Difficult Merits*

L. Flóra Drucker†

October 1, 2024

Click here for the latest version.

Abstract

Empirical and experimental evidence supports that people, on average, find performance-based income inequality fair and inequality due to pure luck unfair. In reality, performance can also depend on factors that are, like pure luck, not necessarily under one's control. I study in an experiment how transparent inequalities in two such factors – difficulty of the task and ability in the task – affect redistributive decisions. In the experiment, participants perform a simple task and then decide how to redistribute the income earned for their performance. I find that participants compensate those with more difficult tasks but do not compensate the ones with lower ability. On the other hand, participants deciding about their own payoffs treat inequality in both factors in a self-serving way. A structural estimation of the shares of heterogeneous fairness views reveals that more than a quarter of participants consider unequal difficulties in their decisions, none of them unequal ability; more than half still leave performance-based inequality unchanged, and the rest always equalize the incomes.

*I thank Botond Kőszegi for his continuous supervision of this project. I am indebted to Erik Sørensen for his generous help with the structural estimation, including sharing his code. I thank Peter Andre, Zsófia Bárány, Alexander Cappelen, Dietmar Fehr, Attila Gáspár, Marc Kaufmann, Hubert János Kiss, Mats Köster, Robert Lieli, Tímea Laura Molnár, Hannah Schildberg-Hörisch and Bertil Tungodden for discussions and helpful suggestions. I am also grateful to various seminar and conference participants for their comments and suggestions. I am indebted to Central European University for funding the project. The project received ethical approval from the CEU Ethical Research Committee (Reference number: 2020-2021/3/EX). This RCT was registered as AEARCTR-0006709.

†Düsseldorf Institute for Competition Economics, Heinrich Heine University Düsseldorf. Email: drucker@dice.hhu.de

1 Introduction

The level of redistribution in a society correlates with the sources of inequality its members tolerate. Experimental and empirical evidence shows that people, on average, prefer to reduce inequalities that arise from factors outside one's control (Konow, 2000; Fong, 2001; Alesina and Angeletos, 2005). In a situation where luck and merit can both determine income, this means that most people want to reduce inequality due to pure luck but keep the inequality based on merit (Cappelen et al., 2007, 2010; Almås et al., 2010, 2020; Durante et al., 2014). This body of research equates merit by performance in a task and infers attitudes toward merit-based differences from attitudes toward performance differences. An important reality of life, however, is that achieving a given level of performance is transparently more difficult and, consequently, takes more effort for some than for others. For example, a student who has a quiet room for studying can, in the same amount of time, do her homework more successfully than a student who is crammed into a small corner with her siblings running around.[1] At an even more basic level, a student more talented in a subject can do more exercises in a given time than a less talented one.[2] We know little about what people consider as merit in these cases: the final performance or simply the exerted effort. Do they still find rewarding performance fair, even though there were unequal opportunities due to external or internal factors?

I address this question in an experiment that compares how people redistribute performance-based inequality when there are unequal external difficulties to when there are unequal abilities in the task. The variation of external difficulties I introduce is similar to studying in different home environments: I randomly assign easier or harder tasks to subjects, which induces external variation in performance in a given time. Unequal ability is analogous to talent in the study example: subjects may have different abilities in the same task, again resulting in variation in performance within a period. While both factors are outside one's control when arriving at the experiment, I find that subjects treat them very differently: they redistribute income towards the person with more difficult tasks but do not redistribute to the one with lower ability. This result suggests that a significant share of people do not think one should be held responsible for external factors affecting her performance. In contrast, ability might be considered inherent to the person and, as such, her responsibility.

I describe the experimental design in Section 2. I run two experiments to test how people in-

[1] It is well established that the home environment plays a significant role in the development of cognitive and non-cognitive skills, and affects later life outcomes as well (Bradley et al., 2000; Mott, 2004; Falk et al., 2021).

[2] Some also regard talent as a result of lucky circumstances (Rawls, 1971; Littler, 2017).

corporate the two factors into their redistribution decisions. Both experiments have a Production and a Redistribution stage, and in the Production stage, participants perform a simple letter encryption task Benndorf et al. (2018). In the *Difficulty experiment*, I randomly assign easy, medium or hard tasks to subjects, while in the *Ability experiment*, everyone has similar, medium-difficulty tasks. I first measure participants' production potential – the number of tasks they can do within a minute with maximal effort – by asking them to do ten tasks as fast as possible. Based on this measure, I classify participants into ability levels (low, medium, high) in the Ability experiment, analogously to the difficulty levels (easy, medium, hard) in the Difficulty experiment. Next, participants work on the task for 15 minutes and receive ten experimental tokens per task completed. In the Redistribution stage, they decide how to redistribute the joint income of a randomly formed pair. They are informed about the performance and the difficulty or ability level (depending on the experiment) of both participants in the pair. They make decisions in either a *spectator* or a *stakeholder* role: Spectators redistribute the income of two subjects as an outside observer, while stakeholders redistribute their and another subject's joint income.

Section 3 presents descriptive statistics about both stages. In the Difficulty experiment, both production potential and performance within 15 minutes decrease in task difficulty. Similarly, low-ability participants perform significantly worse during the 15 minutes than medium- and high-ability participants in the Ability experiment. Though both factors affect performance, subjects only compensate for external difficulties in their redistributive decisions. In the Difficulty experiment, spectators and stakeholders redistribute around five percent of the total income to participants with more difficult tasks within the pair. In the case of equal difficulties, there is no redistribution on average.[3] In contrast, neither spectators nor stakeholders compensate the participant with lower ability in the Ability experiment: the average redistribution is zero at both equal and unequal ability levels.

Section 4 explores the results in a reduced-form regression framework. In this section, I also investigate heterogeneities in spectators' treatment of ability differences depending on their views on talent and ability. Those who think it is entirely unfair if talent determines income inequality do compensate the low-ability participant in the pair. However, most people regard talent as a fair source of inequality, and there is no compensation from their side, hence the average zero compensation. Furthermore, I study whether stakeholders make self-serving redistributive decisions depending on whether they have more or less difficult tasks or higher or lower ability within the

[3]There is no redistribution to a *randomly chosen participant* within the pair when difficulties are equal. There is, however, redistribution on average *to the participant with lower performance* because a fraction of participants follow the egalitarian principle and want to reduce inequality regardless of its source.

pair. The results in the Difficulty experiment are in line with the literature finding self-serving redistributive decisions among stakeholders (Rodriguez-Lara and Moreno-Garrido, 2012; Eisenkopf et al., 2013; Deffains et al., 2016; Fehr and Vollmann, 2022). Although all stakeholders redistribute income towards themselves, those with more difficult tasks take an extra amount as compensation for their disadvantage. In contrast, those with easier tasks in a pair do not compensate the other participant. They take as much as stakeholders in equal-difficulty pairs as if their performance only resulted from hard work. Interestingly, lower ability in the Ability experiment does not serve as a justification for stakeholders to compensate themselves. On the other hand, high-ability stakeholders take more than other stakeholders, suggesting that they believe high ability deserves extra reward on top of what they get for their performance.

Finally, in Section 5, to uncover the heterogeneity in individual fairness preferences behind the average decisions, I structurally estimate the shares of different fairness views among spectators and stakeholders (following Mollerstrom et al. 2015 and Andre 2024). I build on the model of fairness views in Almås et al. (2010) that distinguishes between libertarian, egalitarian and meritocratic views. I extend the model to include those compensating for inequality in either external difficulties or abilities. I find that the majority holds a performance-meritocratic view, i.e., accepting performance-based inequality regardless of potential differences in external difficulties or abilities. However, there is a large share of both spectators and stakeholders that compensate for external difficulties. The structural results also support the reduced-form ones by showing a zero share of those compensating for ability differences.[4] The share of egalitarians, who always equalize the incomes no matter the source of the inequality, is similar in both experiments to those found in the literature.

This paper contributes to the literature addressing the factors people hold each other responsible for. The existing literature finds a substantial prevalence of the meritocratic fairness view: many people find inequalities based on merit – or in the experimental papers, performance – and personal choices fair, but inequalities due to pure luck unfair (see a review in Cappelen et al., 2020a). However, luck usually operates in more subtle ways. Even if someone advances through their performance, performance is often shaped by inequalities in external factors – such as incentives, external circumstances or educational opportunities – or internal ones – like ability or productivity. Compensating for unequal external factors aligns with procedural fairness (Trautmann, 2022), but compensating for unequal ability is not as straightforward. People generally find performance-

[4]Since I address the two factors separately, I can only identify if someone compensates for either external difficulties or low ability. In reality, someone may compensate for both or only one of the two, but I cannot observe behavior towards both factors simultaneously.

based inequality fair when it comes from specific abilities or knowledge (Cappelen et al., 2010; Eisenkopf et al., 2013; Durante et al., 2014). However, ability often depends on prior access to information, practice of the task, or innate talent – factors that are also heavily influenced by luck and circumstances (Rawls, 1971; Eisenkopf et al., 2013).

Few experimental papers address the role of external circumstances in performance. Andre (2024) finds that spectators see performance-based reward as fair even with unequal incentives. A significant share, however, cares about unequal incentives once they learn the counterfactual performance under equal incentives. In Dong et al. (2022), workers either receive unequal education before the quiz determining their income or face quizzes of different lengths, affecting the total number of possible correct answers. They find that spectators redistribute more toward the disadvantaged worker in both treatments compared to the equal opportunities benchmark but less compared to when outcomes are purely luck-based. Eisenkopf et al. (2013) also study situations where participants have unequal education opportunities before a quiz determining their income. They find that participants treat unequal opportunities similarly to luck when they have a long time to study the questions beforehand. With only a short learning time, educated and uneducated participants differ significantly in whether they attribute their performance to luck or skill and, accordingly, in their redistribution decisions. Preuss et al. (2024) compare luck-based outcomes to those affected by lucky opportunities, where workers' performance might receive an unequal multiplier. They also find that redistribution is significantly lower with lucky opportunities than with lucky outcomes.

I add to these papers first by studying another type of unequal opportunity: similar to disturbing external circumstances while working that make the task at hand more difficult, some participants are assigned easier while others harder tasks, but ultimately, the reward goes for their final performance. By measuring the production potential at the beginning and displaying it during the redistributive decisions, participants can see exactly how difficult the task was for the pair and, thus, how much effort they must have exerted to achieve their performance. Providing this information allows me to separate whether participants reward each other for their effort or performance with both the reduced-form and the structural estimations.

Even fewer studies attempt to distinguish between effort and ability or productivity as factors behind performance. Cappelen et al. (2010) define participants' effort choice as their chosen working time and their productivity as the number of tasks done per minute of the working time. They find that only a small, non-significant share of participants hold others responsible for their choice of working time but not for their productivity. Schildberg-Hörisch et al. (2023) use a similar measure of productivity to mine: they measure how many tasks people can do when given only a short time,

arguing that people in experiments tend to exert maximum effort in short time frames. They use this measure to see whether participants view an affirmative action policy favoring low-productivity individuals as fair. Their results show that all other affirmative action policies studied (favoring those facing discrimination and those who choose a shorter working time) are perceived as fairer. Moreover, affirmative action favoring low-productivity individuals and no affirmative action are viewed as equally fair. Pogliano (2024) takes a different approach and uses a mathematical task in which innate talent plays a role. He then provides an information treatment, telling spectators about the large or small role of genetics in math performance. In contrast to the previous findings, he finds that those spectators who learned that genetics play a major role in workers' performance redistribute significantly more towards the less productive worker than those who learned about a minor role of genetics.

My paper also adds to both strands of literature by directly comparing how participants treat external difficulties and unequal abilities in their redistributive decisions. In the Ability Experiment, as well, I first measure the production potential. Then, I show participants' ability group together with the production potential during the redistributive decisions to help them see how much effort was needed for the particular performance of the pair. This way, compensating for external difficulties and ability is directly comparable. Ability in this task likely depends on how fast someone types, how fast they find patterns on the screen, and whether they have already encountered this task. Since the Benndorf et al. (2018) task is relatively new, it is unlikely that many participants have seen it before. Therefore, they did not know how good they were at it when they agreed to participate in the experiment. Though the fact that someone is better or worse in this particular task is also affected by luck, participants treat the two factors very differently. Many people see having to exert higher effort for the same performance due to external circumstances as unfair. However, having to exert higher effort simply because of a lower ability is considered completely fair. These results can explain that even though there exist policies to compensate students from lower socioeconomic backgrounds or with serious health disadvantages, e.g., dyslexia, differences in ability that cannot be traced back to a particular illness or disability are not seen as something to be compensated.

Finally, the papers mentioned above look at either spectator decisions, without own payoffs at stake, or stakeholder decisions, where participants' decisions may directly impact their payoffs. I contribute to these papers by comparing the decisions made in the two roles in the same experiment and same task. According to the theory described in Cappelen et al. (2020a), stakeholders act on the same fairness views as spectators. The papers documenting self-serving redistributive decisions (e.g.,

Eisenkopf et al., 2013; Deffains et al., 2016) suggest otherwise. My results show that, on average, stakeholders' and spectators' behavior towards the two factors are indeed similar. However, when looking at redistributive decisions depending on one's own advantage or disadvantage, the treatment of unequal external difficulties and abilities is very different. This result limits the extent of possible compensating policies, as people might agree on compensating external circumstances, but not at their own expense.

2 Experimental design

I study whether people compensate for external difficulties and low ability by running two parallel experiments. The Difficulty experiment addresses unequal external circumstances, and the Ability experiment addresses unequal ability. Both have the same structure: There is a Production stage and a Redistribution stage. The main stage concerning the question of this paper is the second stage, where participants make redistributive decisions. The Production stage is needed to create situations with different sources of inequality. I explain the two stages separately, highlighting the differences between the Difficulty and Ability experiments.

2.1 Production

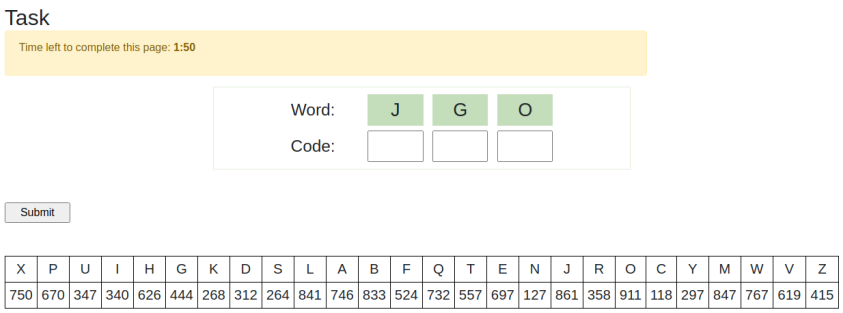
In this stage, participants work on a letter encryption task developed by Benndorf et al. (2018). See an example in Figure 1. The task consists of encrypting letter combinations to numbers, where the letter-number pairs and the order of the pairs in the encryption key are randomized between each correctly solved task. Benndorf et al. (2018) find that this double randomization minimizes learning in the task. The little scope for learning ensures that unequal ability to learn does not affect performance in the task, which is crucial for the way I separate production potential and effort made during production.

I measure participants' production potential by asking them to do ten tasks as fast as possible.[5] In the Difficulty experiment, I randomly assign participants a task length – two, three or four letters – to create external variation in task difficulty. Therefore, participants' production potential depends significantly on their assigned task length. In the Ability experiment, every participant works on three-letter tasks. Even though the task is simple and does not require specific skills, if someone types faster, has better short-term memory, or can find patterns on the screen faster, she can be better at it. So in the Ability experiment, how fast participants solve ten tasks depends on

[5]One three-letter task takes around 10-15 seconds to complete.

their ability in the task (this measure is similar to the baseline productivity measure in Schildberg-Hörisch et al., 2023).[6]

After measuring the production potential, participants work on the task for 15 minutes and receive ten experimental tokens for each correctly solved task. Participants in the Difficulty experiment work on tasks of their assigned length, while participants in the Ability experiment work on medium-length tasks. The number of tasks participants complete within 15 minutes is their production.



X	P	U	I	H	G	K	D	S	L	A	B	F	Q	T	E	N	J	R	O	C	Y	M	W	V	Z
750	670	347	340	626	444	268	312	264	841	746	833	524	732	557	697	127	861	358	911	118	297	847	767	619	415

Figure 1: Example (three-letter) task

Note: Benndorf et al. (2018) task. In the example task, the letter J corresponds to the number 861, G to 444, and O to 911, so the participant has to enter these three numbers into the boxes.

2.2 Redistribution

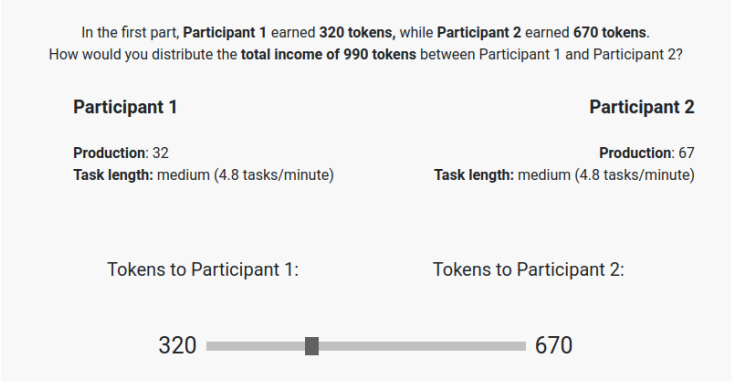
The Redistribution stage is the stage of interest. Participants in the same experiment are randomly assigned into pairs and decide on redistributing the income earned in the first stage. Every participant is either a *spectator* or a *stakeholder*, where spectators redistribute the income of another pair of subjects, and stakeholders redistribute income within their pair.[7] Spectator decisions show what participants find fair in a given situation, while stakeholder decisions also involve selfishness concerns since their monetary gains are at stake (Cappelen et al., 2020a).

Participants make redistributive decisions knowing all the details of how the incomes emerged. They know the production of both participants in the pair and the average production potential of their difficulty or ability level, depending on the experiment. The three task lengths in the Difficulty experiment determine the three difficulty levels. To have a comparable measure in the Ability experiment, I classify participants into ability terciles (low, medium and high) based on how fast they solved the ten tasks.

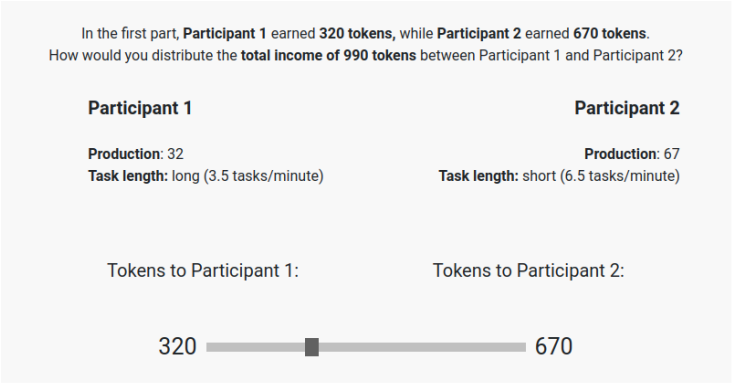
[6] See a discussion on how well this measure reflects true ability in Appendix Section A.4.

