), 411-419.

Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly journal of economics*, *110*(3), 551-584.

Piketty, T. (2020). *Capital and ideology*. Harvard University Press.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 1281-1302.

Rapoport, A. (1988). Provision of step-level public goods: Effects of inequality in resources. *Journal of Personality and Social Psychology*, *54*(3), 432.

Reindl, I. (2022). Wealth and Vulnerability to Climate Change: An Experimental Study on Burden Sharing among Heterogeneous Agents. *Environmental and Resource Economics*, *82*(4), 791-823.

Roth, A. E., & Murnighan, J. K. (1982). The role of information in bargaining: An experimental study. *Econometrica: Journal of the Econometric Society*, 1123-1142.

Samuelson, C. D., & Allison, S. T. (1994). Cognitive factors affecting the use of social decision heuristics in resource-sharing tasks. *Organizational Behavior and Human Decision Processes*, *58*(1), 1-27.

Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good?*. Penguin UK.

Sandler, T. (1992). *Collective action: Theory and applications*. University of Michigan Press.

Schoemaker, P. J. (2011). *Brilliant mistakes: Finding success on the far side of failure*. University of Pennsylvania Press.

Segbroeck, S. V., Pacheco, J. M., Lenaerts, T., & Santos, F. C. (2013). Evolution of fairness and conditional cooperation in public goods dilemmas. In *Proceedings of the European Conference on Complex Systems 2012* (pp. 827-830). Springer, Cham.

Son Hing, L. S., Bobocel, D. R., Zanna, M. P., Garcia, D. M., Gee, S. S., & Orazietti, K. (2011). The merit of meritocracy. *Journal of personality and social psychology*, *101*(3), 433.

Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*(4), 1-7.

Stewart, N., Brown, G. D., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, *112*(4), 881.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1-26.

Stroebe, W., & Frey, B. S. (1982). Self-interest and collective action: The economics and psychology of public goods. *British Journal of Social Psychology*, *21*(2), 121-137.

Sugden, R. (1984). Reciprocity: the supply of public goods through voluntary contributions. *The Economic Journal*, *94*(376), 772-787.

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, *108*(29), 11825-11829.

Taylor, D. M., Moghaddam, F. M., Gamble, I., & Zellerer, E. (1987). Disadvantaged group response to perceived inequality: From passive acceptance to collective action. *The Journal of Social Psychology*, *127*(3), 259-272.

The Sunday Times. (2021). Marcus Rashford tops The Sunday Times Giving List 2021. Accessible online: <https://www.thetimes.co.uk/article/marcus-rashford-sunday-times-giving-list-2021-l9td90n0l>

United Nations. (2015). *Paris Agreement*. Accessed online at <https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement>

United Nations Framework Convention on Climate Change. (2021). Accessed online at <https://unfccc.int/news/cop26-update-to-the-ndc-synthesis-report>

Van Dijk, E., & Wilke, H. (1993). Differential interests, equity, and public good provision. *Journal of Experimental Social Psychology*, *29*(1), 1-16.

Van Dijk, E., & Wilke, H. (1994). Asymmetry of wealth and public good provision. *Social Psychology Quarterly*, 352-359.

Van Lange, P. A., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*(2), 125-141.

Vasconcelos, V. V., Santos, F. C., Pacheco, J. M., & Levin, S. A. (2014). Climate policies under wealth inequality. *Proceedings of the National Academy of Sciences*, *111*(6), 2212-2216.

Vicens, J., Bueno-Guerra, N., Gutiérrez-Roig, M., Gracia-Lázaro, C., Gómez-Gardeñes, J., Perelló, J., ... & Duch, J. (2018). Resource heterogeneity leads to unjust effort distribution in climate change mitigation. *PloS one*, *13*(10), e0204369.

Wade-Benzoni, K. A., Tenbrunsel, A. E., & Bazerman, M. H. (1996). Egocentric interpretations of fairness in asymmetric, environmental social dilemmas: Explaining harvesting behavior and the role of communication. *Organizational Behavior and Human Decision Processes*, *67*(2), 111-126.

Walster, E., Berscheid, E., & Walster, G. W. (1976). New directions in equity research. In *Advances in experimental social psychology* (Vol. 9, pp. 1-42). Academic Press.

Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review*, *8*(3), 281-307.

World Bank. (2019). Gini index (World Bank estimate) <https://data.worldbank.org/indicator/SI.POV.GINI?most_recent_value_desc=false> (accessed 29 October 2021)

Young, M. (1958). *The rise of the meritocracy*. Routledge.