[7] The randomization occurs at the pair level, so each pair is either a spectator or a stakeholder pair.

To elicit participants' preferences over different sources of inequality, they make ten decisions about ten different pairs. I study compensation for the disadvantaged participant – i.e., the one with harder tasks or lower ability – by comparing decisions made over pairs where the difficulty or ability levels are equal (Figure 2a) to those where they are unequal (Figure 2b). Those participants who accept performance-based inequality leave the income distribution unchanged in these two situations. Those who wish to eliminate all inequalities equalize the incomes in both situations. However, those who want to compensate the disadvantaged participant leave the incomes unchanged if the difficulty or ability levels are equal but redistribute income to the disadvantaged participant if they are unequal.



(a) Equal difficulty levels



(b) Unequal difficulty levels

Figure 2: Example decision screens

Note: Example spectator decision screen in the Difficulty experiment. At stakeholder decisions, decision-makers' data was presented under 'You'. Spectators and stakeholders in the Ability experiment saw 'ability group' instead of 'task length' on the screen, with labels 'low', 'medium' and 'high' instead of 'long', 'medium' and 'short'.

2.3 Technical details

The experiment ran online in three sessions between 1-9 December 2021, with 100-250 participants per session. I recruited participants on Prolific and coded the experiment using oTree (Chen et al., 2016). The experimental design, the hypotheses, and the main empirical analyses were pre-registered in the AEA RCT Registry (Drucker, 2021). Appendix Section A.11 presents all experimental instructions the participants received. I recruited participants currently living in the United States (see Appendix Table A.1 for demographic data of the participants). They earned about 8.5 USD on average in total.

The two stages – Production and Redistribution – took place on two consecutive days and took together around 35 minutes to complete. Participants did the Production stage on the first day; the next day, they could return to do the Redistribution stage at any time between 6 AM and midnight. Separating the two stages allowed subjects not to have to sit in front of the computer simultaneously, excluding attrition due to having to wait for other subjects' moves. On the other hand, it introduced attrition between the two stages. There were 594 participants in the Production stage and 500 in the Redistribution stage, which means there was 16 percent attrition between the two days. However, not returning to the second stage did not correlate with any feature of the experiment or any demographic characteristic (see Appendix Table A.2).

Participants learned at the beginning of the Production stage that their income from this stage (except for the show-up fee) would not be their final income, but it might change according to their or other participants' decisions in the Redistribution stage. I followed the literature (e.g., Durante et al., 2014; Rey-Biel et al., 2018; Cappelen et al., 2022) in not revealing further details about the Redistribution stage until they arrived at that stage to minimize the effect of knowing about a possible redistribution on production.[8] Participants in the Difficulty experiment did not know about other task lengths until the Redistribution stage, so the effect of unequal difficulties on effort would be negligible.[9] Before consenting to participate, participants completed a comprehension check to ensure that they understand the experiment and the payment structure.

At the beginning of the Redistribution stage, participants learned about their task performance in the first stage. This was also when they learned about other possible task lengths in the Difficulty experiment and the ability groups in the Ability experiment. Here, I explained that other participants might have had an easier or harder job than them due to their task difficulty or ability

[8]Erkal et al. (2011) find that knowledge about subsequent redistribution induces more selfish participants to self-select to high-income ranks by exerting high effort and then choosing much less redistribution than other participants.

[9]Indeed, the number of individual letters encrypted did not differ significantly across the three difficulty levels.

levels (see the exact wording in Appendix Section A.11). Then, participants made redistributive decisions using the strategy method: They made ten decisions corresponding to ten random pairs from the first stage, but only one was the true pair assigned to the participant. Participants knew this but did not know which one was the true one. Two participants made a redistribution decision for all pairs – either the spectators assigned to that pair or the participants constituting the pair as stakeholders. At the end of the experiment, one decision out of the two was chosen randomly and implemented. I converted the experimental tokens to British Pounds (at 250 tokens = £1), the currency in which Prolific participants are paid.

At the end of the experiment, participants completed a survey about how hard they worked on different parts of the experiment, how fair they found different sources of inequality (questions from Cappelen et al., 2022), and their views regarding how changeable talent and ability are (based on Dweck, 2006). Appendix Figure A.1 shows the exact survey questions.

3 Descriptive statistics

3.1 Production

The data from the Production stage shows that differences in both external difficulties and ability induce inequality in performance. Table 1 shows the mean production potential at each difficulty and ability level, measured by how quickly participants solved ten tasks when working as fast as possible. The exogenous task length induced higher variation in production potential than people's abilities. Still, the differences across levels are large: high-ability participants could do, on average, 1.62 times more tasks within a minute than low-ability ones, while participants with easy tasks could do 1.87 times more than those with hard tasks. The means at the Medium difficulty/ability level are the same (4.74 tasks) across the experiments since there were three-letter tasks at both levels. Easy tasks allowed people to do more tasks within a minute than high ability in three-letter tasks (6.54 vs. 6.00, $p = 0.000$), and hard tasks slightly less than low ability in three-letter tasks (3.49 vs. 3.72, $p = 0.000$).

Table 2 reveals that the differences in production potential across difficulty and ability levels translate to differences in production, i.e., the number of tasks done within 15 minutes. The variance of production is also higher in the Difficulty experiment than in the Ability experiment. However, the ability-induced differences are also large and significant. In the Difficulty experiment, the mean production of participants with hard tasks is 47.6, with medium tasks 68.4, and with easy tasks 96.1. The means in the Ability experiment are 55.8 for low ability, 69.4 tasks for medium, and 83

Level	Difficulty experiment	Ability experiment
High diff./ Low ab.	3.489	3.715
Medium diff./ab.	4.744	4.737
Low diff./ High ab.	6.539	6.001

Table 1: Mean production potential by difficulty and ability level

Note: Each cell shows the average production potential (tasks/minute) measured by how fast people do ten tasks when asked to work as fast as possible. In the Difficulty experiment, the levels are determined by the assigned task length – high difficulty means four-letter tasks and low difficulty two-letter tasks. In the Ability experiment, the levels are terciles of the distribution of participants' time to solve the ten tasks. Low ability means the slowest tercile and high ability the fastest.

for high ability.

Level	Difficulty experiment	Ability experiment
High diff./ Low ab.	47.61	55.79
Medium diff./ab.	68.38	69.41
Low diff./ High ab.	96.06	82.97

Table 2: Mean production by difficulty and ability level

Note: Each cell shows the average production (tasks within 15 minutes) at a difficulty or ability level. In the Difficulty experiment, the levels are determined by the assigned task length – high difficulty means four-letter tasks and low difficulty two-letter tasks. In the Ability experiment, the levels are terciles of the distribution of participants' time to solve ten tasks when asked to work as fast as possible. Low ability means the slowest tercile and high ability the fastest.

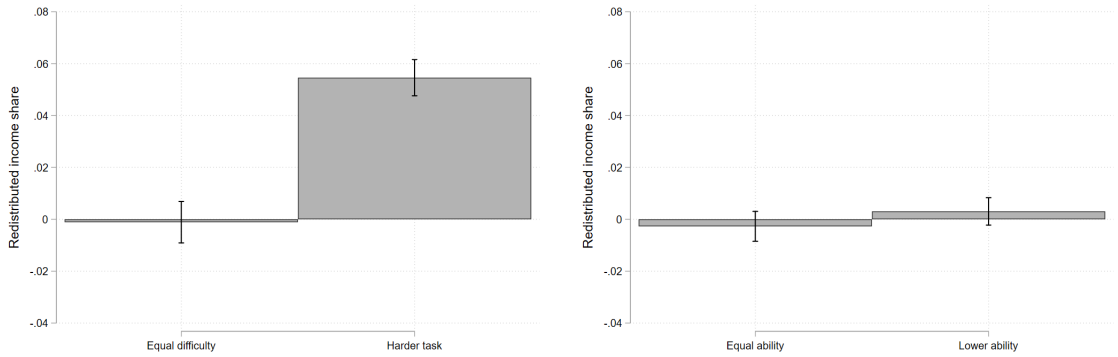
3.2 Redistribution

The paper's central question is how relative difficulties and abilities within the pair affect allocation decisions. Figures 3a and 3b present the average spectator decisions in the two experiments. Both figures distinguish two situations: equal and unequal difficulty/ability levels within the pair.[10] In the Difficulty experiment, participants with harder tasks within the pair received a 5.5 percentage points higher income than their production share. In decisions with equal difficulty levels, there is no redistribution on average.[11] In contrast, in the Ability experiment, we see no compensation for differences in ability: participants, on average, do not get more or less than their production share.

[10] See Appendix Figures A.6-A.9 for the distributions of production share and the income share redistributed in all situations.

[11] Note that, in case of unequal difficulties (abilities), the participant with harder tasks (lower ability) is likely also the participant with lower production in the pair. In case of equal difficulties (abilities), the randomly chosen participant is equally likely to be the one with higher or lower production. The regressions in Section 4 correct for this difference by controlling for the production share of the randomly chosen participant.

Even though the differences in production across difficulty levels were lower in the Ability experiment than in the Difficulty experiment, there would still have been scope for compensation. The average production share of the participant with longer tasks in the Difficulty experiment is 0.39, while the average production share of the participant with lower ability in the Ability experiment is 0.43. The average allocated income share to these participants is 0.44 in both experiments, suggesting that participants want to reduce inequalities but consistently up to the level determined by ability and effort.

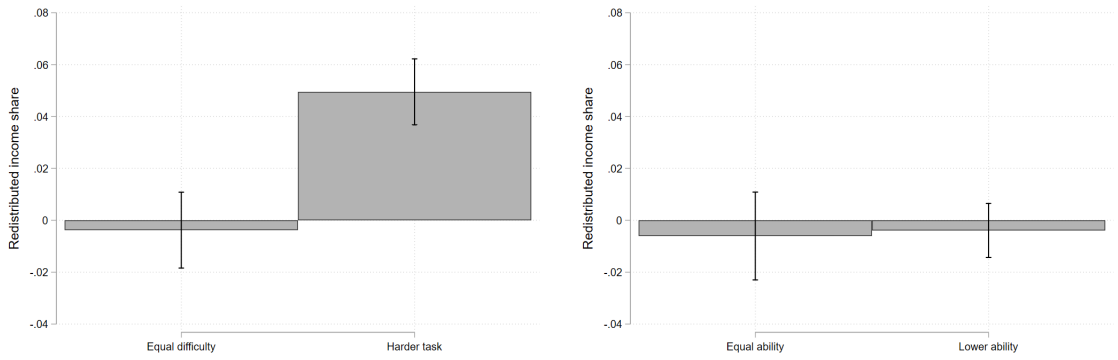


(a) Difficulty experiment (b) Ability experiment

Figure 3: Redistributed income share by spectators

Note: Spectator decisions. The figures show the redistributed income share (allocated income share minus production share) to a randomly chosen participant in the pair in two situations: with equal difficulty levels and with unequal difficulty levels. The spikes indicate 95 percent confidence intervals.

Figure 4 shows the redistributed income share to a randomly chosen participant by stakeholders in the same two situations: with equal and unequal difficulties. Here the random participant is either the decision-maker stakeholder or the other participant in the pair. The average redistribution by stakeholders is similar to that of spectators: stakeholders redistribute five percent of the total income to the participant with harder tasks, and there is no significant redistribution with equal difficulty levels (Figure 4a). In the Ability experiment, just like with spectators, there is no redistribution on average in either situation (Figure 4b). In the next section, I first explore these relationships in a reduced-form regression framework to see if the results are robust to different specifications. Then I study heterogeneities in the effect of unequal difficulties and abilities on stakeholders' decisions.



(a) Difficulty experiment

(b) Ability experiment

Figure 4: Redistributed income share by situation (stakeholder decisions)

Note: The figures show the redistributed income share (allocated income share minus production share) to a randomly chosen participant in the pair in two situations: with equal difficulty levels and with unequal difficulty levels. The decisions are made by stakeholders. The spikes indicate 95 percent confidence intervals.

4 Reduced-form analysis

4.1 Average compensation for higher difficulty or lower ability

I first look at the redistribution to the participant in the pair who had a higher difficulty (or lower ability) level compared to situations with equal difficulty/ability levels. For both spectators and stakeholders, the outcome variable is the redistributed income share – allocated income share minus production share – to a participant in the pair who is either the disadvantaged participant or a randomly chosen equally advantaged participant (Participant 1 or P1 henceforth). I estimate the following equation:

$$r_{1,i,p} = \alpha_0 + \alpha_1(\theta_1 < \theta_2)_p + \alpha_2 \cdot x_{1,sh,p} + \epsilon_{i,p}, \tag{1}$$

where $r_{1,i,p}$ is the share of income spectator or stakeholder i redistributes to Participant 1. θ_i is the average production potential of i, so $\theta_1 < \theta_2$ denote the cases where P1 had a higher difficulty or lower ability level than P2. $x_{1,sh,p}$ is the production share of P1 in pair p.

The first four columns of Table 3 present the results for spectators in the Difficulty experiment. The first column only adds a session fixed effect as control. Spectators, on average, redistribute 5.7 percent of the total income to the participant with more difficult tasks in the pair. Column 2 controls for the production share of the participant to see if, comparing two participants with equal production shares, spectators still redistribute more to the one with harder tasks. Column 3 adds basic demographic controls: age, gender, whether the spectator was born in the US, and whether she

has US nationality.[12] Column 4 adds participant fixed effects to control for any individual-specific allocation behavior that does not depend on the production share or relative difficulties. The effects of relative difficulties remain: spectators redistribute a 3.9-4.5 percentage points higher income to the participant with more difficult tasks at a given production share.

The last four columns cover spectator decisions in the Ability experiment. Unlike unequal difficulties, unequal abilities do not matter in the decisions of spectators, either on average or when controlling for the production share. Some spectators may want to compensate low-ability participants, while others think these participants did not exert high effort and take money from them, resulting in a null effect on average. I will explore such potential heterogeneities in Section 4.2.

	Difficulty experiment				Ability experiment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Situation, ref. Equal difficulty/ability level								
Higher difficulty / Lower ability level	0.0570***	0.0393***	0.0406***	0.0446***	0.00594	-0.000446	0.0000890	0.00164
	(0.00724)	(0.00639)	(0.00638)	(0.00741)	(0.00498)	(0.00457)	(0.00441)	(0.00501)
Constant	-0.00564	0.0743***	0.0825***	0.0764***	-0.00585	0.0461*	0.0135	0.0422*
	(0.00930)	(0.0172)	(0.0206)	(0.0134)	(0.00994)	(0.0274)	(0.0331)	(0.0229)
Observations	1210	1210	1210	1210	1170	1170	1170	1170
Share in total production	no	yes	yes	yes	no	yes	yes	yes
Participant fixed effect	no	no	no	yes	no	no	no	yes
Demographic controls	no	no	yes	no	no	no	yes	no
Session fixed effect	yes	yes	yes	no	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Spectator decisions

Note: The outcome variable is the excess income share given to one participant in a pair on top of her production share by spectators in the Difficulty experiment (1-4) and the Ability experiment (5-8). Columns 1 and 5 only control for whether the participant had more difficult tasks or lower ability than the other participant. Columns 2-4 and 6-8 also control for the production share of Participant 1. Columns 3 and 7 include demographic controls: age, gender, whether the spectator was born in the US, and whether she has US nationality. Columns 4 and 8 add participant fixed effects.

The first four columns of Table 4 look at stakeholder decisions in the Difficulty experiment. Column 1 shows that, on average, stakeholders also redistribute about 5.2 percent of the total income to P1 if she had a higher difficulty level. However, this result is not robust to controlling

[12]Student status and employment status were also available in the Prolific database. However, there is a larger share of missing values in these variables, as these data had expired for some participants by the time of the experiment. Appendix Tables A.6 and A.7 show that the results are robust to using only participants with complete demographic data.

for the production share of the participant. The redistribution is still positive but much lower and non-significant. This result suggests that there may be heterogeneities in how stakeholders treat unequal difficulties depending on whether they are advantaged or disadvantaged. I will explore these heterogeneities in Section 4.3.

The last four columns of Table 4 present the results for stakeholders in the Ability experiment. Similarly to spectators, stakeholders do not redistribute income to participants with lower ability either.[13] If anything, the coefficients on lower ability are negative (but non-significant), as if stakeholders redistributed towards the higher-ability participant instead of compensating the lower-ability one. Section 4.3 also discusses heterogeneities of stakeholder decisions in the Ability experiment.

	Difficulty experiment				Ability experiment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Situation, ref. Equal difficulty/ability level								
Higher difficulty / Lower ability level	0.0524***	0.0132	0.0138	0.0135	0.00224	-0.0287	-0.0296	-0.0276
	(0.0138)	(0.0170)	(0.0172)	(0.0160)	(0.0146)	(0.0180)	(0.0179)	(0.0213)
Constant	0.0151	0.188***	0.220***	0.145***	-0.00182	0.254**	0.201**	0.214**
	(0.0213)	(0.0473)	(0.0514)	(0.0394)	(0.0146)	(0.0985)	(0.0991)	(0.0922)
Observations	1250	1250	1250	1250	1210	1210	1210	1210
Share in total production	no	yes	yes	yes	no	yes	yes	yes
Participant fixed effect	no	no	no	yes	no	no	no	yes
Demographic controls	no	no	yes	no	no	no	yes	no
Session fixed effect	yes	yes	yes	no	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Stakeholder decisions

Note: The outcome variable is the excess income share given to one participant in a pair on top of her production share by stakeholders in the Difficulty experiment (1-4) and the Ability experiment (5-8). Columns 1 and 5 only control for whether the participant had more difficult tasks or lower ability than the other participant. Columns 2-4 and 6-8 also control for the production share of Participant 1. Columns 3 and 7 include demographic controls: age, gender, whether the stakeholder was born in the US, and whether she has US nationality. Columns 4 and 8 add participant fixed effects.

[13] Appendix Tables A.8 and A.9 look at redistribution as the share of the original income difference to address the concern that differences in production across ability levels were lower than across difficulty levels. The results for spectators are similar: they redistribute 40-50 percent of the original income difference to the disadvantaged participant in the Difficulty experiment, while the coefficients are low and insignificant in the Ability experiment. The redistribution toward the disadvantaged participant by stakeholders is insignificant in both experiments, though the magnitudes are still large in the Difficulty experiment.

4.2 Heterogeneous treatment of ability

Participants may differ in how they think about ability: some might think ability is fixed, so one should not be held responsible for it. Others may believe that ability is malleable, so participants' ability in the task is entirely under their control. These people will not want to compensate for low ability. To elicit views on ability, participants answered some questions about talent and ability in the survey at the end of the experiment. They had to rate on a zero to ten scale how strongly they agreed with statements. The first question is part of three questions regarding how fair they find different sources of inequality (from Cappelen et al., 2019). The second and third questions are proxies for how malleable people think talent and ability are (growth mindset questions based on Dweck, 2006). These statements were

1. "I find it fair if talent determines income inequality."
2. "The harder I work on something, the better I will be at it."
3. "Talent in an area is something about me that I can't change very much."[14]

Appendix Figures A.11, A.13 and A.14 show the distribution of the ratings of participants. On average, participants find it relatively fair if talent determines income inequality (mean = 6.76, median = 7). They find talent somewhat less fair than hard work as a determinant of income inequality (mean = 7.38, median = 8), and much fairer than luck (mean = 2.68, median = 2), consistent with the findings of Cappelen et al. (2022). Most participants agree with the statement that the harder they work on something, the better they become at it (mean = 8.1, median = 8), while the opinions vary greatly about talent being something that cannot change (mean = 4.63, median = 5).[15] My hypotheses were that people who find it rather unfair if talent determines income inequality will compensate the person with lower ability. Also, those who think hard work can make them better at something might regard ability as a result of past efforts and not compensate for it. In contrast, those who think talent cannot be changed would compensate for low ability because they believe it is not under one's control.

Table 5 presents the spectators' redistribution towards the lower-ability participant depending on views about talent and ability. There are indeed differences in the decisions by different views.

[14]The last two questions were not pre-registered, and I only added them to the second and third sessions, so this part of the analysis is rather exploratory.

[15]The distributions of spectators' answers are similar across ability levels (see Appendix figures A.15-A.17). If anything, high-ability spectators seem to agree more, on average, with talent being a fair source of inequality than medium-ability ones (6.88 vs. 5.51, $p = .049$). The difference between low- and high-ability spectators' average opinions is insignificant (6.44 vs. 6.88, $p = .441$).

Column 1 shows the average (non-)compensation for low ability among spectators. According to Column 2, however, those who disagree with talent being a fair source of income inequality redistribute 2.5 percent of the tokens to the lower-ability participant. Each one-point increase in agreement decreases the compensation by 0.3 percentage points. Column 3 shows that the more a spectator agrees that hard work makes her better at something, the less redistribution she chooses on average. This behavior aligns with the view that ability results from past effort, so performance-based income inequality is fair. However, there is no difference in the choices depending on whether ability levels within the pair are equal or unequal. Finally, the more spectators agree with talent being malleable (Column 4), the more they give to the lower-ability participant, in contrast to what I expected. However, this relationship is weak, only significant at the ten percent level.

4.3 Self-servingness in stakeholder decisions

Previous literature has shown that when uncertain about whether luck – difficulty of the task – or effort determined success, successful and unsuccessful participants attribute it to different factors and choose redistribution levels accordingly. (Deffains et al., 2016; Fehr and Vollmann, 2022). Table 6 shows that, even with full information about the sources of income, participants with more and less difficult tasks treat inequalities differently and choose allocations that benefit them the most. Stakeholders with harder tasks in the pair always redistribute more to themselves than in situations with equal task lengths. In contrast, stakeholders with easier tasks do not redistribute less to themselves once we take production share into account. Controlling for production share, disadvantaged stakeholders compensate themselves to a similar extent as spectators compensate them. It is the stakeholders in advantaged situations who behave differently from spectators by disregarding unequal difficulties completely.

The self-compensation in disadvantageous situations remains even when including participant fixed effects (Column 4). This suggests that the magnitude of self-compensation depends on the relative difficulty within the pair and not the absolute one. Indeed, as Appendix Table A.10 shows, stakeholders with long tasks do not give themselves significantly more than those with medium or short tasks once controlling for the production share. These results indicate that circumstances alone do not induce stakeholders to behave differently. What matters is the type of the situation: stakeholders with long tasks take more when they have a justification – a disadvantaged position –, and stakeholders with short tasks keep more despite the transparent inequality in the difficulty levels.

Table 7 explores stakeholder decisions depending on their relative ability in the pair in the Ability

	Redistributed income share			
	(1)	(2)	(3)	(4)
Situation, ref. equal ability levels				
Lower ability	0.00594	0.0249**	-0.0364	-0.0156
	(0.00498)	(0.0125)	(0.0357)	(0.0110)
Finds fair if talent determines income inequality		0.000297		
		(0.00106)		
Lower ability × Finds fair if talent determines income inequality		-0.00302**		
		(0.00143)		
Hard work makes me better at something			-0.00415**	
			(0.00205)	
Lower ability × Hard work makes me better at something			0.00447	
			(0.00402)	
Talent in an area is something I can change				-0.00132
				(0.00105)
Lower ability × Talent in an area is something I can change				0.00304*
				(0.00166)
Constant	-0.00585	-0.00796	0.0363*	0.00853
	(0.00994)	(0.0119)	(0.0196)	(0.00820)
Observations	1170	1170	950	950
Participant fixed effect	no	no	no	no
Demographic controls	no	no	no	no
Session fixed effect	yes	yes	yes	yes

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Spectators' decisions by views on ability

Note: The outcome variable is the excess income share given to one participant in a pair on top of her production share by spectators in the Ability experiment. Column 1 shows the average compensation. Column 2 adds agreement with "I find it fair if talent determines income inequality". Column 3 adds agreement with "The harder I work on something, the better I will be at it". Column 4 adds agreement with "Talent in an area is something about me I *can* change." The scale is reversed here compared to the original question for easier interpretation. Agreement with the statements is on a 0-10 scale. The last two questions were only added to Sessions 2 and 3, hence the fewer observations in the last two columns.

experiment. In contrast to harder tasks, lower ability within the pair did not serve as a reason for stakeholders to compensate themselves. In contrast, when controlling for production shares, stakeholders with higher ability redistribute significantly more to themselves than stakeholders in other situations. This difference becomes insignificant when including participant fixed effects, suggesting that participants with high ability might differ from participants with low ability in this task. Indeed, as Appendix Table A.11 shows, the asymmetric treatment of unequal ability situations comes exclusively from high-ability stakeholders, who redistribute 8-9 percentage points more to

	Redistributed income share to self			
	(1)	(2)	(3)	(4)
Situation, ref. equally difficult tasks				
Stakeholder had harder tasks	0.0828***	0.0476**	0.0454**	0.0252***
	(0.0206)	(0.0201)	(0.0194)	(0.00824)
Stakeholder had easier tasks	-0.0214*	0.0154	0.0138	0.000523
	(0.0129)	(0.0161)	(0.0160)	(0.00861)
Constant	0.0300	0.196***	0.241***	0.210***
	(0.0213)	(0.0371)	(0.0702)	(0.0251)
Observations	1250	1250	1250	1250
Share in total production	no	yes	yes	yes
Participant fixed effect	no	no	no	yes
Demographic controls	no	no	yes	no
Session fixed effect	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Stakeholders' decisions by own relative difficulty

Note: The outcome variable is the excess income share stakeholders in the Difficulty experiment give to themselves on top of their production share. Column 1 controls for whether the stakeholder had easier or harder tasks than her partner in the pair (the baseline is equally difficult tasks). Columns 2-4 also control for the production share of the stakeholder. Column 3 includes demographic controls: age, gender, whether the stakeholder was born in the US, and whether she has US nationality. Column 4 adds participant fixed effects.

themselves than other stakeholders.

High-ability participants may be more selfish than lower-ability ones. However, it is also possible that classification as high-ability induced more selfish choices. Because the cutoffs between the ability levels differ a little across the three sessions I ran, I can look at whether participants with similar production potential near the cutoffs between medium and high ability levels redistribute differently depending on how I classified them.[16] Running the same regression on this very small sample (25 participants), the coefficient on high ability is still large and significant (0.064, $p < 0.01$ with non-clustered standard errors). This suggests that classification to the high-ability group indeed induced selfish choices, though I cannot exclude that high-ability participants are generally more selfish as well.

[16] This comparison was not pre-registered.

	Redistributed income share to self			
	(1)	(2)	(3)	(4)
Situation, ref. equal ability				
Stakeholder had lower ability	-0.00430	-0.0323	-0.0279	0.00563
	(0.0184)	(0.0233)	(0.0214)	(0.00736)
Stakeholder had higher ability	0.00183	0.0326**	0.0298**	0.0120
	(0.0155)	(0.0153)	(0.0143)	(0.00964)
Constant	0.0424*	0.299***	0.411***	0.141***
	(0.0253)	(0.0955)	(0.102)	(0.0265)
Observations	1210	1210	1210	1210
Share in total production	no	yes	yes	yes
Participant fixed effect	no	no	no	yes
Demographic controls	no	no	yes	no
Session fixed effect	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Stakeholders' decisions by own relative ability

Note: The outcome variable is the excess income share stakeholders in the Ability experiment give to themselves on top of their production share. Column 1 controls for whether the stakeholder had lower or higher ability than her partner in the pair (the baseline is equal ability). Columns 2-4 also control for the production share of the stakeholder. Column 3 includes demographic controls: age, gender, whether the stakeholder was born in the US, and whether she has US nationality. Column 4 adds participant fixed effects.

5 Structural analysis

We saw that spectators and stakeholders, on average, compensate in their decisions for more difficult tasks but not for lower ability. However, this average behavior masks very different individual preferences. The literature on fairness and inequality shows that people hold heterogeneous fairness views that they consistently act on – at least as spectators – in their choices (see a review in Cappelen et al., 2020a). In this section, I structurally estimate the shares of different fairness views among the participants.

I use the model of fairness preferences by Almås et al. (2010), applied to my setting, for the structural analysis. According to the model, a spectator with fairness preference type k finds the following allocation of joint income Y fair:

$$t_1^k(\mathbf{x},\, \theta) = \frac{f^k(x_1, \theta_1)}{f^k(x_1, \theta_1) + f^k(x_2, \theta_2)} Y(\mathbf{x}(\theta)) \tag{2}$$

Here $t_1^k(\mathbf{x},\, \theta)$ is the number of tokens a spectator with fairness preference type k finds fair to give to Participant 1 in the pair – in my analysis, a randomly selected participant of the pair.[17] $f^k(x_i, \theta_i)$ is a function that shows how a spectator with fairness preference k values the contribution of participant i. x_i is the production level of participant i, and θ_i is the average production potential at the difficulty or ability level of participant i.

The most common fairness views in the literature that can be distinguished in my sample are the following:

- *Performance-meritocratic:*[18] does not redistribute performance-based inequality

$$t_1^{PM}(\mathbf{x},\, \theta) = \frac{x_1}{x_1 + x_2} Y \tag{3}$$

- *Egalitarian:* always redistributes to equality

$$t_1^{E}(\mathbf{x},\, \theta) = \frac{1}{2} Y \tag{4}$$

Participants with the performance-meritocratic fairness view find inequalities based on performance fair, so they leave the performance-based inequality unchanged. Participants with the egalitarian

[17] Alternatively to a random participant from the pair, I could have used the share given to the participant with lower initial income (as in e.g., Almås et al., 2020). Since the allocations within a pair are symmetric, we can choose either definition without loss of generality.

[18] This view is usually called the meritocratic view, but I want to distinguish it from the compensating view, which is also partly meritocratic.

fairness view find any sources of inequality unfair, so they will redistribute to equality in all situations. A third common fairness view is the *libertarian* view which finds all inequalities, even those coming from pure luck, fair. In the current design, since there is no pure luck component in the income, people with this view make identical decisions to those with the performance-meritocratic view. I made this simplification because I was interested in deviations from the performance-meritocratic allocations.

To test if there are people who consistently follow a different rule by compensating for harder tasks or lower ability, I introduce a fourth (in the sample, third) fairness view:

- *Compensating meritocratic:*

$$t_1^{CM}(\mathbf{x},\, \theta) = \frac{x_1/\theta_1}{x_1/\theta_1 + x_2/\theta_2} Y \tag{5}$$

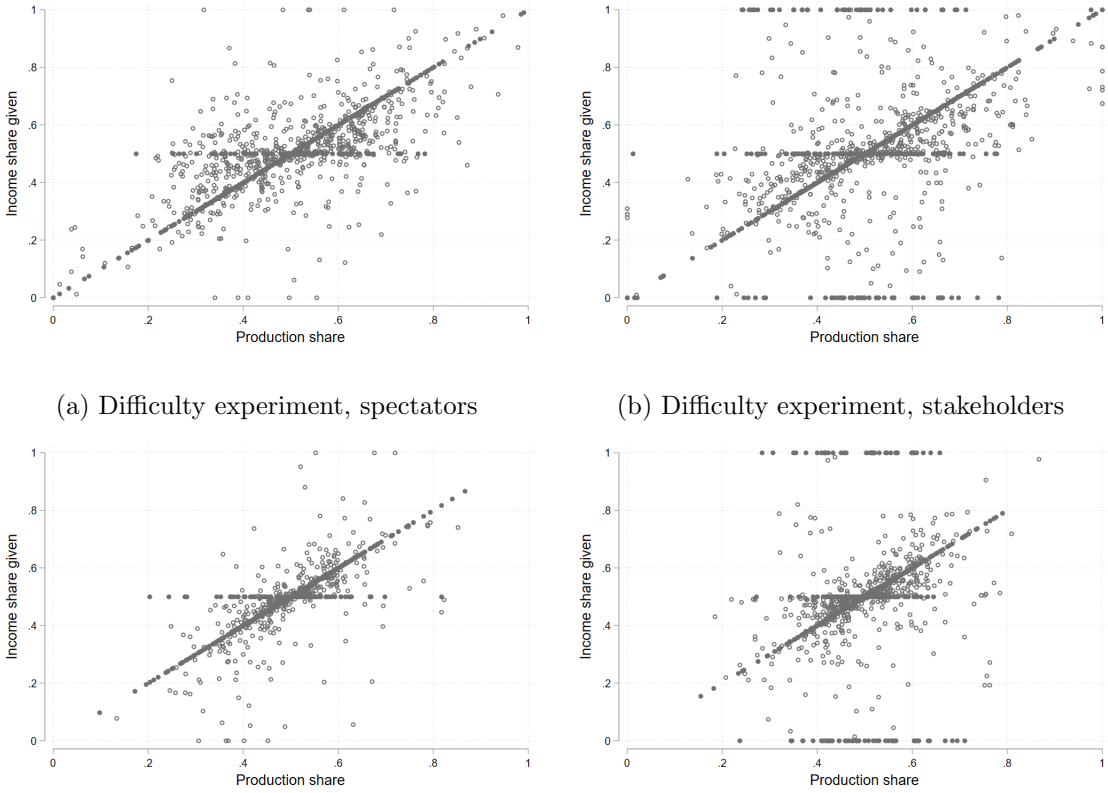
People with this fairness view want to reward performance but also want to compensate those who had a disadvantage in performance. I capture such preferences by assuming that the fair allocation is based on the production weighted by the average production potential of the person (x_i/θ_i). This assumption is a starting point for separating compensation from purely performance-meritocratic decisions, but I also test other functional forms in Appendix Section A.10.3.

If we look at redistributive decisions separately, they indeed show significant heterogeneity. Figure 5 presents the decisions of spectators and stakeholders in each experiment. The decisions are plotted as the share of the total income allocated to a randomly chosen participant in the pair against the same participant's share in the total production of the pair.

In the spectator figures, we can clearly distinguish two types of allocations: the egalitarian ones that equalize the income between the two participants and the performance-meritocratic ones that distribute the income proportional to the production share. A third type of decision is also salient in the stakeholder figures: distributing all or none of the total income to the random participant. Since the randomly chosen participant is either the decision-maker stakeholder or her partner, these decisions are, in fact, the ones where the stakeholder allocated all tokens to herself (see Appendix Figures A.5a and A.5b for decisions expressed as a share of income given to self).

A significant portion of decisions is outside these clearly defined shares, which, in the stakeholder figures, may only indicate partly selfish stakeholders giving slightly more to themselves than half or than their production share. However, in the spectator figures, these allocations suggest that some participants might follow rules other than the egalitarian or the performance-meritocratic ones. It is also visible that the allocations in the Ability experiment are less spread out and less different from the performance-meritocratic, egalitarian and entirely selfish allocations than in the

Difficulty experiment. In the structural estimation, I distinguish between two types of compensating meritocrats: difficulty-compensating meritocrats compensate for unequal external difficulties while ability-compensating meritocrats compensate for low ability. I can estimate the share of the first type in the Difficulty experiment and the second type in the Ability experiment.



(a) Difficulty experiment, spectators

(b) Difficulty experiment, stakeholders

(c) Ability experiment, spectators

(d) Ability experiment, stakeholders

Figure 5: Production share and allocated income share

Note: The figures show the share of tokens allocated to a randomly chosen participant in the pair by spectators and by stakeholders, plotted against the share of the same participant in the total production of the pair. The upper two figures show the spectator and stakeholder decisions in the Difficulty experiment and the lower two in the Ability experiment. One point indicates one decision. Filled circles indicate the decision types previously identified in the literature: performance-meritocratic, egalitarian and purely selfish.

5.1 Results

Tables 8a and 8b show the results of the structural estimation for spectators. Even with transparent inequalities in task difficulty, 55 percent of spectators hold a performance-meritocratic view and leave the performance-based inequality unchanged. Fifteen percent hold the egalitarian view and always redistribute to equality. However, a large share, 29 percent of the spectators, can be classified as difficulty-compensating meritocrats, who compensate the participant with more difficult tasks.

In Table 8b, we can see that if the only factor determining performance besides effort is ability, 87 percent of spectators leave the performance-based incomes unchanged. Thirteen percent make egalitarian decisions and redistribute to equality, and the estimated share of ability-compensating meritocrats is zero.[19]

Since ability differences alone induced lower differences in production than the exogenous task difficulty, there was less scope for compensating for low ability without fully equalizing the incomes. However, making ability differences salient could have induced more egalitarian decisions among the subjects. This was not the case: the shares of participants with the egalitarian view are similar across the two experiments.

	Estimate	Std. error
σ	112.813	2.284
λ_{PM}	0.553	0.047
λ_E	0.152	0.034
λ_{DCM}	0.294	0.046

(a) Difficulty experiment

	Estimate	Std. error
σ	82.905	1.716
λ_{PM}	0.871	0.032
λ_E	0.129	0.032
λ_{ACM}	0.000	0.004

(b) Ability experiment

Table 8: Structural estimation results – Spectators

Note: Spectator results. σ is the standard deviation of the response error, λ_{PM} is the estimated share of performance meritocrats, λ_E is the share of egalitarians, and λ_{DCM} and λ_{ACM} are the shares of difficulty-compensating and ability-compensating meritocrats. Parameters estimated in R using the `stats4` package. The parameters are estimated with unconstrained optimization and transformed within the likelihood function. Standard errors are calculated with the delta method.

Tables 9a and 9b present the results for stakeholders. The estimated shares are similar to the spectator shares: 55 percent of stakeholders hold the performance-meritocratic view, 25 percent are difficulty-compensating meritocrats, and 20 percent are egalitarians. The point estimate for the share of ability-compensating meritocrats is positive, 5 percent, but insignificant on any standard significance level.