Supplementary Materials

Supplementary Figures

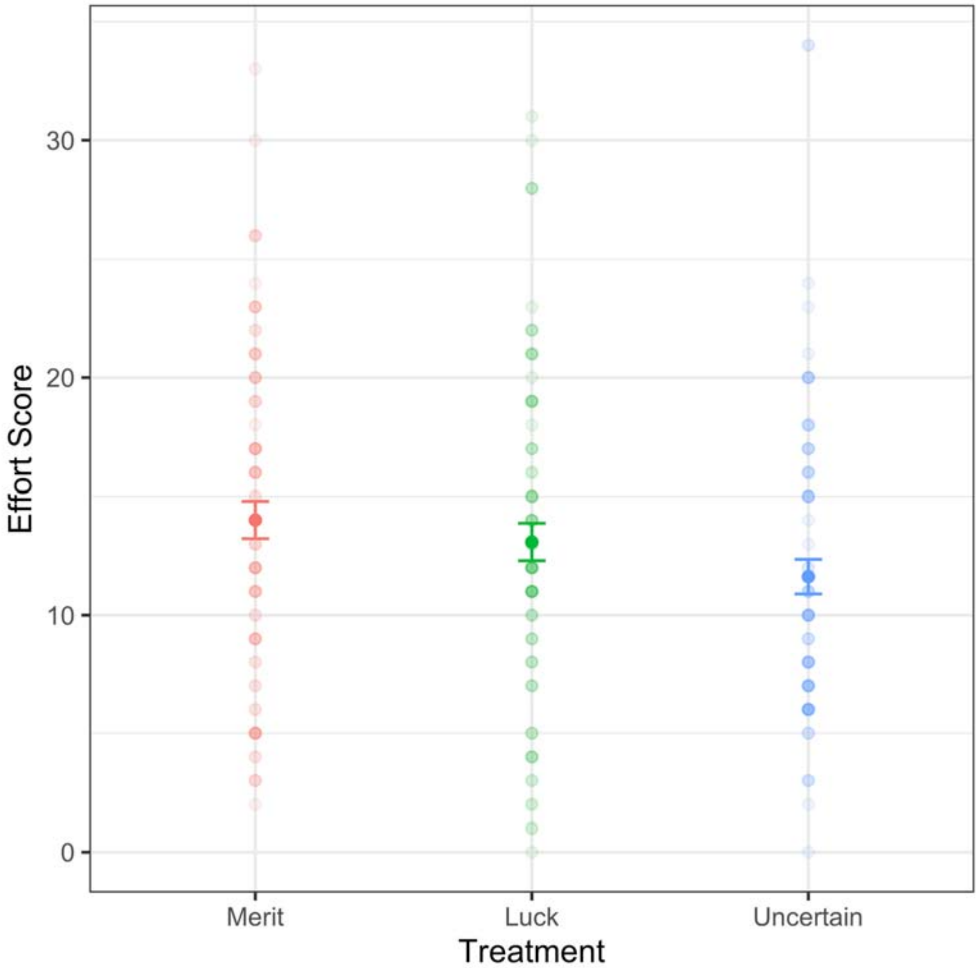


Figure S1. *Effort Scores Between Treatment Conditions*

Participants' scores in our five-minute effort task, split by treatment. Coloured points represent the mean, with coloured bars representing the standard error. Grey points in the background represent raw data. Differences between treatments suggest that participants understood the different implications of their performance in each condition, but we did not detect a significant treatment effect $F(2, 237) = 2.5$, $p = .088$.