The parameter β shows stakeholders' average weight on fairness compared to their monetary gain. On average, they place a considerable weight on fairness. However, seven percent of the stakeholders in the Difficulty experiment and 6.5 percent in the Ability experiment took all the money in at least eight of their ten decisions. It is, therefore, impossible to categorize them into any fairness type. Appendix Tables A.14a and A.14b show the estimates excluding these participants. In this not entirely selfish sample, the standard deviation of the response error is much smaller in

[19]The tables also show the estimated standard deviation of the response error of spectators, σ. To put the response error in context, the average tokens earned were around 720 in the Difficulty experiment and 700 in the Ability experiment.

both experiments, closer to the spectator estimates. The estimated share of difficulty-compensating meritocrats is slightly lower, 18 percent, in line with the finding that only disadvantaged stakeholders compensated themselves. The share of performance meritocrats is 60 percent and of egalitarians 17-23 percent. The share of ability-compensating meritocrats is virtually zero.

	Estimate	Std. error
σ	220.357	4.445
β	15.816	1.279
λ_{PM}	0.546	0.069
λ_E	0.206	0.049
λ_{DCM}	0.248	0.073

(a) Difficulty experiment

	Estimate	Std. error
σ	193.13	3.941
β	22.196	2.041
λ_{PM}	0.807	0.052
λ_E	0.139	0.047
λ_{ACM}	0.053	0.036

(b) Ability experiment

Table 9: Structural estimation results – Stakeholders

Stakeholder results. σ is the standard deviation of the response error, β is the average weight stakeholders put on fairness as opposed to their monetary gain, λ_{PM} is the estimated share of performance meritocrats, λ_E is the share of egalitarians, and λ_{DCM} and λ_{ACM} are the shares of difficulty- and ability-compensating meritocrats. Parameters estimated in R using the `stats4` package. The parameters are estimated with unconstrained optimization and transformed within the likelihood function. Standard errors are calculated with the delta method.

These results are consistent with the shares found in the literature while also identifying the new, compensating meritocratic types. Andre (2024), studying redistribution among workers with unequal incentives, finds in a sample representative to the US population 37 percent actual choice meritocrats, 23 percent libertarians, 14 percent egalitarians and 26 percent comparable choice meritocrats. The latter group redistributes income proportionately to participants' counterfactual production with equal incentives. In contrast, actual choice meritocrats leave the performance-based inequality unchanged no matter what the workers would have done with equal incentives. The share of actual choice meritocrats and libertarians together is roughly the same as the estimated share of performance meritocrats in my sample. This is consistent with the fact that performance meritocrats and libertarians make identical decisions in my setting. The shares of comparable choice meritocrats in Andre (2024) and difficulty-compensating meritocrats in my sample are also similar. Both groups aim to compensate people for factors that affect their performance but are outside their control. The estimated shares are also consistent with those found for the US sample in Almås et al. (2020).

6 Conclusion

In an online experiment with 500 participants from the United States, I studied if people redistribute performance-based income in situations with transparent inequality in external difficulties or abilities in the task. Spectators who decided about other participants' income compensated those with harder tasks but did not compensate those with lower ability. Stakeholders who decided about their own income, though behaving similarly to spectators on average, made very different decisions based on whether they were advantaged or disadvantaged: Stakeholders with harder tasks compensated themselves, while stakeholders with easier tasks did not compensate them. In contrast, stakeholders with high ability in the task behaved more selfishly than other stakeholders, while low ability did not justify taking more.

Although the core idea of meritocracy is that people should not be held responsible for factors outside their control, only for their choices, these are often not clearly separable. Circumstances influence choices or even limit them – as in the experiment, participants with longer tasks could not choose to do as many tasks as participants with shorter tasks. Andre (2024) finds that even though circumstances influence choices, spectators disregard unequal circumstances and hold the workers fully responsible for their choices. My result shows that when circumstances so transparently limit choices, a large share of people realize their effect on outcomes and redistribute accordingly. Interestingly, participants did not think about ability differences the same way. Although ability in this task can also result from lucky prior circumstances, participants held each other responsible for differences in their performance due to unequal ability. Stakeholders even rewarded themselves for high ability on top of their performance-based income.

These results have important implications. In situations where the person deciding about the income distribution is not one of the income recipients, such as a teacher rewarding students or a manager rewarding workers, learning more about external difficulties might help them make more aligned decisions with what they find fair. Without this information, they can only base their choices on observed performance, which is heavily influenced by external factors. However, if students or workers had to decide on their own reward, those with unequal difficulties or abilities would make conflicting decisions, even with similar fairness views and complete information about pre-existing inequalities.

The results of this paper lead to various potential directions for further research. First, it is worth exploring what kinds of abilities people find more exogenous, so compensating them by redistribution is seen as fairer than compensating for other abilities. Second, people's views on external

difficulties and abilities might be intertwined. For example, people may prefer more redistribution towards someone with disadvantageous circumstances if she exhibits high talent in something at the same time. Furthermore, sometimes, external circumstances are less transparent than in my experiment. It would be interesting to see how spectators redistribute with uncertainty about the role of difficulties and effort in performance. Cappelen et al. (2022) find among US and Norwegian participants that in situations where earnings could result from pure luck or performance, those with the meritocratic view make more egalitarian redistribution decisions. Since I could identify a share of participants who compensate for external difficulties, it would be worth studying if they make more meritocratic or more egalitarian decisions when the role of difficulties versus effort is uncertain. Finally, people in the United States hold more libertarian and less egalitarian views than Norwegians (Almås et al., 2020). These views align with higher redistribution in Norway than in the US. It would be interesting to look at whether the shares of difficulty-compensating and ability-compensating meritocrats are also higher in countries with higher redistribution levels than I found among US participants.

References

Alesina, Alberto and George-Marios Angeletos, "Fairness and Redistribution," *American Economic Review*, September 2005, *95* (4), 960–980.

Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden, "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?," *Journal of Political Economy*, 2020, (262675).

Almås, Ingvild, Alexander W. Cappelen, Erik Ø. Sørensen, and Bertil Tungodden, "Fairness and the Development of Inequality Acceptance," *Science*, 2010, *328* (5982), 1176–1178.

Andre, Peter, "Shallow Meritocracy," *The Review of Economic Studies*, 04 2024, p. rdae040.

Benndorf, Volker, Holger A. Rau, and Christian Sölch, "Minimizing learning behavior in repeated real-effort tasks," *Center for European, Governance and Economic Development Research Discussion Papers*, 3 2018.

Bradley, Robert, Robert Corwyn, Bettye Caldwell, Leanne Whiteside-Mansell, Gail Wasserman, and Iris Mink, "Measuring the Home Environments of Children in Early Adolescence," *Journal of Research on Adolescence*, 07 2000, *10*, 247–288.

Cappelen, Alexander W., Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden, "The Pluralism of Fairness Ideals: An Experimental Approach," *American Economic Review*, June 2007, *97* (3), 818–827.

_ , Erik Sørensen, and Bertil Tungodden, "Responsibility for what? Fairness and individual responsibility," *European Economic Review*, 2010, *54* (3), 429–441.

_ , Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden, "A Meritocratic Origin of Egalitarian Behavior," *SSRN Electronic Journal*, 2019, pp. 1–62.

Cappelen, Alexander W, Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden, "A Meritocratic Origin of Egalitarian Behaviour," *The Economic Journal*, 04 2022, *132* (646), 2101–2117.

Cappelen, Alexander W., Ranveig Falch, and Bertil Tungodden, "Fair and Unfair Income Inequality," *Handbook of Labor, Human Resources and Population Economics*, 2020, pp. 1–25.

_ , **Thomas de Haan, and Bertil Tungodden**, "Fairness and limited information: Are people Bayesian meritocrats?," 2020. Presentation.

Chen, Daniel L., Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88 – 97.

Deffains, Bruno, Romain Espinosa, and Christian Thöni, "Political self-serving bias and redistribution," *Journal of Public Economics*, 2016, *134*, 67–74.

Dong, Lu, Lingo Huang, and Jaimie Lien, ""They Never Had a Chance": Unequal Opportunities and Fair Redistributions," Discussion Papers 2022-11, The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham 2022.

Drucker, L. Flóra, "Meritocracy and Persistent Luck," December 2021.

Durante, Ruben, Louis Putterman, and J. van der Weele Joël, "Preferences for redistribution and perception of fairness: An experimental study," *Journal of the European Economic Association*, 2014, *12* (4), 1059–1086.

Dweck, Carol Susan, *Mindset: The New Psychology of Success.*, Random House Publishing Group, 2006.

Eisenkopf, Gerald, Urs Fischbacher, and Franziska Föllmi-Heusi, "Unequal opportunities and distributive justice," *Journal of Economic Behavior and Organization*, 2013, *93*, 51–61.

Erkal, Nisvan, Lata Gangadharan, and Nikos Nikiforakis, "Relative Earnings and Giving in a Real-Effort Experiment," *American Economic Review*, December 2011, *101* (7), 3330–48.

Falk, Armin, Fabian Kosse, Pia Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers, "Socio-Economic Status and Inequalities in Children's IQ and Economic Preferences," *Journal of Political Economy*, 04 2021, *129*.

Fehr, Dietmar and Martin Vollmann, "Misperceiving Economic Success: Experimental Evidence on Meritocratic Beliefs and Inequality Acceptance," *CESifo Working Paper*, September 2022, (9983).

Fong, Christina, "Social Preferences, Self-Interest, and the Demand for Redistribution," *Journal of Public Economics*, 11 2001, *82*, 225–246.

Konow, James, "Fair Shares : Accountability and Cognitive Dissonance in Allocation Decisions," *The American Economic Review*, 2000, *90* (4), 1072–1091.

Littler, J., *Against Meritocracy*, Abingdon, UK: Routledge, August 2017. This book has been made available under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 license.

Mollerstrom, Johanna, Bjørn Atle Reme, and Erik T. Sørensen, "Luck, choice and responsibility - An experimental study of fairness views," *Journal of Public Economics*, 2015, *131*, 33–40.

Mott, Frank L., "The Utility of the HOME-SF Scale for Child Development Research in a Large National Longitudinal Survey: The National Longitudinal Survey of Youth 1979 Cohort," *Parenting*, 2004, *4* (2-3), 259–270.

Pogliano, Andrea, "Born that way: Beliefs about genetics' importance and redistribution preferences," Tinbergen Institute Discussion Paper TI 2024-017/I, Amsterdam and Rotterdam 2024.

Preuss, Marcel, Germán Reyes, Jason Somerville, and Joy Wu, "Inequality of Opportunity and Income Redistribution," 2024.

Rawls, John, *A Theory of Justice: Original Edition*, Harvard University Press, 1971.

Rey-Biel, Pedro, Roman Sheremeta, and Neslihan Uler, "When Income Depends on Performance and Luck: The Effects of Culture and Information on Giving," in "Experimental Economics and Culture," Vol. 20, Emerald Publishing Ltd, 2018, pp. 167–203.

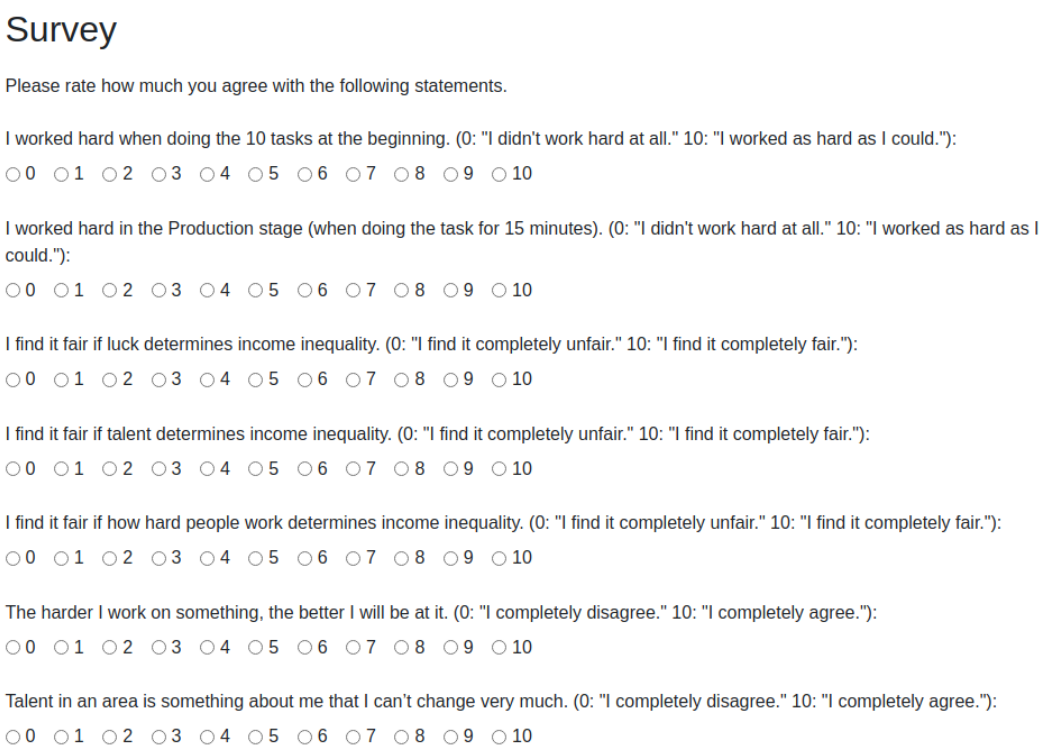
Rodriguez-Lara, Ismael and Luis Moreno-Garrido, "Self-interest and fairness: Self-serving choices of justice principles," *Experimental Economics*, 2012, *15* (1), 158–175.

Schildberg-Hörisch, Hannah, Marco A Schwarz, Chi Trieu, and Jana Willrodt, "Perceived Fairness and Consequences of Affirmative Action Policies," *The Economic Journal*, 08 2023, *133* (656), 3099–3135.

Trautmann, Stefan T., "Procedural fairness and equality of opportunity," *Journal of Economic Surveys*, 2022, pp. 1–18.

A Appendix

A.1 Survey



Survey

Please rate how much you agree with the following statements.

I worked hard when doing the 10 tasks at the beginning. (0: "I didn't work hard at all." 10: "I worked as hard as I could."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

I worked hard in the Production stage (when doing the task for 15 minutes). (0: "I didn't work hard at all." 10: "I worked as hard as I could."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

I find it fair if luck determines income inequality. (0: "I find it completely unfair." 10: "I find it completely fair."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

I find it fair if talent determines income inequality. (0: "I find it completely unfair." 10: "I find it completely fair."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

I find it fair if how hard people work determines income inequality. (0: "I find it completely unfair." 10: "I find it completely fair."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

The harder I work on something, the better I will be at it. (0: "I completely disagree." 10: "I completely agree."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Talent in an area is something about me that I can't change very much. (0: "I completely disagree." 10: "I completely agree."):

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Figure A.1: Screenshot of the survey at the end of the experiment

A.2 Participants

The participants are not a representative sample of the US population, so I present their demographic characteristics compared to the US population in Table A.1. The average age in the sample is 35 years, and the median is 33, which is lower than the median age in the US, 38.5. The sample has a larger share of females than the population (56 percent vs. 51 percent). There is a higher share of immigrants (80 percent born in the US vs. 85 percent in the population). There are more students among the participants than in the US population (31 percent vs. 10 percent among 18-year-olds or above). The share of employed people is similar to the population (62 vs. 61 percent), but the shares of unemployed and out of the labor force are different (16 vs. 3 percent and 20 vs. 36 percent, respectively).

	Mean	SD	N	US Mean
Age	34.93	(13.10)	585	38.50*
Female	0.56	(0.50)	589	0.51
Born in the US	0.80	(0.40)	582	0.85
Currently studying	0.31	(0.46)	484	0.10
Employment status				
Employed (part-time or full-time) or about to start employment	0.62	(0.49)	458	0.61
Unemployed	0.16	(0.37)	458	0.03
Not in labor force	0.20	(0.40)	458	0.36

Table A.1: Demographic characteristics of participants

Note: Demographic characteristics of the participants. Participants give these data and other background information to Prolific upon registration to the platform. The data presented here were available for the researcher to download. Some participants revoked their consent for the researcher to see the data, or the data expired by the time of the experiment, hence the varying number of observations across the rows. Source of the US population statistics: United States Census data, 2019 (https://data.census.gov/cedsci/). The labor market status data from the Census covers people aged 16 or above, while in the sample, the minimum age is 18. *Median age is presented in the US population instead of the mean age.

A.3 Attrition

	Did not come back	Did not come back
Low production potential	-0.0410	0.0188
	(0.0535)	(0.0627)
High production potential	-0.0488	-0.0492
	(0.0535)	(0.0603)
Difficulty experiment	-0.0468	-0.0375
	(0.0519)	(0.0590)
Low production potential	-0.0200	-0.0707
× Difficulty experiment	(0.0744)	(0.0862)
High production potential	0.0953	0.0773
× Difficulty experiment	(0.0736)	(0.0847)
Production	-0.00121	-0.000728
	(0.000756)	(0.000968)
Age		-0.000461
		(0.00165)
Female		0.00901
		(0.0380)
Currently studying		-0.00319
		(0.0445)
Employment status		
Full-Time		-0.140
		(0.122)
Not in paid work (e.g.		-0.201
homemaker', 'retired or disabled)		(0.130)
Other		-0.191
		(0.132)
Part-Time		-0.138
		(0.124)
Unemployed (and job seeking)		-0.147
		(0.126)
Constant	0.198***	0.328***
	(0.0373)	(0.127)
Observations	594	451

Standard errors in parentheses

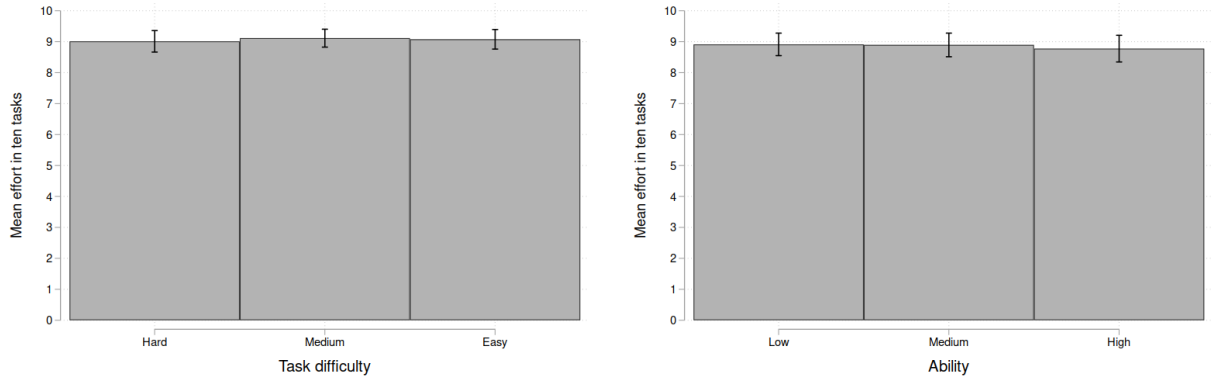
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Attrition from the Production to the Redistribution stage

Note: The table shows the effect of features of the experiment and own performance on attrition from the Production to the Redistribution part. The baseline category in the groups is the medium production potential group in the Ability experiment. Age and production are demeaned. The second column controls for demographic variables. The number of observations is lower in the second column because demographic data was not available for all participants.

A.4 Ability measure

In the Ability experiment, the measure of ability is how fast people do ten tasks when asked to do them as fast as possible. The idea is that if participants exert maximal effort in this part, remaining differences in their speed can only stem from their current ability in the task (as argued in Schildberg-Hörisch et al., 2023). Since I did not provide monetary incentives in this part, some participants might not have exerted maximal effort. This is only problematic if participants believe the measure reflects effort instead of ability when redistributing the income in the second stage. I aimed to minimize the differences in the effort people put into doing the ten tasks in both experiments by asking them to do them as fast as possible and emphasizing that their performance would provide important information about how fast these tasks can be done. This part also had a time limit, and participants knew they would be excluded from the experiment if they did not finish ten tasks within the time limit (see the exact instructions in Appendix Section A.11). At the end of the experiment, participants had to rate on a zero to ten scale how hard they worked on each part (ten tasks and production) of the first stage. Figure A.2 shows the average self-reported effort levels exerted in the ten tasks in the Difficulty and Ability experiments. The mean effort level in both experiments is around 9 out of 10, and it does not differ significantly by difficulty and ability level. Table A.3 validates the self-reported effort measure by showing that, controlling for the difficulty/ability level of the participant, self-reported effort in the production part still affects production. At the same time, self-reported effort in the ten tasks part does not affect production significantly (only on the ten percent level in the Ability experiment).



(a) Difficulty experiment

(b) Ability experiment

Figure A.2: Mean self-reported effort by difficulty/ability level

Note: The figures show the means of the self-reported effort levels of participants when doing ten asks as fast as they can, by difficulty/ability level. The spikes show 95 percent confidence intervals.

	Difficulty experiment	Ability experiment
	Production	Production
High difficulty/Low ability	-20.55***	-13.06***
	(3.369)	(2.028)
Low difficulty/High ability	27.39***	14.26***
	(3.283)	(2.016)
Worked hard on production	5.019***	4.285***
	(1.185)	(0.589)
Worked hard on 10 tasks	-1.056	-1.214*
	(1.161)	(0.551)
Constant	68.73***	69.80***
	(2.331)	(1.455)
Observations	257	243

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: Production by difficulty/ability level and self-reported effort

Note: The table shows the effect of own difficulty/ability level on production, including controls for self-reported effort in both parts. Effort level is measured on a 0 to 10 scale. Both effort variables are demeaned. The reference category in the levels in both columns is the medium difficulty/ability level. The constant therefore shows the production at the medium levels with average effort in both parts.

Table A.4 further explores the possible role of effort in the ability measure. The first column shows how the time taken for the ten tasks in the Ability experiment correlates with demographic characteristics. Subjects of the average age (35 years) who are full-time employed took about two minutes for the ten tasks – each additional year of age added about half a second to the time taken. Part-time workers and currently unemployed participants also took more time by 19-20 seconds. There was no difference in the time taken by gender, student status, or whether someone was born in the US.

Older people might have taken more time due to less practice in typing or worse vision, so even with high effort, they might be slower in this task than younger participants. It is not as straightforward why unemployed and part-time-employed participants took more time than others. They might be less skilled, but it is also possible they exerted less effort in the ten tasks since there were no monetary incentives. As Column 2 shows, unemployed participants reported significantly lower effort levels than others in the ten tasks part. Part-time workers also exerted less effort, although this difference is only significant at the ten percent level. In contrast, Column 3 shows that the self-reported effort level of these two groups in the production part does not differ from

that of other participants. Comparing their potential production – predicted production had they completed the same number of tasks per minute as in the ten tasks part – to their actual production, unemployed and part-time-employed participants also sped up significantly. Therefore, it is likely that those participants who needed the payment from the experiment more exerted less effort in the ten tasks stage and more effort in the production stage. This is a threat to identification if these participants regarded the ability measure as a measure of effort when making redistributive decisions or if other participants did so.

To address this threat to identification, I split the sample into participants with and without full-time employment and study whether they made different redistributive decisions in the Ability experiment. Participants without full-time employment might think the ability measure measures effort and ignore it. In contrast, participants with full-time employment might see it as a measure of ability and compensate low-ability participants. Table A.5 shows the results of this comparison.[20] There is no difference between the redistribution choices of employed and non-employed participants: neither group compensates for low ability. Unfortunately, employment status is missing for 25 percent of the sample. Therefore, I also compare the results to participants' decisions with missing employment data. These participants do not compensate for low ability, either.

Based on these results, some participants might have regarded the ability measure as another measure of effort but did not care about it and rewarded only production. However, suppose a large share of participants saw the ability measure as measuring effort. In that case, we should see redistribution towards the *higher-ability* participant due to spectators thinking she exerted high effort. Nevertheless, there is no such redistribution by spectators. Most participants were, therefore, likely to have thought about it as a measure of ability but did not feel that they should compensate for its differences.

[20]The full-time employed category includes those with full-time employment and those due to start a new job within the next month because they also have good employment prospects. The latter category only includes two percent of the sample. The non-employed category includes part-time employed and unemployed people, those not in paid work, and non-employed for other, non-specified reasons.

	Time 10 tasks	Effort in 10tasks	Effort in production	Prod. - pot.prod.
Age	0.454**	0.0213	0.0105	-0.0261
	(0.192)	(0.0139)	(0.0123)	(0.0865)
Female	0.737	0.141	0.0685	0.718
	(4.207)	(0.304)	(0.269)	(1.894)
Currently studying	3.564	0.671*	-0.153	-6.233***
	(5.147)	(0.372)	(0.329)	(2.317)
Born in the US	0.0478	-0.437	-0.715**	-3.373
	(4.609)	(0.333)	(0.295)	(2.075)
Employment status, ref.: Full-time employed				
Due to start a new job within the next month	-12.21	-1.279	-1.048	-6.898
	(13.98)	(1.010)	(0.894)	(6.295)
Not in paid work (e.g. homemaker', 'retired or disabled)	8.874	0.0794	0.0752	4.951*
	(6.434)	(0.465)	(0.411)	(2.897)
Other	5.150	-0.265	0.337	4.926
	(7.993)	(0.577)	(0.511)	(3.599)
Part-Time	19.03***	-0.728*	-0.129	4.927*
	(5.716)	(0.413)	(0.365)	(2.574)
Unemployed (and job seeking)	19.91***	-0.950**	-0.263	6.319**
	(5.760)	(0.416)	(0.368)	(2.593)
Constant	118.1***	9.179***	9.613***	-0.788
	(4.864)	(0.351)	(0.311)	(2.190)
Observations	186	186	186	186

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Exploring effort and ability

The table explores the role of effort in the ability measure in the Ability experiment. The first column shows the correlates of individual characteristics with the time taken for the ten tasks (in seconds). The second column shows the correlates of characteristics with the self-reported effort level in the ten tasks part. The third column looks at the self-reported effort level in the production part. The dependent variable in the fourth column is the difference between actual and potential production, where potential production is 15 times the individual production potential measure from the ten tasks part.

	Redistributed income share		
	(1)	(2)	(3)
	Employed	Non-employed	Missing data
Situation, ref. equal difficulty levels			
Lower ability	-0.00118	0.00818	0.0107
	(0.00701)	(0.00600)	(0.0112)
Constant	0.0138	-0.0145**	-0.00854
	(0.00971)	(0.00675)	(0.0270)
Observations	430	540	220
Participant fixed effect	no	no	no
Demographic controls	no	no	no
Session fixed effect	yes	yes	yes

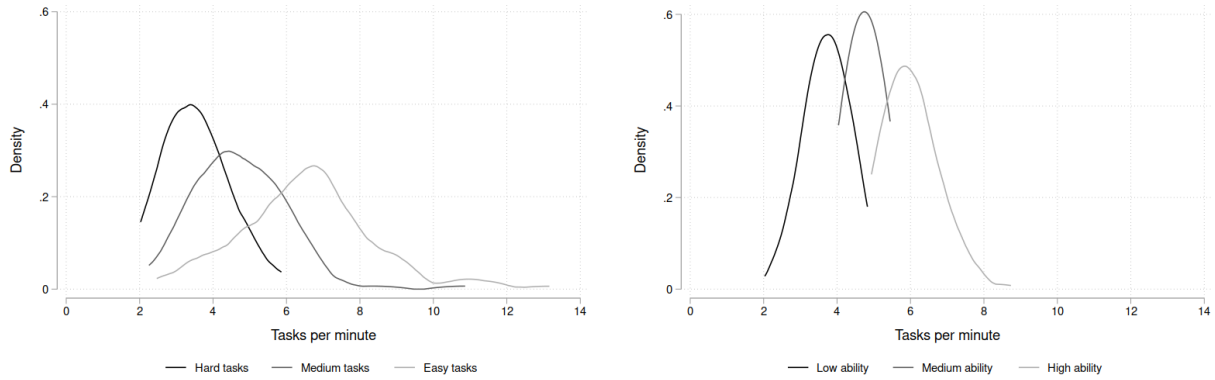
Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Spectator decision correlations with employment status

Note: The outcome variable is the excess income share given to a randomly chosen participant in a pair on top of her production share by spectators in the Ability experiment. Column 1 shows the redistribution among those participants who have full-time employment or are about to start working. Column 2 includes participants who are part-time employed, unemployed, not in paid work, or other non-employed. Column 3 includes those whose employment data are missing. Standard errors are not clustered due to the low number of subjects in each column.

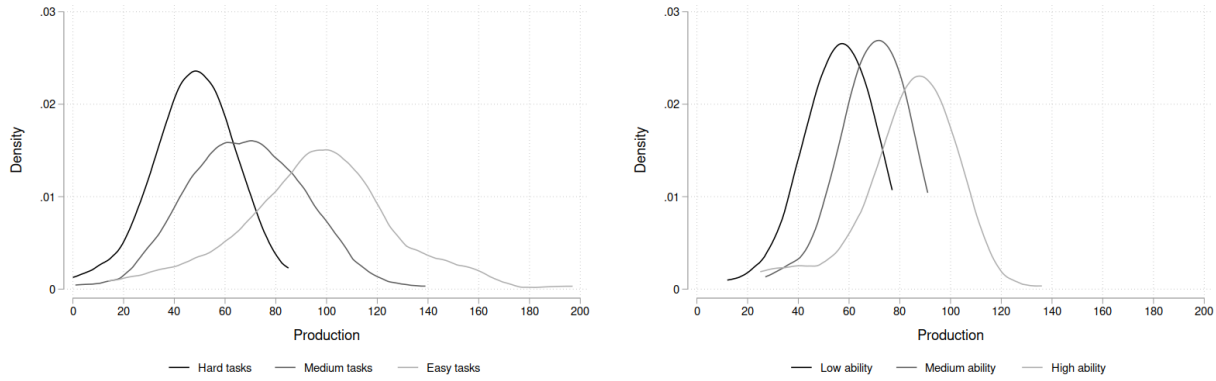
A.5 Distribution of production potential and production by difficulty/ability level



(a) Difficulty experiment (b) Ability experiment

Figure A.3: Distribution of production potential by difficulty/ability level

Note: The figures show the distribution of the individual production potential (tasks/minute) by difficulty and ability levels. There is an overlap in the distributions in the Ability experiment as well because the thresholds between the groups differ across the three sessions.

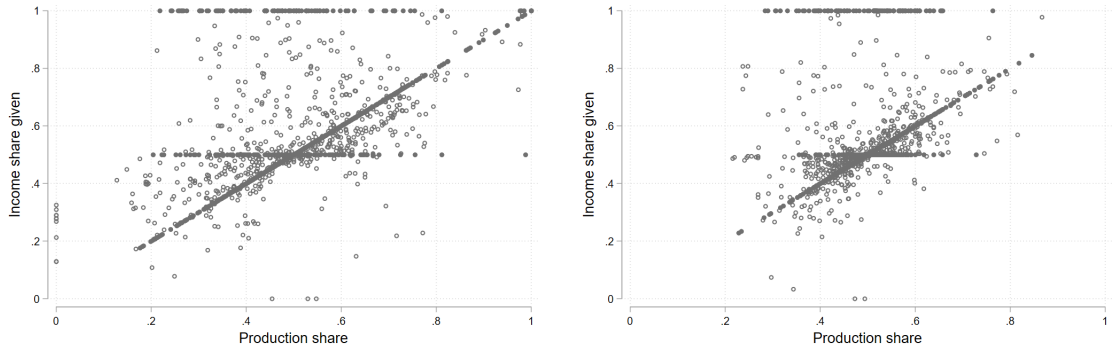


(a) Difficulty experiment (b) Ability experiment

Figure A.4: Distribution of production by difficulty/ability level

Note: The figures show the distribution of the individual production (tasks done within 15 minutes) by difficulty and ability levels.

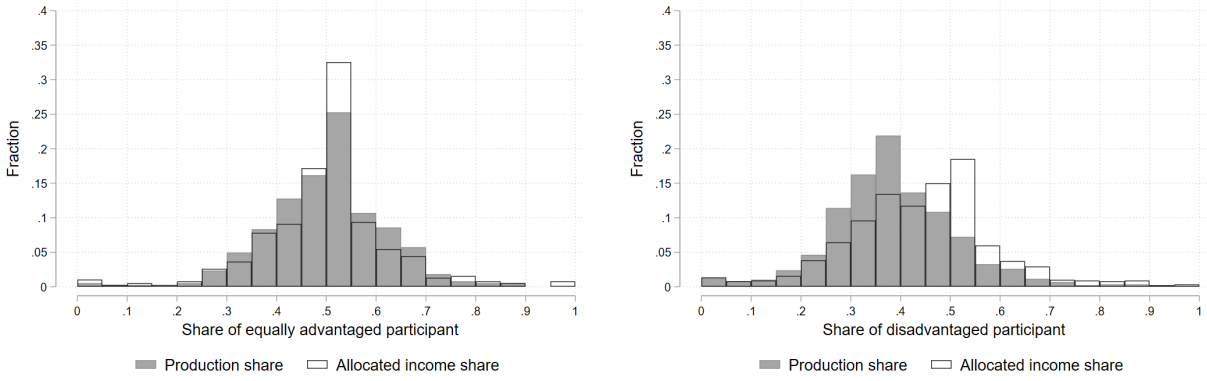
A.6 Redistribution



(a) Stakeholders, Difficulty experiment (b) Stakeholders, Ability experiment

Figure A.5: Share allocated to self in the pair

Note: The figures show the share of tokens allocated to a self in the pair by stakeholders, plotted against the production share of the stakeholder. One point indicates one decision.

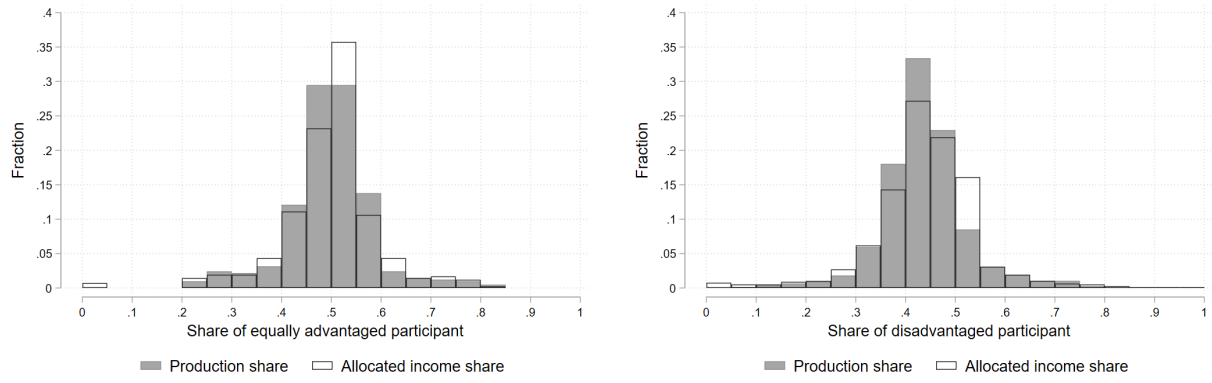


(a) Equal difficulty levels within pair
0.50 vs 0.50 (p=0.715)

(b) Harder tasks within pair
0.39 vs 0.44 (p=0.000)

Figure A.6: Distribution of shares among spectators in the Difficulty experiment

Note: The figures show the distribution of the production share within the pair and the income share allocated by spectators in the Difficulty experiment. The first panel shows the distribution of shares of a randomly chosen participant in the pair with equal difficulty levels. The second panel shows the distribution of the shares of the disadvantaged participant within the pair in case of unequal difficulties. The numbers show the mean production share vs. the mean allocated income share.

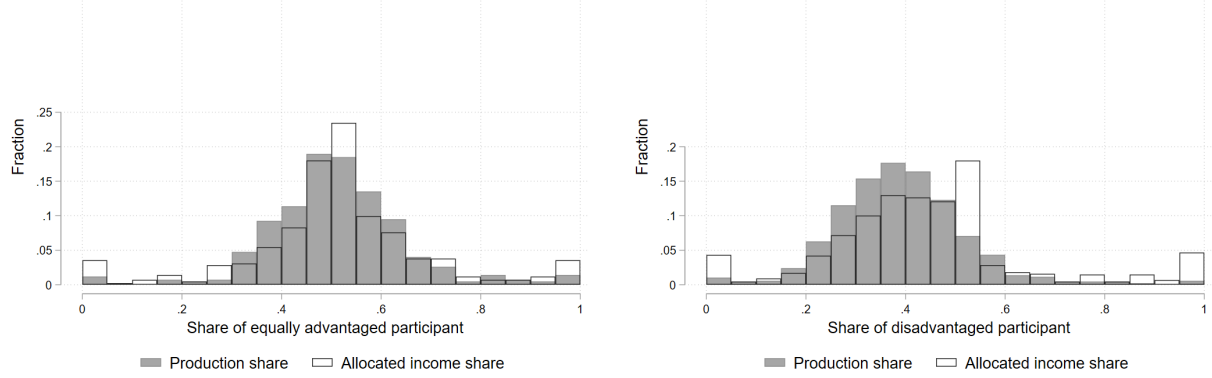


(a) Equal ability levels within pair
0.49 vs 0.50 (p=0.393)

(b) Lower ability level within pair
0.43 vs 0.44 (p=0.504)

Figure A.7: Distribution of shares among spectators in the Ability experiment

Note: The figures show the distribution of the production share within the pair and the income share allocated by spectators in the Ability experiment. The first panel shows the distribution of shares of a randomly chosen participant in the pair with equal difficulty levels. The second panel shows the distribution of the shares of the disadvantaged participant within the pair in case of unequal difficulties. The numbers show the mean production share vs. the mean allocated income share.