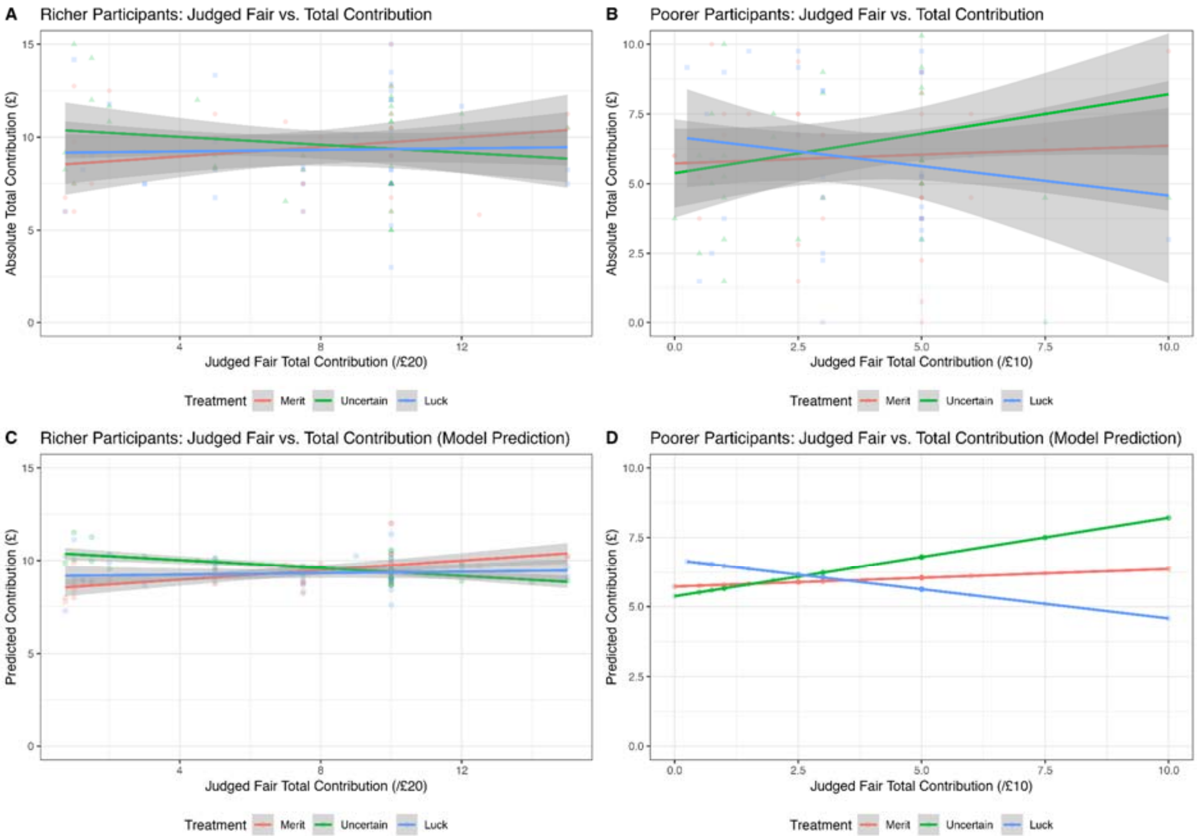


Figure S2. *Contributions as a Function of Fairness Judgements*

Plots illustrate the relationship between: x) richer and poorer participants' judgements about what was fair for someone of their wealth level to contribute; and y) their total contribution (their actual total contribution in plots A and B vs. their contribution predicted by our model in plots C and D). Plots A and B show the relationship between participants' judgements and what they actually contributed within the game in absolute terms, with points in the background representing individuals, separate linear regression lines for each treatment, and shaded areas representing the standard error. Plots C and D show predicted contributions from multilevel models that included total contribution as the dependent variable, fixed fairness judgement and treatment effects, and random group effects to take account of the nested structure of the data. These models did not detect any effect of fairness judgements for either richer participants ($t(107) = 0.7$, $p = .513$) or poorer participants ($t(112) = 0.4$, $p = .735$). Points in the background represent predictions for individuals and coloured lines are generalised linear regression lines for each treatment.