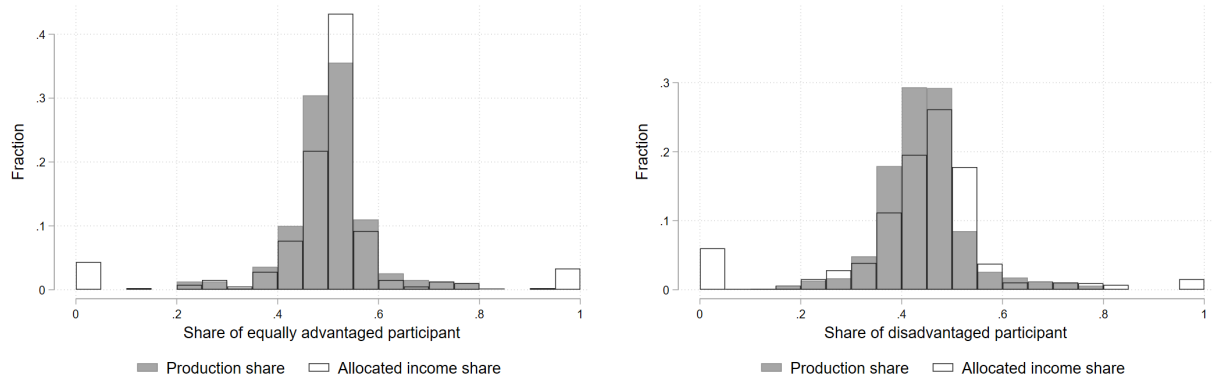


(a) Equal difficulty levels within pair
0.50 vs 0.50 (p=0.319)

(b) Harder tasks within pair
0.39 vs 0.44 (p=0.001)

Figure A.8: Distribution of shares among stakeholders in the Difficulty experiment

Note: The figures show the distribution of the production share within the pair and the income share allocated by stakeholders in the Difficulty experiment. The first panel shows the distribution of shares of a randomly chosen participant in the pair with equal difficulty levels. The second panel shows the distribution of the shares of the disadvantaged participant within the pair in case of unequal difficulties. The numbers show the mean production share vs. the mean allocated income share.



(a) Equal ability levels within pair

0.50 vs 0.51 (p=0.289)

(b) Lower ability level within pair

0.44 vs 0.44 (p=0.759)

Figure A.9: Distribution of shares among stakeholders in the Ability experiment

Note: The figures show the distribution of the production share within the pair and the income share allocated by stakeholders in the Ability experiment. The first panel shows the distribution of shares of a randomly chosen participant in the pair with equal difficulty levels. The second panel shows the distribution of the shares of the disadvantaged participant within the pair in case of unequal difficulties. The numbers show the mean production share vs. the mean allocated income share.

A.7 Robustness checks of reduced-form results

A.7.1 Sample with all demographic controls

	Difficulty experiment				Ability experiment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Situation, ref. Equal difficulty/ability level								
Higher difficulty / Lower ability level	0.0554***	0.0397***	0.0408***	0.0450***	0.00445	-0.00312	-0.00184	-0.000696
	(0.00823)	(0.00725)	(0.00723)	(0.00833)	(0.00604)	(0.00509)	(0.00495)	(0.00584)
Constant	-0.00550	0.0658***	0.0799***	0.0685***	-0.00489	0.0528*	0.00643	0.0494*
	(0.00979)	(0.0190)	(0.0233)	(0.0142)	(0.0104)	(0.0293)	(0.0353)	(0.0251)
Observations	980	980	980	980	960	960	960	960
Share in total production	no	yes	yes	yes	no	yes	yes	yes
Participant fixed effect	no	no	no	yes	no	no	no	yes
Demographic controls	no	no	yes	no	no	no	yes	no
Session fixed effect	yes	yes	yes	no	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Spectator decisions

Note: The outcome variable is the excess income share given to one participant in a pair on top of her production share by spectators in the Difficulty experiment (1-4) and the Ability experiment (5-8). Columns 1 and 5 only control for whether the participant had more difficult tasks or lower ability than the other participant. Columns 2-4 and 6-8 also control for the production share of Participant 1. Columns 3 and 7 include demographic controls: age, gender, whether the spectator was born in the US, whether she has US nationality, labor market status and student status. Columns 4 and 8 add participant fixed effects.

	Difficulty experiment				Ability experiment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Situation, ref. Equal difficulty/ability level								
Higher difficulty / Lower ability level	0.0632***	0.0351*	0.0350*	0.0283	-0.00322	-0.0247	-0.0254	-0.0303
	(0.0168)	(0.0187)	(0.0188)	(0.0179)	(0.0184)	(0.0250)	(0.0248)	(0.0300)
Constant	0.0126	0.139***	0.222***	0.111***	-0.000751	0.164	0.0965	0.183
	(0.0262)	(0.0484)	(0.0649)	(0.0303)	(0.0177)	(0.106)	(0.109)	(0.126)
Observations	960	960	960	960	880	880	880	880
Share in total production	no	yes	yes	yes	no	yes	yes	yes
Participant fixed effect	no	no	no	yes	no	no	no	yes
Demographic controls	no	no	yes	no	no	no	yes	no
Session fixed effect	yes	yes	yes	no	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Stakeholder decisions

Note: The outcome variable is the excess income share given to one participant in a pair on top of her production share by stakeholders in the Difficulty experiment (1-4) and the Ability experiment (5-8). Columns 1 and 5 only control for whether the participant had more difficult tasks or lower ability than the other participant. Columns 2-4 and 6-8 also control for the production share of Participant 1. Columns 3 and 7 include demographic controls: age, gender, whether the spectator was born in the US, whether she has US nationality, labor market status and student status. Columns 4 and 8 add participant fixed effects.

A.7.2 Share of original income difference redistributed

	Difficulty experiment				Ability experiment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Situation, ref. Equal difficulty/ability level								
Higher difficulty / Lower ability level	0.417***	0.482***	0.485***	0.502***	0.0349	0.0148	0.0201	0.0312
	(0.0923)	(0.123)	(0.122)	(0.127)	(0.0822)	(0.0928)	(0.0915)	(0.101)
Constant	-0.0222	-0.312	-0.150	-0.242	-0.0895	0.0724	-0.221	0.0527
	(0.0942)	(0.228)	(0.148)	(0.186)	(0.108)	(0.208)	(0.219)	(0.195)
Observations	1200	1200	1200	1200	1149	1149	1149	1149
Share in total production	no	yes	yes	yes	no	yes	yes	yes
Participant fixed effect	no	no	no	yes	no	no	no	yes
Demographic controls	no	no	yes	no	no	no	yes	no
Session fixed effect	yes	yes	yes	no	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Spectator decisions

Note: The outcome variable is the share of the original income difference redistributed to one participant in a pair by spectators in the Difficulty experiment (1-4) and the Ability experiment (5-8). Columns 1 and 5 only control for whether the participant had more difficult tasks or lower ability than the other participant. Columns 2-4 and 6-8 also control for the production share of Participant 1. Columns 3 and 7 include demographic controls: age, gender, whether the spectator was born in the US, and whether she has US nationality. Columns 4 and 8 add participant fixed effects.

	Difficulty experiment				Ability experiment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Situation, ref. Equal difficulty/ability level								
Higher difficulty / Lower ability level	0.247	0.222	0.219	0.223	-0.213	-0.350	-0.364	-0.303
	(0.289)	(0.370)	(0.366)	(0.300)	(0.261)	(0.299)	(0.304)	(0.316)
Constant	0.145	0.254	1.220	-0.0909	0.145	1.270*	0.670	1.504
	(0.239)	(0.632)	(0.763)	(0.447)	(0.272)	(0.681)	(0.667)	(1.005)
Observations	1242	1242	1242	1242	1189	1189	1189	1189
Share in total production	no	yes	yes	yes	no	yes	yes	yes
Participant fixed effect	no	no	no	yes	no	no	no	yes
Demographic controls	no	no	yes	no	no	no	yes	no
Session fixed effect	yes	yes	yes	no	yes	yes	yes	no

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Stakeholder decisions

Note: The outcome variable is the share of the original income difference redistributed to one participant in a pair by stakeholders in the Difficulty experiment (1-4) and the Ability experiment (5-8). Columns 1 and 5 only control for whether the participant had more difficult tasks or lower ability than the other participant. Columns 2-4 and 6-8 also control for the production share of Participant 1. Columns 3 and 7 include demographic controls: age, gender, whether the spectator was born in the US, and whether she has US nationality. Columns 4 and 8 add participant fixed effects.

A.8 Survey answers

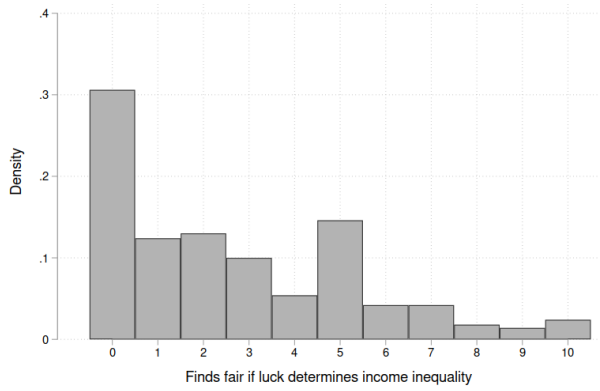


Figure A.10: Distribution of answers to fairness question – luck

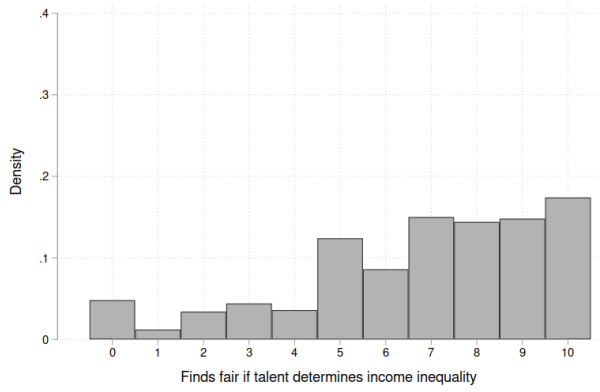


Figure A.11: Distribution of answers to fairness question – talent

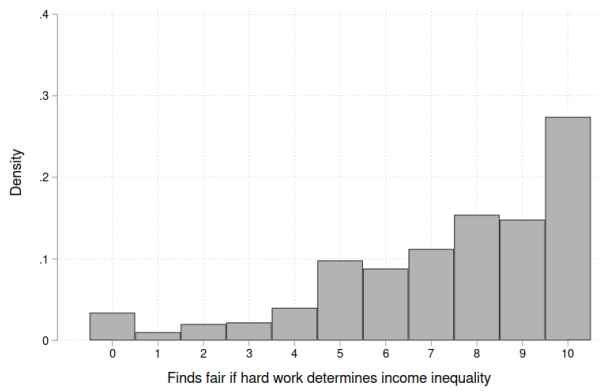


Figure A.12: Distribution of answers to fairness question – hard work

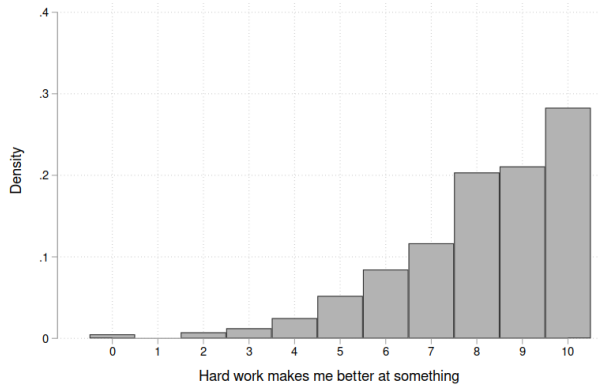


Figure A.13: Distribution of answers to growth mindset question – hard work makes me better

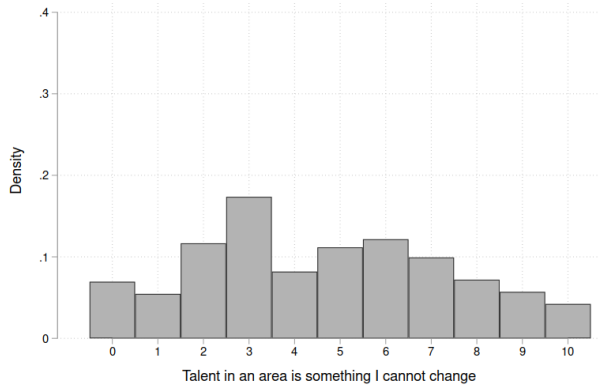


Figure A.14: Distribution of answers to growth mindset question – talent cannot change

A.8.1 Survey answers by ability level

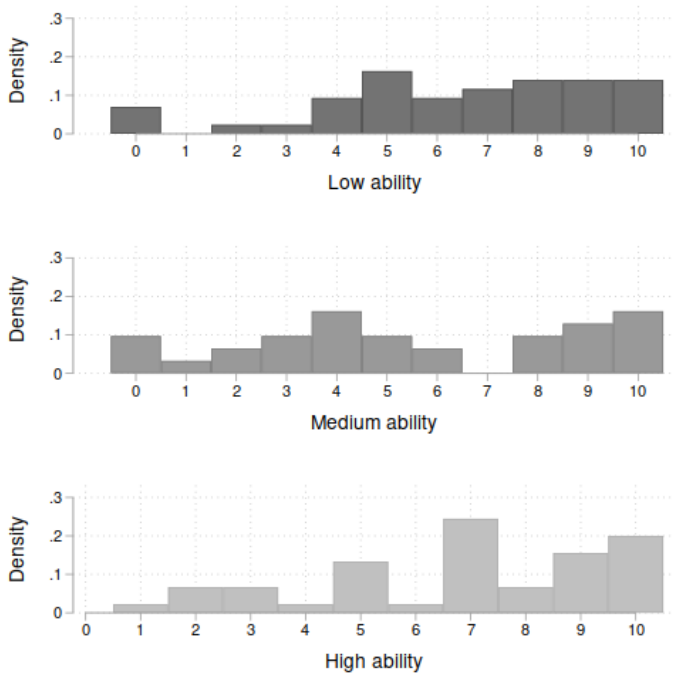


Figure A.15: Finds it fair if talent determines inequality (Ability experiment, spectators)

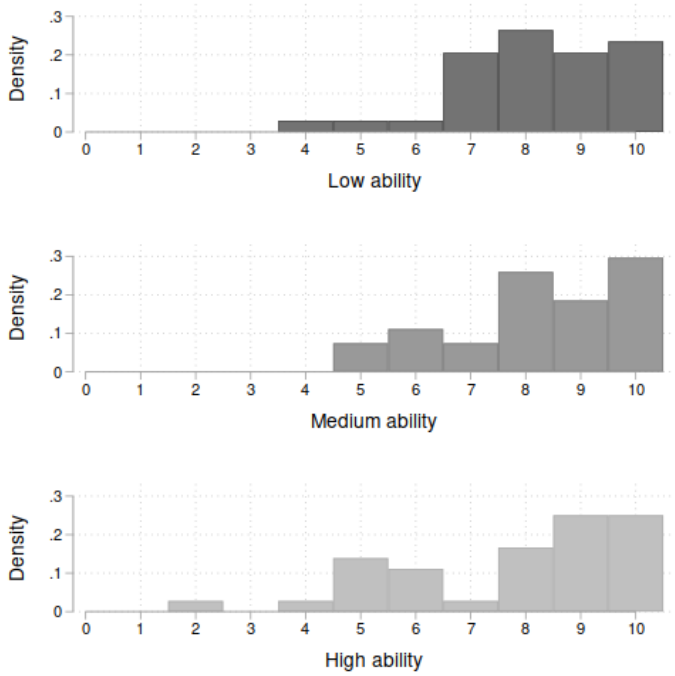


Figure A.16: Hard work makes one better at something (Ability experiment, spectators)

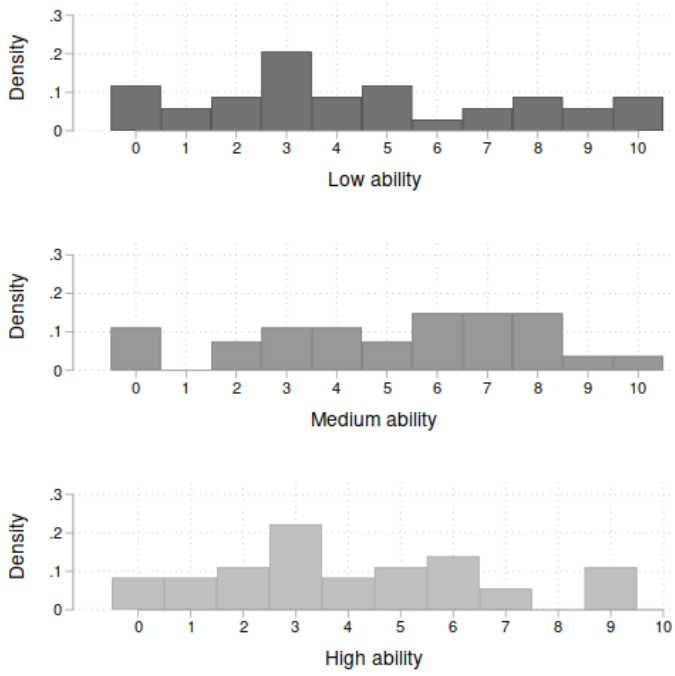


Figure A.17: Talent is something one cannot change (Ability experiment, spectators)

A.9 Heterogeneity in stakeholder decisions

	Redistributed income share to self		
	(1)	(2)	(3)
Own difficulty level, ref. medium-difficulty tasks			
Long tasks	0.0786**	0.0448	0.0420
	(0.0353)	(0.0378)	(0.0357)
Short tasks	-0.00292	0.0222	0.0178
	(0.0297)	(0.0268)	(0.0265)
Constant	0.0277	0.205***	0.240***
	(0.0256)	(0.0471)	(0.0732)
Observations	1250	1250	1250
Share in total production	no	yes	yes
Participant fixed effect	no	no	no
Demographic controls	no	no	yes
Session fixed effect	yes	yes	yes

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Stakeholders' decisions by own difficulty

Note: The outcome variable is the excess income share stakeholders in the Difficulty experiment give to themselves on top of their production share. Column 1 controls only for the whether the stakeholder had hard, medium, or easy tasks. Columns 2-3 also control for the production share of the stakeholder. Column 3 includes demographic controls: age, gender, whether the stakeholder was born in the US, and whether she has US nationality.

	Redistributed income share to self		
	(1)	(2)	(3)
Own difficulty level, ref. medium ability			
Low ability	0.0408*	0.0145	0.0314
	(0.0231)	(0.0234)	(0.0239)
High ability	0.0734**	0.0875***	0.0923***
	(0.0314)	(0.0323)	(0.0310)
Constant	0.0124	0.248***	0.357***
	(0.0220)	(0.0782)	(0.0815)
Observations	1210	1210	1210
Share in total production	no	yes	yes
Participant fixed effect	no	no	no
Demographic controls	no	no	yes
Session fixed effect	yes	yes	yes

Standard errors are clustered on the participant level. Clustered standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: Stakeholders' redistribution by own ability

Note: he outcome variable is the excess income share stakeholders in the Ability experiment give to themselves on top of their production share. Column 1 controls only for the whether the stakeholder had hard, medium, or easy tasks. Columns 2-3 also control for the production share of the stakeholder. Column 3 includes demographic controls: age, gender, whether the stakeholder was born in the US, and whether she has US nationality.

A.10 Structural estimation

A.10.1 Utility maximization of spectators and stakeholders

We assume that both spectators and stakeholders hold one of the fairness views described in Section 5, but their optimization problems in redistributive decisions differ. The utility maximization problem of a spectator is the following (from Cappelen et al., 2010, 2020b):

$$U_i(\mathbf{x},\, \theta, k_i) = -\frac{(t_{1,i} - t_1^{k_i})^2}{2Y} \tag{6}$$

Here, $t_{1,i}$ is the decision variable of spectator i – the number of tokens allocated to Participant 1 in the pair, while $t_1^{k_i}$ is what she finds fair to allocate given her fairness view k. Since spectators do not have any monetary gain from the decision, they simply choose the allocation that aligns with their fairness views. The optimal spectator decision is, therefore:

$$t_{1,i}^* = t_1^{k_i} \tag{7}$$

Stakeholders, on the other hand, consider both the fairness of the allocation and their individual monetary gain. Their problem can be written as follows:

$$U_j(\mathbf{x},\, \theta, k_j, \beta_j) = t_{own} - \beta_j \frac{(t_{own,j} - t_{own}^{k_j})^2}{2Y}, \tag{8}$$

Here, $t_{own,j}$ is the number of tokens stakeholder j gives herself in the decision, while $t_{own}^{k_j}$ is the tokens she finds fair to give according to her fairness view k. β_j is her weight on fairness relative to her final income. Therefore, the optimal stakeholder decision is:

$$t_{own,j}^* = t_{own}^{k_j} + \frac{Y}{\beta_j} \tag{9}$$

Stakeholders give themselves the number of tokens they find fair to give and some extra tokens depending on how selfish they are.

A.10.2 Maxmimum likelihood estimation

Based on the model described above, I use a maximum likelihood estimation to estimate the shares of fairness views among spectators and stakeholders separately (following Mollerstrom et al. 2015 and Andre 2024). The fact that every participant made multiple decisions in different situations makes classifying them into separate fairness preference types possible.

I assume that spectators redistribute income to a randomly chosen participant within the pair based on what they find fair with a normally distributed response error:

$$t_{1,i,p} = t_{1,p}^{k_i} + \epsilon_{i,p}, \tag{10}$$

where $\epsilon_{i,p} \sim i.i.d.\, N(0, \sigma^2)$. One participant made ten decisions for ten different random pairs. The total likelihood of a participant given that she is of type k is therefore

$$L_i(k) = \prod_p \phi\left(t_{1,i,p} - t^k_{1,p}\,,\, \sigma\right), \tag{11}$$

where $\phi()$ is the probability density function of the normal distribution, and p denotes a decision made for pair p. The total likelihood of a participant is the weighted sum of the conditional likelihoods given a fairness view weighted by the share of that fairness view among the participants:

$$L_i = \lambda_{PM} \cdot L_i(PM) + \lambda_E \cdot L_i(E) + \lambda_{CM} \cdot L_i(CM) \tag{12}$$

λ_{PM} denotes the share of performance meritocrats, λ_E the egalitarians, and λ_{CM} the compensating meritocrats. Using the independence of the error terms across participants, the total log-likelihood the estimation maximizes is

$$LogL = \sum_i log(L_i) \tag{13}$$

The parameters to estimate are the standard deviation of the response error, σ, and the shares of the types, λ_{PM}, λ_E and λ_{CM}.

I assume a similar choice structure for stakeholders as for spectators. Stakeholders redistribute income to themselves according to what they find fair, plus an extra amount depending on how selfish they are. There is also a normally distributed response error in their decision:

$$t_{own,j,p} = t^{kj}_{own,p} + \frac{1}{\beta} Y_p + \epsilon_{j,p}, \tag{14}$$

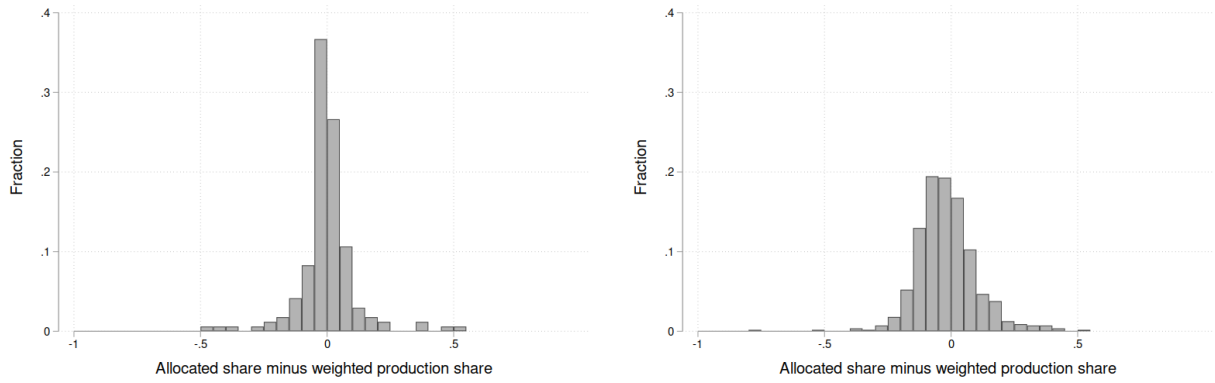
where $\epsilon_{j,p} \sim N(0, \sigma^2)$, Y_p is the total income of pair p, and β is the average weight stakeholders put on fairness. The total likelihood of a stakeholder given that she is of type k is, analogously to spectators,

$$L_j(k) = \prod_p \phi\left(t_{own,j,p} - t^k_{own,p} - \frac{1}{\beta} Y_p\,,\, \sigma\right). \tag{15}$$

The total likelihood of a participant and the total log-likelihood in the sample to maximize is the same as the spectator equations 12 and 13.

A.10.3 Testing other functional forms for compensators

Looking at the actual decisions in the Difficulty experiment that were not purely performance-meritocratic or egalitarian, in unequal difficulty situations, the average redistribution is 1.8 percentage points lower than what compensation based on the weighted production share would predict.[21] Figure A.18 shows the distribution of the differences between actual compensation and the weighted-performance-based one in these decisions. Though the difference between actual and predicted compensation is low, assuming this functional form can lead to underestimating the share of compensating meritocrats.



(a) Equal difficulty levels

(b) Harder tasks

Figure A.18: Allocated income share compared to weighted production share

Note: The histograms show the distributions of income allocations compared to those based on the weighted production shares (by the average production potential) of the participant. The sample includes only spectators in the Difficulty experiment and only those decisions that are not purely performance-meritocratic or egalitarian. Panel (a) shows allocation decisions made for a random participant with equal difficulty levels, while panel (b) shows allocation decisions made for the participant with a higher difficulty level. A Kolmogorov-Smirnov test rejects that the two distributions are equal.

Tables A.12 and A.13 show the results of the structural estimation of shares assuming different functional forms for compensators. Each column tests a different power α in the compensating function

$$t_1^{CM}(\mathbf{x},\, \theta) = \frac{x_1 \cdot \theta_1^{\alpha}}{x_1 \cdot \theta_1^{\alpha} + x_2 \cdot \theta_2^{\alpha}} Y.$$

The first column corresponds to the power used in the main structural estimation, $\alpha = -1$. The subsequent columns test $\alpha = -2/3, -1/2$ and $-1/3$. The estimated shares of performance meritocrats and difficulty-compensating meritocrats in the Difficulty experiment are sensitive to the functional form definition. However, if we look at the not strictly egalitarian and performance-

[21] This is a rigorous comparison as it classifies every performance-meritocratic or egalitarian decision made with some error as neither of these types.

meritocratic decisions, the magnitude of compensation is somewhere between those defined by the first two columns. Therefore, the share of difficulty-compensating meritocrats is between 29 and 36 percent, while the share of performance meritocrats is between 55 and 48 percent. The estimated shares in the Ability experiment are robust to the definition of the compensating function.

	$\alpha = -1$		$\alpha = -2/3$		$\alpha = -1/2$		$\alpha = -1/3$	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
σ	112.813	2.284	115.112	2.347	118.041	2.411	121.424	2.472
λ_{PM}	0.553	0.047	0.478	0.053	0.385	0.063	0.140	0.090
λ_{E}	0.152	0.034	0.165	0.035	0.176	0.036	0.191	0.037
λ_{DCM}	0.294	0.046	0.357	0.053	0.438	0.065	0.669	0.097

Table A.12: Difficulty experiment, spectators – different functional forms

Note: The table shows the results of a test for different functional forms for compensators in the Difficulty experiment. The functional form for compensators is $(x_1 \cdot \theta_1^\alpha)/(x_1 \cdot \theta_1^\alpha + x_2 \cdot \theta_2^\alpha)$, where x_j is the production and θ_j is the average production potential of participant j. Each column tests a different power α.

	$\alpha = -1$		$\alpha = -2/3$		$\alpha = -1/2$		$\alpha = -1/3$	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
σ	82.905	1.716	82.906	1.712	82.906	1.711	82.907	1.711
λ_{PM}	0.871	0.032	0.871	0.032	0.871	0.032	0.871	0.032
λ_{E}	0.129	0.032	0.129	0.032	0.129	0.032	0.129	0.032
λ_{ACM}	0.000	0.004	0.000	0.001	0.000	0.001	0.000	0.001

Table A.13: Ability experiment, spectators – different functional forms

Note: The table shows the results of a test for different functional forms for compensators in the Ability experiment. The functional form for compensators is $(x_1 \cdot \theta_1^\alpha)/(x_1 \cdot \theta_1^\alpha + x_2 \cdot \theta_2^\alpha)$, where x_j is the production and θ_j is the average production potential of participant j. Each column tests a different power α.

A.10.4 Structural results for not fully selfish participants

	Estimate	Std. error
σ	137.168	2.878
β	45.041	6.662
λ_{PM}	0.593	0.055
λ_E	0.229	0.045
λ_{DCM}	0.177	0.050

(a) Difficulty experiment

	Estimate	Std. error
σ	102.713	2.127
β	57.243	6.233
λ_{PM}	0.809	0.041
λ_E	0.171	0.041
λ_{ACM}	0.019	0.018

(b) Ability experiment

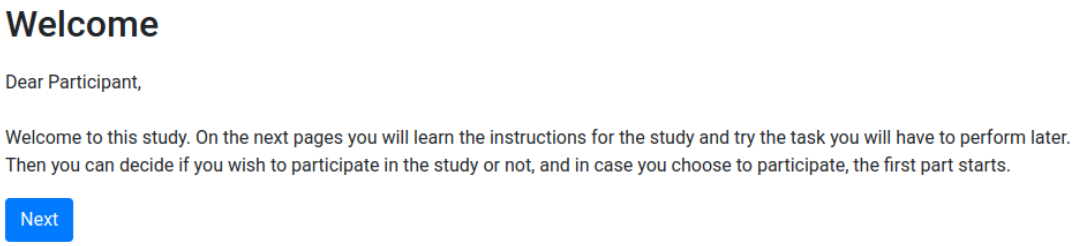
Table A.14: Structural estimation results – Stakeholders excluding fully selfish individuals

Note: Stakeholder results excluding participants who took all the money in their decisions. σ is the standard deviation of the response error, β is the average weight stakeholders put on fairness, λ_M is the estimated share of meritocrats, λ_E is the share of egalitarians, and λ_C is the share of compensating meritocrats (compensating for external difficulties in the Difficulty experiment and compensating for low ability in the Ability experiment). Parameters estimated in R using the `stats4` package. Standard errors are calculated with the delta method.

A.11 Experimental instructions

A.11.1 Production

The example task in the screenshots is a 4-letter task in the Difficulty experiment, but the Production stage instructions were the same for all participants across treatments.

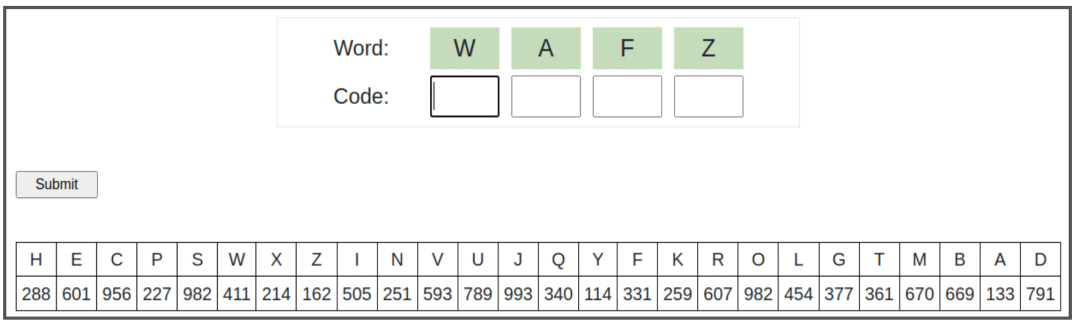


Instructions

This study will take place entirely online, in two parts. If you decide to take part in the study, the first part will start immediately after you give consent to participate. You can complete the second part anytime tomorrow between 6 AM and midnight (Pacific Time). You will receive a **completion bonus of £2.50** for completing the first part, **plus £1.00** if you complete the second part as well. You can earn **additional income** that will depend on your performance in the first part and your decisions or on other participants' decisions in the second part. The average additional income will be **around £3.00.** The first part will take around 25-30 minutes, and the second part around 5-10 minutes to complete.

First part - TODAY

In the first part you will have to do a simple task. In this task you have to encrypt letter combinations (= "words") into numbers, based on the key you see at the bottom of the page. For reference, see **a screenshot of the task** below.



H	E	C	P	S	W	X	Z	I	N	V	U	J	Q	Y	F	K	R	O	L	G	T	M	B	A	D
288	601	956	227	982	411	214	162	505	251	593	789	993	340	114	331	259	607	982	454	377	361	670	669	133	791

In this example the letter "W" corresponds to the number 411, "A" to 133, "F" to 331, and "Z" to 162, so you have to type 411, 133, 331, and 162 into the boxes under the letters, respectively. You can submit your response by clicking the Submit button below the task. If any of the entries are incorrect, the computer will tell you the number of entries that are wrong but not which ones are wrong. Then you will be able to revise your solution. Once the correct numbers are entered, after clicking the Submit button, you will see a new word to encrypt. After each correctly solved task the computer generates a new encryption key, too. The order of the letters in the key is also shuffled between the tasks.

When a new word is generated, **the first box will become active,** so you can start entering the first number. The fastest way to navigate from one box to the next is to use the **tabulator key (Tab)** on your keyboard. From the last box you can navigate to the Submit button directly with the Tab key and hit Enter. The tab key looks like this on your keyboard on PC (1) and on Mac (2 or 3):

(1) (2) (3)

On the next page you can try the task you will do in the experiment. You will be automatically navigated to the following page after completing the task or after 2 minutes.

Next

Try task

Time left to complete this page: **1:50**

Word:	G	A	R	B
Code:				

Submit

D	F	U	P	C	X	O	L	S	Y	M	B	A	H	G	W	R	V	J	I	N	E	Q	Z	K	T
124	227	474	628	521	484	448	672	767	853	598	831	440	631	993	695	562	922	733	303	879	526	735	788	794	220

Reminder: The fastest way to navigate from one box to the next is to use the **tabulator key (Tab)** on your keyboard. From the last box you can navigate to the Submit button directly with the Tab key and hit Enter. The tab key looks like this on your keyboard on PC (1) and on Mac (2 or 3):

(1) (2) (3)

Instructions - continued

First part - TODAY

The first part will take around 25-30 minutes to complete and will have **two main stages**. In both stages you will have to perform the same task you just tried. You will receive £2.50 for completing this part.

1. 10 tasks

First, I will ask you to do **10 tasks, as fast as you can,** to measure how quickly you can perform the tasks. Although you do not receive payment for this part in particular, it will serve as **information about how fast these tasks can be done.** It may also affect your additional income (on top of the completion bonuses) determined in the second part. You will have 5 minutes for this part. You will only be able to continue with the study if you **complete the 10 tasks within the 5 minutes.** Before doing the 10 tasks you will be able to practice for 2 minutes.

2. Production

In this part, you will have to **do the task for 15 minutes.** The **number of tasks you correctly completed** within these 15 minutes will be **your production.**

Earnings from the first part

You will earn **10 tokens for each completed task** in the Production stage.

Note that this is **not how much you will eventually earn in addition to the completion bonuses**. Your final additional income may change according to your decisions or other participants' **decisions in the second part**.

Second part - TOMORROW

The second part will take place tomorrow, and you can complete it anytime between 6 AM and midnight (Pacific Time). It will take around 5-10 minutes to complete. You will receive a completion bonus of £1.00 for completing this part. You will receive an **invitation tomorrow via the Prolific emailing system with the link to the second part**. If you complete the first part, when entering the second part the computer will automatically recognize you by your Prolific ID and let you start.

At the beginning of this part, you will be **randomly paired with another participant** who also completed the first part. In this part you will have to make **decisions about the distribution of earnings from the first part** within pairs. Your final additional income, that is on top of the fixed completion bonuses, will therefore depend on your decisions or on other participants' decisions made in the second part. The **completion bonuses** for the first and the second parts **will not be affected in any way by either your or other participants' decisions. The decisions will affect only the additional income you will earn.** The final additional earnings will be converted to British Pounds at the end of the study with **250 tokens = £1.00**. You will receive more detailed instructions about the second part at the beginning of the second part.

Payment for the study

You will receive **all payments within 3 days after the second part** via the Prolific payment system. **If you only complete the first part, you will receive £2.50 within 3 days after the corresponding second part.** Due to the interactive nature of the study we will have to wait for all participants' answers to draw the final payments. Because of this, **you will only learn how much you earned in total when you receive your payments for the study.**

On the next page you can test your comprehension of the study and you can decide if you wish to participate. You will see the instructions again under the consent form.

Further Information

This study is conducted by Luca Flora Drucker and financed by Central European University. If you accept to participate in the study, you may still change your mind and quit at any time. However, please note that you only receive the completion bonus for the first part and the completion bonus for the second part and the additional income if you complete the first and the second parts, respectively.

Participation in this study is not associated with any foreseeable risk or benefit. Your answers will be collected confidentially and anonymously (the researcher will not be able to link decisions and participants' identity beyond the Prolific ID provided). At the data analysis stage your Prolific ID will be changed to a random identifying number, and the Prolific IDs will be deleted. In case the results of the study are published, there can be no references to your identity. Data anonymity is guaranteed.

This study received a research ethics approval from the Ethical Research Committee of Central European University.

If you have any questions or concerns regarding this study, please contact me at lucafloradrucker.research@gmail.com.

Next

Comprehension check and consent form

Please answer the following questions before you decide if you wish to participate, to make sure you understand the details of the study. You can see the instructions again at the bottom of this page for reference. If you give an incorrect answer to any of the questions and click the Next button, the computer will notify you that the answer is wrong.

When is the second part of the study?

- ○ Today.
- ○ Tomorrow.
- ○ In two days.

If you only complete the first part, what will be your payment?