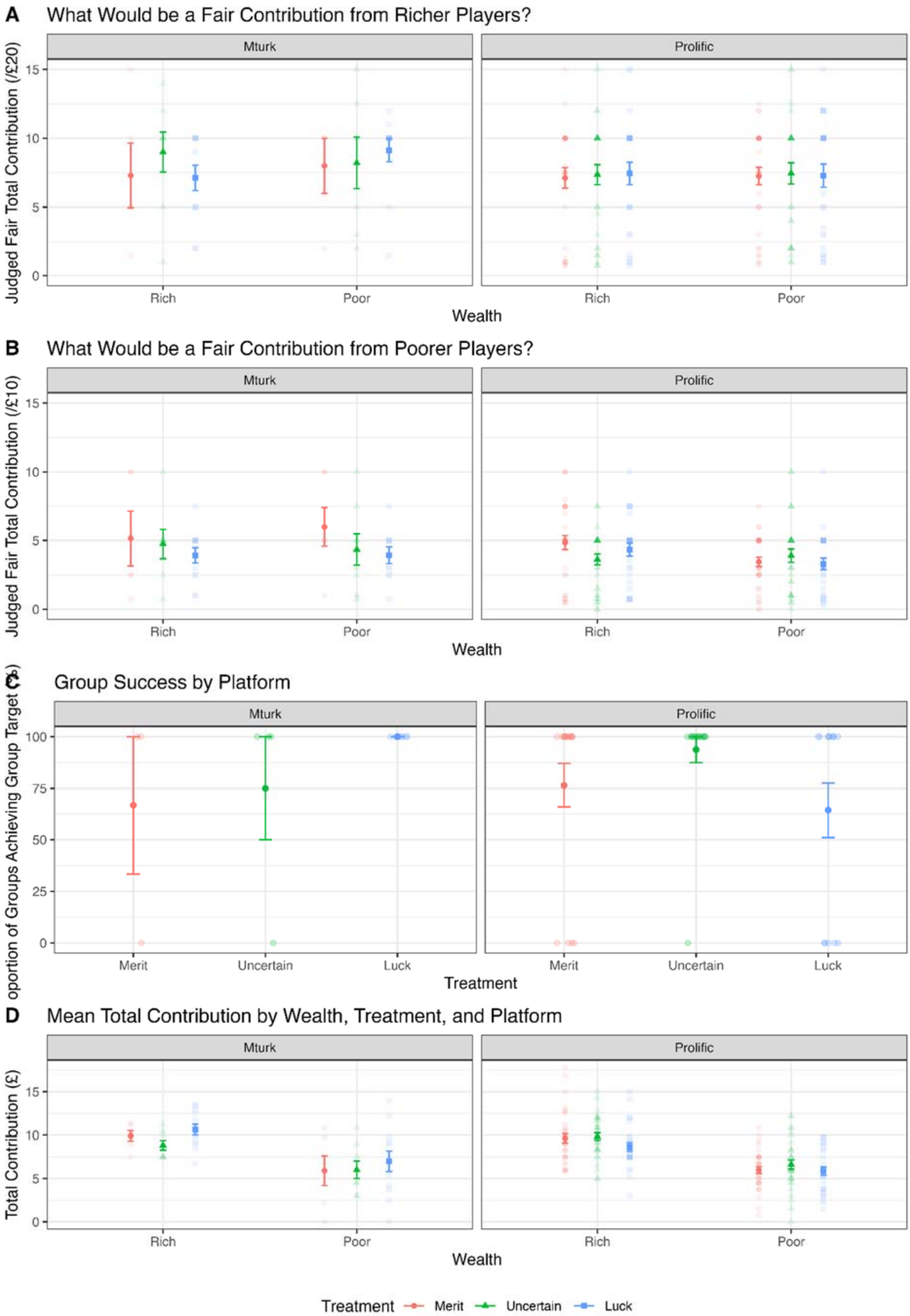

A What Would be a Fair Contribution from Richer Players?
Mturk
Prolific
Judged Fair Total Contribution (/£20)
Rich
Poor
Wealth
B What Would be a Fair Contribution from Poorer Players?
Judged Fair Total Contribution (/£10)
C Group Success by Platform
Proportion of Groups Achieving Group Target (%)
Merit
Uncertain
Luck
Treatment
D Mean Total Contribution by Wealth, Treatment, and Platform
Total Contribution (£)
Treatment Merit Uncertain Luck

Figure S3. *Main Outcome Variables by Data Collection Platform (MTurk and Prolific)*

Plots A and B show responses to the two questions put to participants in the pre-game questionnaire: 'In your opinion, what would be a fair total contribution in £ to the group account during the game?' for players starting with £20 (plot A) and players starting with £10 (plot B). ANOVAs did not detect a platform effect for judgements about richer players ($F(1, 216) = 1.31$, $p = .254$) or poorer players ($F(1, 224) = 3.32$, $p = .070$). Points in the background represent individuals, with summary points showing the mean response and error bars representing the standard error. Plot C shows the proportion of groups who achieved the target in each treatment, split by platform. Points in the background represent groups. Summary points show the mean success rate where 100 represents group success and 0 represents group failure, with error bars representing the standard error. Chi-squared tests indicated that the proportion of groups that were successful in each treatment did not differ significantly between platforms (merit: χ^2 (1) = 0, $p = 1$; luck: χ^2 (1) = 1.27, $p = .260$; uncertain: χ^2 (1) = 0.03, $p = .852$). Plot D shows absolute contributions to the group account by wealth and treatment levels and platform, with points in the background representing individuals and summary points showing the mean response (and error bars represent the standard error). An ANOVA did not detect a significant platform effect ($F(1, 228) = 0.40$, $p = .526$).