- ○ The completion bonus for the first part plus the payment after the number of tasks you do.
- ○ Only the completion bonus for the first part.
- ○ Nothing.

How will you earn additional income on top of the completion bonuses?

- ○ The additional income will be based on the tasks you perform in the production stage, but will be determined by decisions in the second part.
- ○ The number of tasks you perform in the production stage will be your additional income.
- ○ The additional income will entirely be determined by luck.

Consent Form

Do you wish to participate in the study?

- ○ Yes
- ○ No

Next

Practice

Welcome to the study! Thank you for taking part.
Now you can **practice the task for 2 minutes.**

Next

10 tasks

Now please do **10 tasks as fast as you can,** to measure how fast you can solve the tasks. You will have 5 minutes for this part. You have to finish the 10 tasks within the 5 minutes **to be able to continue the study.**

Next

Beginning of production

Thank you for completing the 10 tasks. Next, I will ask you to do the same task for **15 minutes.** The timer starts when you click the Next button on this page. The **number of tasks you do within the 15 minutes will be your production.** You will earn 10 tokens per task correctly done. The resulting amount will be your **income to distribute in the second part.** You will only receive information about your exact production at the beginning of the second part.

Please click on the Next button to proceed.

Next

End of first part

This is the end of the first part. Thank you for participating!

Tomorrow you will receive an invitation to the second part, which will be open between 6 AM and midnight (Pacific Time).

Please click HERE to **go back to Prolific, proving that you completed this part of the study.** Alternatively, you can go back to Prolific, and manually enter the following **completion code:**

A.11.2 Redistribution – Difficulty experiment, spectators

Instructions for Part 2

Your performance and more information about the first part

In the first part you performed 45 tasks, therefore your earnings from the first part are 450 tokens.

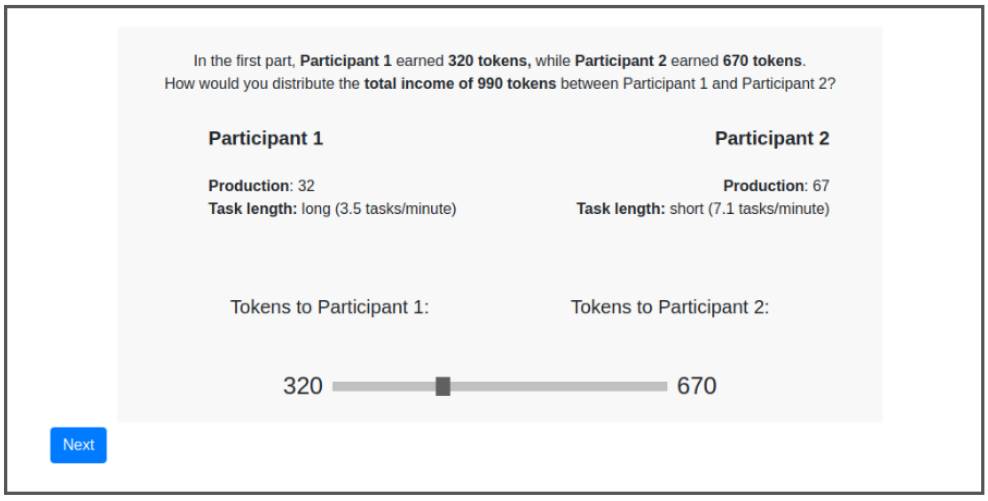
During the first part, you had to work on **3-letter tasks**. However, only about a third of the other participants had 3-letter tasks as well, another **third of participants had to work on 2-letter and the other third on 4-letter tasks**. The task length was **randomly selected** at the beginning of the study for each participant.

At the beginning of the first part, you had to do 10 tasks, as fast as you could. Based on how fast all participants solved the 10 tasks, the computer calculated **how many tasks participants were able to solve within a minute** on average in each task length group. These are **6.2 tasks per minute for 2-letter tasks, 4.5 for 3-letter tasks, and 3.2 for 4-letter tasks**.

Instructions for the second part

For this part, the computer **paired you randomly with another participant** who also completed the first part yesterday. Besides assigning you to another participant, the computer assigned **another pair of participants** from the first part to your group of two. On the following pages, you will have to **decide how to distribute the joint earnings of these two other participants**. Your partner will also make distribution decisions over the joint earnings of these two other participants. Consequently, another pair of participants will distribute your and your partner's earnings.

For making the decisions you will learn both participants' **earnings** and **production** from the first part, and both participants' **task length**. You will also be reminded about **the average number of tasks per minute** participants in the task length group were able to do **when doing the 10 tasks,** so how easy or difficult it was for both participants to solve the tasks. This is an example of the decision screen (the numbers in the screenshot are hypothetical):



When you arrive to the decision page, **you have to click on the slider to reveal the handle.** If you move the handle, you see on the left and the right of the slider the amounts to the first and the second participant that correspond to the current position of the handle. You will be able to move the handle in 5-token increments. Note that if you want to e.g., **increase the payment to the participant on the left**, you have to move the **handle to the right.** This way the number of tokens given to the participant on the left increases, and the number of tokens to the participant on the right decreases. To make a decision, move the handle to the position of the chosen allocation, and click on the Next button.

You will make **10 decisions** in total, of which only **one corresponds to distribution between the actual two participants assigned to you,** and the others correspond to potential other pairings of participants from the first session that did not take place. We cannot tell you which one is the actual decision, so please consider each decision as it were to be implemented at the end of the study.

After the decisions, I will ask you to fill in a short survey and then the study will end.

After every participant finishes the second part, the computer will select **one decision for each pair randomly from the two decisions made for that pair.** That allocation decision will be implemented, so your final payment will be the income given to you in the allocation decision **made by one of the members of another pair.** Income will be converted to British Pounds with 250 tokens = £1.00. You will learn your final payment when you receive all payments through the Prolific payment system, within three days.

By clicking the Next button you will proceed to the decisions.

Next

A.11.3 Redistribution – Difficulty experiment, stakeholders

Instructions for Part 2

Your performance and more information about the first part

In the first part you performed 45 tasks, therefore your earnings from the first part are 450 tokens.

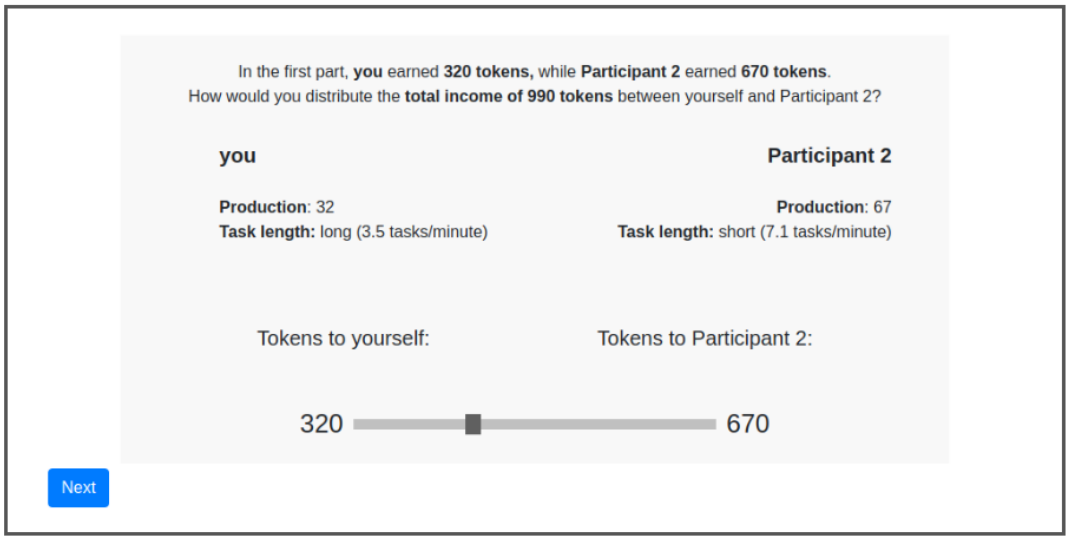
During the first part, you had to work on **4-letter tasks**. However, only about a third of the other participants had 4-letter tasks as well, another **third of participants had to work on 2-letter and the other third on 3-letter tasks**. The task length was **randomly selected** at the beginning of the study for each participant.

At the beginning of the first part, you had to do 10 tasks, as fast as you could. Based on how fast all participants solved the 10 tasks, the computer calculated **how many tasks participants were able to solve within a minute** on average in each task length group. These are **6.2 tasks per minute for 2-letter tasks, 4.5 for 3-letter tasks, and 3.2 for 4-letter tasks**.

Instructions for the second part

For this part, the computer **paired you randomly with another participant** who also completed the first part yesterday. On the following pages, you will have to **decide how to distribute the joint earnings of you and your partner**. Your partner will also make distribution decisions over your joint earnings.

For making the decisions you will learn your partner's **earnings** and **production** from the first part, and your and your partner's **task length**. You will also be reminded about **the average number of tasks per minute** participants in the task length group were able to do **when doing the 10 tasks,** so how easy or difficult it was for you and your partner to solve the tasks. This is an example of the decision screen (the numbers in the screenshot are hypothetical):



When you arrive to the decision page, **you have to click on the slider to reveal the handle.** If you move the handle, you see on the left and the right of the slider the amounts to you and to your partner that correspond to the current position of the handle. You will be able to move the handle in 5-token increments. Note that if you want to e.g., **increase the payment to the participant on the left**, you have to move the **handle to the right.** This way the number of tokens given to the participant on the left increases, and the number of tokens to the participant on the right decreases. To make a decision, move the handle to the position of the chosen allocation, and click on the Next button.

You will make **10 decisions** in total, of which only **one corresponds to distribution between you and your actual partner,** and the others correspond to potential other pairings of participants from the first session that did not take place. We cannot tell you which one is the actual decision, so please consider each decision as it were to be implemented at the end of the study. Note that if in the particular decision you produced less than your partner, your data will be on the **left-hand side**, while if you produced more than your partner, your data will be on the **right-hand side** on the screen.

After the decisions, I will ask you to fill in a short survey and then the study will end.

After every participant finishes the second part, the computer will select **one decision for each pair randomly from the two decisions made for that pair.** That allocation decision will be implemented, so your final payment will be the income given to you in the allocation decision **made by you or your partner.** Income will be converted to British Pounds with 250 tokens = £1.00. You will learn your final payment when you receive all payments through the Prolific payment system, within three days.

By clicking the Next button you will proceed to the decisions.

Next

A.11.4 Redistribution – Ability experiment, spectators

Instructions for Part 2

Your performance and more information about the first part

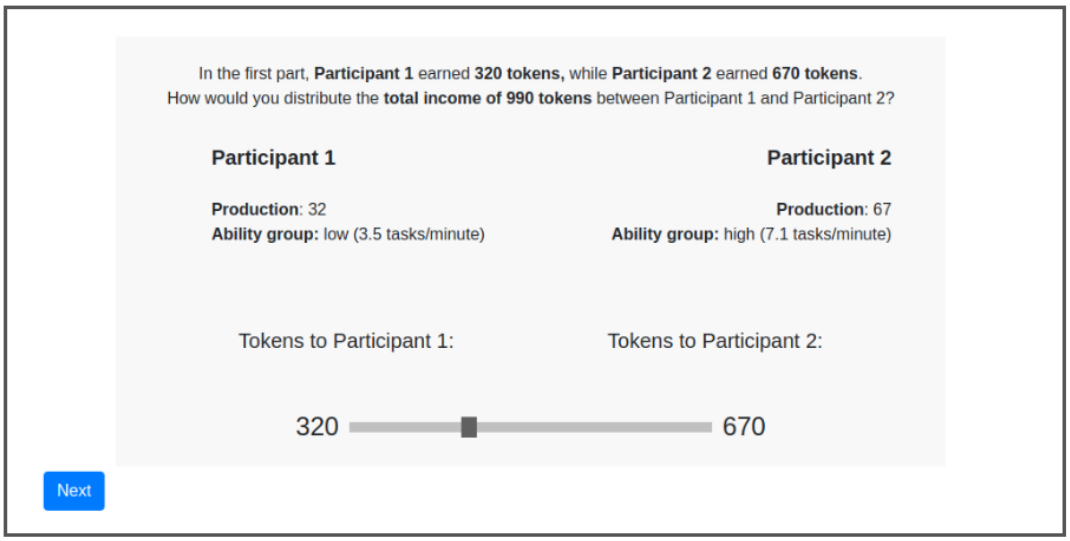
In the first part you performed 45 tasks, therefore your earnings from the first part are 450 tokens.

At the beginning of the first part, you had to do 10 tasks, as fast as you could. Based on how fast all participants solved the 10 tasks, the computer created **three ability groups** that measure **how good participants are in this particular task**. From the time it took for the 10 tasks it also calculated **how many tasks participants were able to solve within a minute** on average in each ability group. These are **3.2 for the low ability group, 4.5 for the medium, and 6.2 for the high ability group.** Based on how fast you solved the 10 tasks, you fell into the third, high ability group in this task.

Instructions for the second part

For this part, the computer **paired you randomly with another participant** who also completed the first part yesterday. Besides assigning you to another participant, the computer assigned **another pair of participants** from the first part to your group of two. On the following pages, you will have to **decide how to distribute the joint earnings of these two other participants**. Your partner will also make distribution decisions over the joint earnings of these two other participants. Consequently, another pair of participants will distribute your and your partner's earnings.

For making the decisions you will learn both participants' **earnings** and **production** from the first part, and both participants' **ability group**. You will also be reminded about **the average number of tasks per minute** participants in the ability group were able to do **when doing the 10 tasks,** so how easy or difficult it was for both participants to solve the tasks. This is an example of the decision screen (the numbers in the screenshot are hypothetical):



When you arrive to the decision page, **you have to click on the slider to reveal the handle.** If you move the handle, you see on the left and the right of the slider the amounts to the first and the second participant that correspond to the current position of the handle. You will be able to move the handle in 5-token increments. Note that if you want to e.g., **increase the payment to the participant on the left**, you have to move the **handle to the right.** This way the number of tokens given to the participant on the left increases, and the number of tokens to the participant on the right decreases. To make a decision, move the handle to the position of the chosen allocation, and click on the Next button.

You will make **10 decisions** in total, of which only **one corresponds to distribution between the actual two participants assigned to you,** and the others correspond to potential other pairings of participants from the first session that did not take place. We cannot tell you which one is the actual decision, so please consider each decision as it were to be implemented at the end of the study.

After the decisions, I will ask you to fill in a short survey and then the study will end.

After every participant finishes the second part, the computer will select **one decision for each pair randomly from the two decisions made for that pair.** That allocation decision will be implemented, so your final payment will be the income given to you in the allocation decision **made by one of the members of another pair.** Income will be converted to British Pounds with 250 tokens = £1.00. You will learn your final payment when you receive all payments through the Prolific payment system, within three days.

By clicking the Next button you will proceed to the decisions.

Next

A.11.5 Redistribution – Ability experiment, stakeholders

Instructions for Part 2

Your performance and more information about the first part

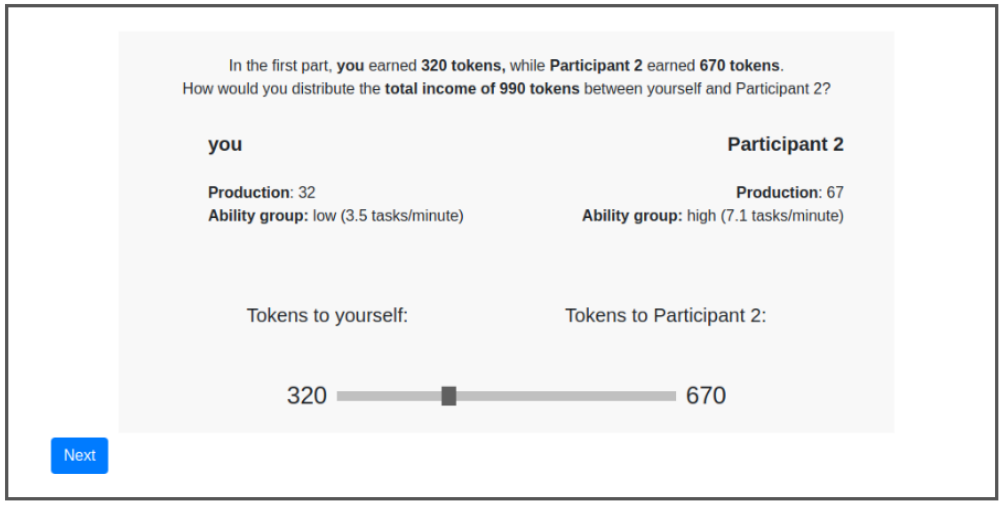
In the first part you performed 45 tasks, therefore your earnings from the first part are 450 tokens.

At the beginning of the first part, you had to do 10 tasks, as fast as you could. Based on how fast all participants solved the 10 tasks, the computer created **three ability groups** that measure **how good participants are in this particular task**. From the time it took for the 10 tasks it also calculated **how many tasks participants were able to solve within a minute** on average in each ability group. These are **3.2 for the low ability group, 4.5 for the medium, and 6.2 for the high ability group.** Based on how fast you solved the 10 tasks, you fell into the third, high ability group in this task.

Instructions for the second part

For this part, the computer **paired you randomly with another participant** who also completed the first part yesterday. On the following pages, you will have to **decide how to distribute the joint earnings of you and your partner**. Your partner will also make distribution decisions over your joint earnings.

For making the decisions you will learn your partner's **earnings** and **production** from the first part, and your and your partner's **ability group**. You will also be reminded about **the average number of tasks per minute** participants in the ability group were able to do **when doing the 10 tasks,** so how easy or difficult it was for you and your partner to solve the tasks. This is an example of the decision screen (the numbers in the screenshot are hypothetical):



When you arrive to the decision page, **you have to click on the slider to reveal the handle.** If you move the handle, you see on the left and the right of the slider the amounts to you and to your partner that correspond to the current position of the handle. You will be able to move the handle in 5-token increments. Note that if you want to e.g., **increase the payment to the participant on the left**, you have to move the **handle to the right.** This way the number of tokens given to the participant on the left increases, and the number of tokens to the participant on the right decreases. To make a decision, move the handle to the position of the chosen allocation, and click on the Next button.

You will make **10 decisions** in total, of which only **one corresponds to distribution between you and your actual partner,** and the others correspond to potential other pairings of participants from the first session that did not take place. We cannot tell you which one is the actual decision, so please consider each decision as it were to be implemented at the end of the study. Note that if in the particular decision you produced less than your partner, your data will be on the **left-hand side**, while if you produced more than your partner, your data will be on the **right-hand side** on the screen.

After the decisions, I will ask you to fill in a short survey and then the study will end.

After every participant finishes the second part, the computer will select **one decision for each pair randomly from the two decisions made for that pair.** That allocation decision will be implemented, so your final payment will be the income given to you in the allocation decision **made by you or your partner.** Income will be converted to British Pounds with 250 tokens = £1.00. You will learn your final payment when you receive all payments through the Prolific payment system, within three days.

By clicking the Next button you will proceed to the decisions.

Next