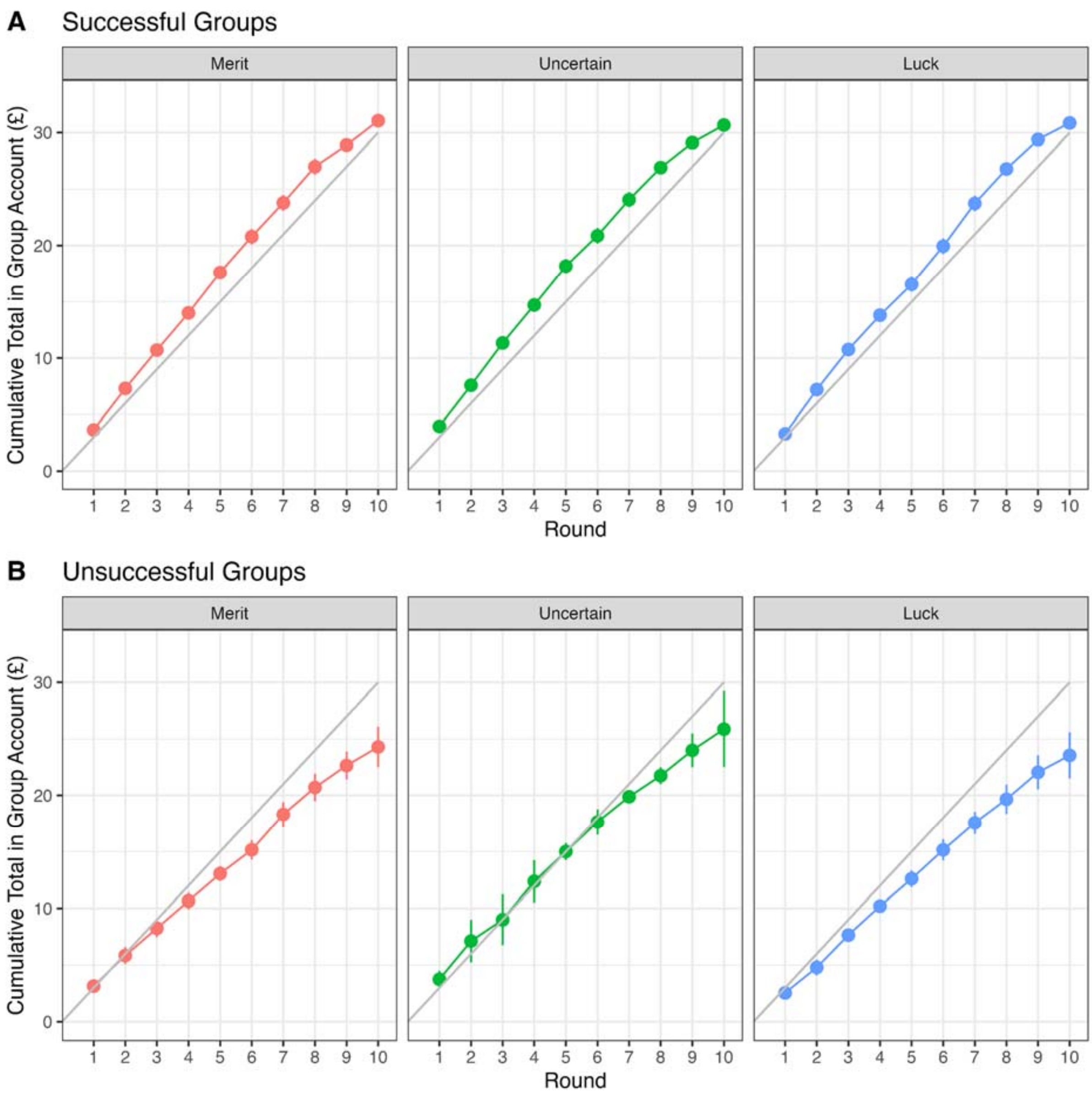


Figure S4. *Cumulative Group Contributions Between Treatments*

Plots show the cumulative total in group accounts in successive rounds. Groups are divided by outcome and treatment, with A plots showing successful groups and B plots showing unsuccessful groups by treatment. Coloured points show the mean cumulative total in the group account in each round, with bars representing the standard error. Grey lines show a linear trajectory towards achieving the group target of £30 within the 10 rounds.

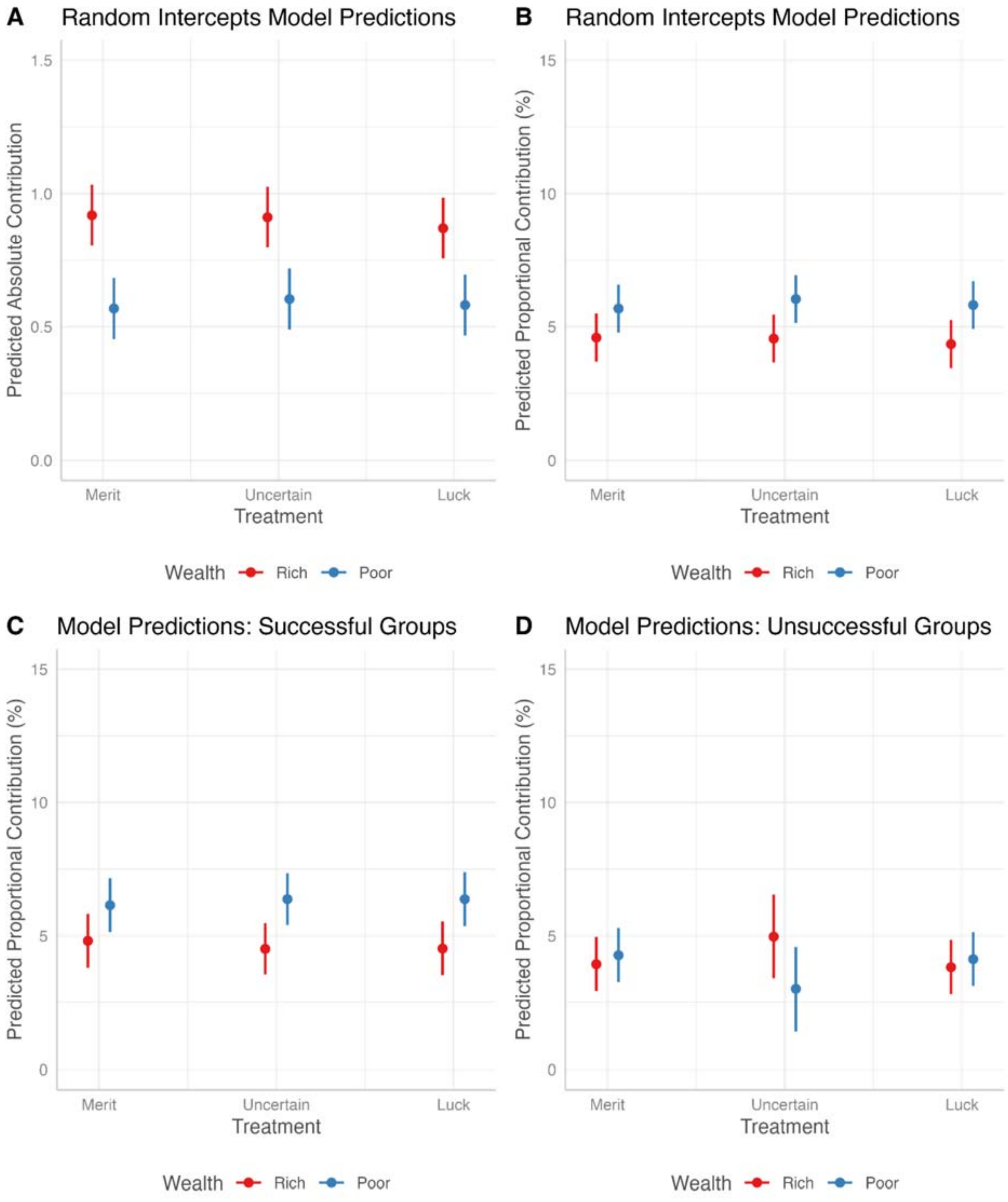


Figure S5. *Multilevel Model Predictions: Contributions by Treatment, Wealth, and Group Success*

Plots show predicted contributions from multilevel models with absolute and relative contributions specified as the dependent variable and treatment and wealth as predictor variables, with random intercepts at the round, individual, and group levels (see full model outputs in Table 1 below). Plot A shows predicted absolute contributions. Plot B shows predicted contributions in relative terms. Plot C shows predicted relative contributions in successful groups, showing higher predictions for poorer participants across all treatments. Plot D shows predicted relative contributions in unsuccessful groups.

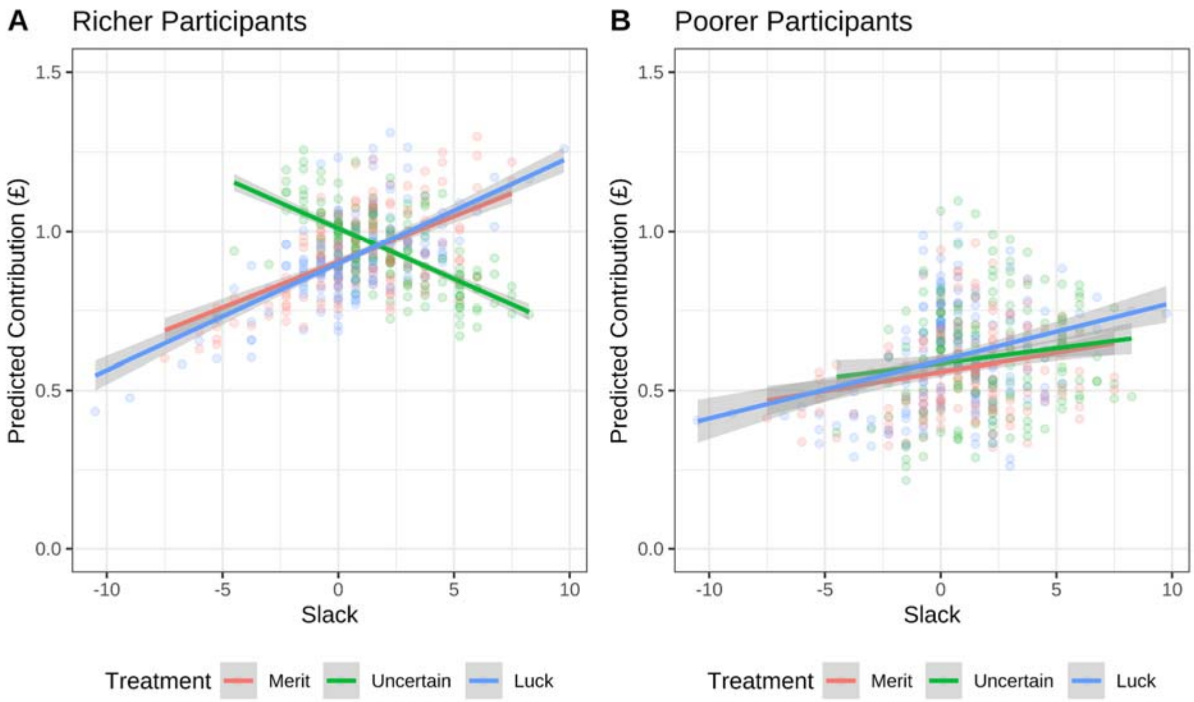


Figure S6. *Multilevel Model Predictions: Contributions as a Function of Slack*

Plots show predicted individual absolute contributions from richer participants (A) and poorer participants (B) as a function of slack (defined as the difference between £3 per round cumulative required contribution and the current group total) , with separate slopes for different treatment levels. Multilevel models included fixed treatment and slack effects and random intercepts at the round and group level. The plots illustrate the generally positive relationship between the level of slack and predicted contributions – with the exception of richer participants in the uncertain treatment, who *increase* their contributions when the group falls behind the required rate.

Supplementary Tables

Table S1. *Relative Contributions: Multilevel Model Parameter Estimates*

The table shows estimates from a multilevel model with relative contributions as the dependent variable, treatment and wealth as fixed effects, and random intercepts at the round, individual, and group levels. The fixed part of the model indicates a significant wealth effect. The random part of the model shows that a greater proportion of variance is accounted for at the round level (6.7%) than at the group level (0.7%).

	Estimate	**Standard Err.**	**Degrees of Freedom**	***t*-value**	***p*-value**
Fixed Effects					
Intercept (Merit Treatment)	4.59	0.46	27	10.05	< .001
Uncertain Treatment	-0.04	0.44	234	-0.09	.932
Luck Treatment	-0.24	0.44	234	-0.56	.578
Wealth Poor	1.09	0.44	234	2.49	.014*
Luck Treatment: Wealth Poor	0.39	0.62	234	0.64	.525
Uncertain Treatment: Wealth Poor	0.38	0.62	234	0.61	.545
Random Effects	**Variance**	**Standard Dev.**			
Intercept (UniqueID)	2.51	1.58			
Intercept (GroupID)	0.00	0.00			
Intercept (Round)	1.13	1.07			
Residual	13.15	3.63			

Table S2. *Contributions as a Function of Slack: Multilevel Model Parameter Estimates*

The table shows estimates from multilevel models on data from richer and poorer participants with absolute contributions as dependent variables, treatment and slack as fixed effects, and random intercepts at the round individual, and group levels.

	Estimate	**Standard Err.**	**Degrees of Freedom**	***t*-value**	***p*-value**
Richer Participants					
Fixed Effects					
Intercept (Merit Treatment)	0.92	0.05	40	19.71	< .001
Uncertain Treatment	0.09	0.07	54	1.44	.155
Luck Treatment	-0.02	0.06	45	-0.30	.767
Slack	0.01	0.01	158	0.47	.638
Uncertain Treatment: Slack	-0.46	0.02	168	-2.47	.015*
Luck Treatment: Slack	0.01	0.02	161	0.59	.559
Random Effects	**Variance**	**Standard Dev.**			
Intercept (UniqueID)	0.02	0.16			
Intercept (GroupID)	0.01	0.09			
Intercept (Round)	0.00	0.06			
Residual	0.21	0.46			
Poorer Participants					
Fixed Effects					
Intercept (Merit Treatment)	0.59	0.06	40	9.57	< .001
Uncertain Treatment	0.07	0.08	44	0.85	.403
Luck Treatment	0.02	0.08	37	0.25	.802
Slack	-0.03	0.01	23	-2.00	.046*
Uncertain Treatment: Slack	-0.00	0.02	25	-0.04	.966
Luck Treatment: Slack	0.01	0.02	24	0.38	.704
Random Effects	**Variance**	**Standard Dev.**			
Intercept (UniqueID)	0.06	0.25			

Intercept (GroupID)	0.01	0.11
Intercept (Round)	0.01	0.09
Residual	0.20	0.44

Table S3. *Prolific Academic Participant Backgrounds*

The table summarises the demographic data (ethnicity and nationality) that were retrospectively available for 180 participants sampled from Prolific Academic. It therefore does not reflect our full sample of 240 participants, and should only be treated as an approximate indication of the make-up of our sample population.

	Count	**%**		**Count**	**%**
Ethnicity			**Nationality**		
Asian	7	4%	Algeria	1	0.6%
Black	17	9%	Australia	1	0.6%
Mixed	16	9%	Austria	1	0.6%
Other	5	3%	Canada	3	1.7%
White	133	74%	Czech Republic	1	0.6%
Missing	2	1%	Egypt	1	0.6%
			Estonia	1	0.6%
			France	2	1.1%
			Germany	1	0.6%
			Greece	2	1.1%
			Hungary	4	2.2%
			Iran	1	0.6%
			Ireland	1	0.6%
			Italy	12	6.7%
			Korea	1	0.6%
			Lebanon	1	0.6%
			Mexico	10	5.6%
			Netherlands	4	2.2%
			Poland	25	13.9%
			Portugal	29	16.1%
			Slovenia	1	0.6%
			South Africa	20	11.1%
			Spain	5	2.8%
			Turkey	1	0.6%

			United Kingdom	41	22.8%
			United States	10	5.6%

Experiment Materials

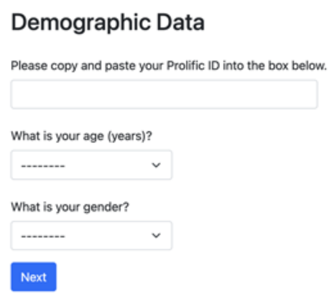
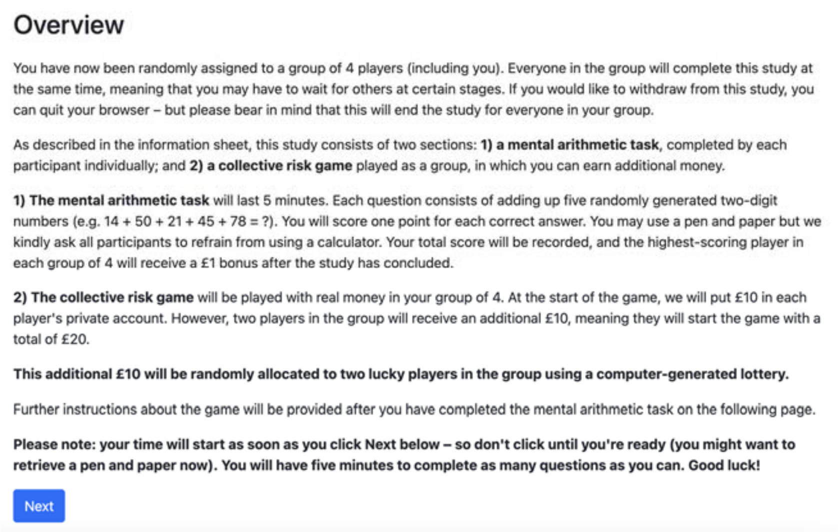


Figure S7. *Demographic Questions*

Drop-down options for age were: 18-24, 25-34, 35-44, 45-54, 55-64, 65+, I'd prefer not to answer. Drop-down options for gender were: Female, Male, Non-binary, Other, I'd prefer not to answer



Overview

You have now been randomly assigned to a group of 4 players (including you). Everyone in the group will complete this study at the same time, meaning that you may have to wait for others at certain stages. If you would like to withdraw from this study, you can quit your browser – but please bear in mind that this will end the study for everyone in your group.

As described in the information sheet, this study consists of two sections: **1) a mental arithmetic task**, completed by each participant individually; and **2) a collective risk game** played as a group, in which you can earn additional money.

1) The mental arithmetic task will last 5 minutes. Each question consists of adding up five randomly generated two-digit numbers (e.g. 14 + 50 + 21 + 45 + 78 = ?). You will score one point for each correct answer. You may use a pen and paper but we kindly ask all participants to refrain from using a calculator. Your total score will be recorded, and the highest-scoring player in each group of 4 will receive a £1 bonus after the study has concluded.

2) The collective risk game will be played with real money in your group of 4. At the start of the game, we will put £10 in each player's private account. However, two players in the group will receive an additional £10, meaning they will start the game with a total of £20.

This additional £10 will be randomly allocated to two lucky players in the group using a computer-generated lottery.

Further instructions about the game will be provided after you have completed the mental arithmetic task on the following page.

Please note: your time will start as soon as you click Next below – so don't click until you're ready (you might want to retrieve a pen and paper now). You will have five minutes to complete as many questions as you can. Good luck!

Next

Figure S8. *Experiment Overview*

Screenshot shows the overview presented to participants in the luck treatment, with the explanation that "This additional £10 will be randomly allocated to two lucky players in the group using a computer-generated lottery." In the merit treatment, this equivalent sentence read: "This additional £10 will be given to the two players who score highest in each group in the mental arithmetic task". In the uncertain treatment, this sentence read: "Whether or not you are one of the players who receives this additional £10 will be determined either by your performance in the mental arithmetic task or by chance – this will be randomly decided by the computer."

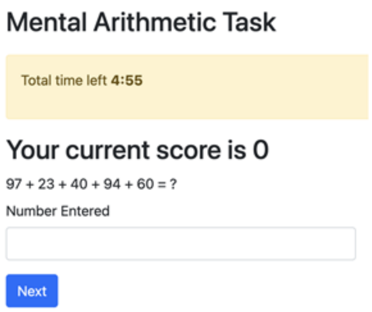


Figure S9. *Mental Arithmetic Task*

Experiment screenshot shows the mental arithmetic task, which lasted five minutes. Participants scored one mark for each correct answer. All two-digit numbers in the questions were randomly generated.

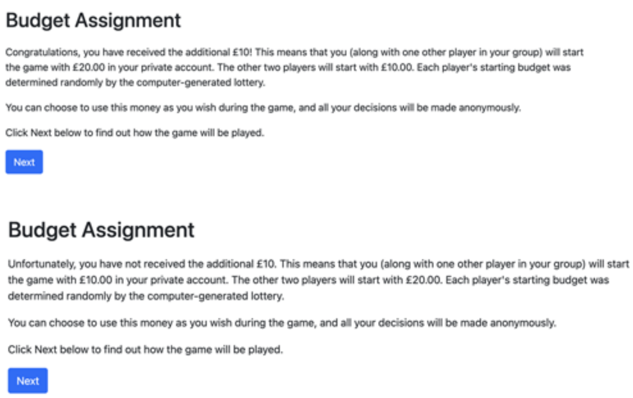


Figure S10. *Budget Assignment*

Screenshots shows budget assignments for richer participants (right) who receive the additional £10 and start the game with £20; and for poorer participants (right) who start the game with just £10.

Pre-Game Quiz

Please answer the following questions. The correct answers to the first three questions about the game instructions will be revealed on the following page.

Question 1

On average, how much in total (in £) over 10 rounds would each player have to transfer to the group account for the group to achieve its target of £30?

○ 0 ○ 3 ○ 7.5 ○ 10

Question 2

Imagine that the group succeeds in reaching the target of £30 after the 10 rounds, and you have £5 remaining in your private account. How much in £ will you take home from the game?

○ 0 ○ 2.5 ○ 5 ○ 10

Question 3

Assume that the group target of £30 is not reached, and you have £5 remaining in your private account. What is the % chance that you will lose this remaining amount in your private account?

○ 0 ○ 10 ○ 50 ○ 100

Question 4

In your opinion, what would be a fair total contribution in £ from players starting the game with £20 and those starting with £10? Please write your answer as a number (without the £ sign).

£20?

£10?

Next

Correct answers

Well done, you got the understanding questions correct. To double-check, here are the correct answers:

1) For the group to achieve its target of £30, on average each player should invest a total of **£7.50** to the group account (an average £0.75 on each round).

2) If the group account reaches the target of £30, and you have £5 remaining in your private account, you will take home **£5**.

3) If the group account does not reach the target of £30, and you have £5 remaining in your private account, you will lose these funds with a probability of **50%**.

The rest of the questions were opinion-based and therefore had no right or wrong answers.

Next

Start of Round Decision

This is round 1 of 10.

So far the group has contributed £0.00 towards the target of £30.00.

You have £20.00 left in your private account.

How much in £ would you like to transfer to the group account on this round?

○ £0.00 ○ £0.75 ○ £1.50

Next

End of Round Results

Results of round 5 of 10:

After 5 rounds, there is now a total of £15.00 in the group account.

Next

Start of Round Decision

This is round 6 of 10.

So far the group has contributed £15.00 towards the target of £30.00.

You have £6.25 left in your private account.

How much in £ would you like to transfer to the group account on this round?

○ £0.00 ○ £0.75 ○ £1.50

Next